AN EXAMINATION OF THE PROOF AND ARGUMENTATION SKILLS

OF EIGHTH-GRADE STUDENTS

MELEK PESEN

BOĞAZİÇİ UNIVERSITY

2018

AN EXAMINATION OF THE PROOF AND ARGUMENTATION SKILLS

OF EIGHTH-GRADE STUDENTS

Thesis submitted to the

Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Primary Education

by

Melek Pesen

Boğaziçi University

2018

DECLARATION OF ORIGINALITY

I, Melek Pesen, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature..

Date ....06.08.2018..

ABSTRACT

An Examination of the Proof and Argumentation Skills of Eighth-Grade Students

This study examined eighth-grade students' proof and argumentation skills and their relationships. The study was a mixed-method study with descriptive, statistical and qualitative analyses. It was conducted with two hundred and forty-two students in eighth- grade. According to the findings, students mostly constructed and appreciated empirical proofs in the algebra tasks. They could not produce any type of proof; but, they preferred analytical proof response in the geometry task. Findings from the argumentation tasks revealed that students mostly produced level 2 arguments, which contained a claim and evidence. Statistical analyses showed that there exists a significant relationship between proof and argumentation skills. Students performed better in proof evaluation part than in poof construction part. There were no gender differences in students' mathematics achievement, proof and argumentation skills. Qualitative findings showed that the students' performances in proof construction tasks were affected by their content knowledge. It was found that students' reasoning for the most convincing proof varied and was compatible with their proof schemes when evaluating proofs. It was found that students' content knowledge, misconceptions and the way they used evidences shaped their argumentation levels. Findings of this study are important contributions in presenting evidences for the relationship between proof and argumentation skills and in revealing specified information about students' proof and argumentation performances.

ÖZET

Sekizinci Sınıf Öğrencilerin İspat ve Argümantasyon Becerilerinin İncelenmesi

Bu çalışmada, sekizinci sınıf öğrencilerinin ispat ve argümantasyon becerileri ve bu becerilerin arasındaki ilişkiler ele alınmıştır. Betimsel, istatistiksel ve nitel analizleri barındıran bir karma yöntem araştırması olan bu çalışma iki yüz kırk iki sekizinci sınıf öğrencisiyle yürütülmüştür. Araştırmanın bulguları, öğrencilerin cebir alıştırmalarında çoğunlukla deneysel ispatlar oluşturup, bu ispatı en ikna edici olarak değerlendirdiklerini göstermektedir. Geometri alıştırmalarında ise öğrencilerin çoğunluğu herhangi bir ispat üretemeyip, analitik ispat türünü daha çok ikna edici olarak bulmuşlardır. Argümantasyon çalışmasının bulgularına göre en çok oluşturulan argümantasyon seviyesinde öğrenciler bir iddia ve gerekçelendirme sunmuşlardır. İstatistiksel analizler ispat ve argümantasyon becerileri arasında anlamlı ilişkiler olduğunu, ispat değerlendirme performansının ispat oluşturma performansından daha iyi olduğunu ve öğrencilerin performanslarında cinsiyet farklılıkları oluşmadığını göstermiştir. Nitel bulgular ispat yapma becerilerinin öğrencilerin içerik bilgilerinden etkilendiğini, ikna edici ispatlar için gösterdikleri gerekçelendirmelerinin ispat değerlendirme şemalarıyla uyum gösterdiğini ve argümantasyon becerilerinin öğrencilerin içerik bilgilerinden, kavram yanılgılarından ve delilleri kullanma biçimlerinden etkilendiğini ortaya koymuştur. Bu çalışmanın bulgularının ispat ve argümantasyon becerilerinin detaylarına ve arasındaki ilişkiye dair bilgi vermesi açısından katkı sunabileceği düşünülmektedir.

ACKNOWLEDGEMENTS

I would first like to express special thanks to my thesis advisor, Prof. Emine Erktin, for her advices, support and encouragement during my master's study. Her directions, feedbacks and constructive criticisms kept my motivation high.

I would also thank my thesis committee Assoc. Prof. Ebru Kaya and Assist. Prof. Zeynep Çiğdem Özcan for their constructive feedbacks and directions. I am very grateful to Assist. Prof. Yeşim İmamoğlu and Assist. Prof. Devrim Güven for their valuable contributions.

I am also thankful to my friends in Boğaziçi University. They provided me support whenever I needed. Among my friends, I would like to thank Büşra Aksöz and Samet Sarıgöl for their help during my data collection. Also, I wish to thank Fatma Acar for her valuable support during my master thesis.

Last but not least, I would like to express my thanks to my family. They have always believed in me and supported me emotionally.

Dedicated to my grandparents,

I always feel their prayers and love…

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Mathematics education does not only intend learners to become proficient in computing and calculations; it aims learners to reason, explore, conjecture and justify both in academic settings and in real life situations (Rumsey & Langrall, 2016). Reasoning through justification is emphasized as an important curricular aim in Turkish mathematics education curricula (Appendix A) and mathematics education organizations. Each year, more and more reasoning and justification tasks and items are included in international comparison exams. The Turkish mathematics curriculum for primary school (MEB, 2013) emphasizes certain basic skills for students such as mathematical process skills, which contain communication, reasoning and association.

Reasoning can be defined as the knowledge acquisition process through using the information at hand as well as thinking techniques such induction, deduction, comparison, and generalization (Heinze & Reiss, 2003). Reasoning skills facilitate academic performance and also performance out of school. Hence the learning environments should be designed taking reasoning skills into consideration. Some of the indicators of reasoning skills in the curriculum are:

- Defending the correctness and validity of inferences
- Making logical generalizations and inferences
- Explaining and using the mathematical relationship while analyzing a mathematical situation (MEB, 2013).

Similar aims exist in the math education worldwide. For example, in the USA the National Council of Teachers of Mathematics (NCTM) set certain academic standards in mathematics to define the required knowledge and skills to be successful in the future. These are both content standards and process standards. One of the process standards of NCTM is labeled as "Reasoning and Proof". According to this standard:

Instructional programs from prekindergarten through grade 12 should enable each student to:

- Recognize reasoning and proof as fundamental aspects of mathematics
- Make and investigate mathematical conjectures
- Develop and evaluate mathematical arguments and proofs
- Select and use various types of reasoning and methods of proof

(NCTM, 2000)

The concepts of proof and argumentation are emphasized in the national curricula and organizations, explicitly or implicitly. However, despite their existence and emphases, there are evidences from the literature and from findings of the international comparison exams that students at all grade levels perform poorly in proof and argumentation related tasks (Bieda & Lepak, 2014). The results of many international exams in mathematics showed that Turkish students' performances are below average. Trends in Mathematics and Science Society (TIMSS) is one of the international exams which looks into mathematical competencies in three cognitive domains: Knowing, applying, and reasoning. The reasoning domain constitutes 20% of fourth grade mathematics test and 25% of eighth grade mathematics test. In the analyses of results for Turkish students, it was observed that for fourth grades average reasoning scores were significantly lower compared to their average overall mathematics scores. In the past four years, the reasoning-average scale score showed an increase, but it was not statistically significant. Also, it was observed that for eighth grades, Turkish students' reasoning scores were significantly higher than their overall mathematics scores (Mullis, Martin, Foy, & Hooper, 2016).

In 2011, TIMSS identified four different mathematical competence levels. For eight grade students "Advanced level" was defined as "Students can reason with knowledge, they can make inferences, make generalizations and solve linear inequalities.". In 2011, 7% of the Turkish students were placed in "Advanced Level", in 2015, 6% of the Turkish students fell into that category (Polat, Gönen, Parlak, Yıldırım, & Özgürlük, 2016).

Program for International Student Assessment (PISA) studies show and compare the mathematical literacy of students in different countries. Mathematical literacy is basically expressed as competencies in use of knowledge and skill, in analysis, in making logical inferences and in establishing efficient communication while identifying, interpreting and solving problems (Taş, Arıcı, Ozarkan & Özgürlük, 2016). It was analyzed that the performances of Turkish students in 2015 was lower than PISA 2009 and PISA 2012 (Taş et al., 2016).

Argumentation refers to the process and product of presenting high-quality arguments individually and often collaboratively (Rumsey & Langrall, 2016). It has been studied in a lot of disciplines, especially in science education for asking students to produce scientific arguments that are well developed (Erduran & Jiménez-Aleixandre, 2008). Dealing with argumentation in mathematics is meaningful for letting learners to construct sound mathematical arguments. Studying argumentation in class both provide better conditions for conceptual understanding and opportunities for producing and distinguishing good arguments in all settings (Staples & Newton, 2016). Argumentation involves acts of validations, justification processes as proof does. Hence, there exist some views about the possible relationship between argumentation and proof skills of students in some studies (Conner, Singletary, Smith, Wagner, & Francisco, 2014; Reiss, Heinze, Renkl, & Groß, 2008; Wood, 1999). Looking whether there exists a relationship between them could be informatory and directive in organizing and implementing mathematics lessons. In other words, when, how, and how often these skills should be presented to students can be decided through examining these skills starting from analyzing their relationships.

The results of TIMSS and PISA reveal that even though in Turkish mathematics curriculum and in NCTM process standards which seem to have influenced the Turkish mathematics curriculum, reasoning, proof, and mathematical argumentation are given importance, Turkish students' performances in these domains are not at a desired level. Students' underachievement in reasoning and proof need to be investigated carefully and their reasoning, proving and argumentation skills should be analyzed to understand why they are not performing well. Also, there is a need for empirical studies which aim to see whether there exists a relationship between these two skills. A meaningful relationship can be important to study further on this issue by looking into cause-effect relationships. By this way, students' proving skill can be developed through the application of argumentation practices, which are more accessible to students and to mathematics teachers due to their less structured and rigorous nature (Reid & Knipping, 2010). On the other hand, students' argumentation skills can be improved through proof instructions so that mathematical practices effect argumentation practices which are required and appear not only in academic settings but also in real life situations .

By considering the importance of proof and argumentation in mathematics education, which are emphasized in research studies, middle school mathematics curricula and international exams, this study aims to examine proof and argumentation skills of eight-grade students and the relationships between these skills and to reveal eight-grade students' proof and argumentation practices in general. Studying with eight-grade students is preferred rather than students from other middle school grades by the researcher. The reasons for this selection mainly stem from two points. Firstly, eight-grade is the terminal year in Turkish middle schools. Students would bring their proof and argumentation skills from middle school to high school in which there is more emphasis on proving in the Turkish high school mathematics curriculum (Liu, Tague, Somayajulu, 2016; Piaget, 1985). Second reason for selecting eighth-grade students is that eighth-grade is more likely to be the year in which students can use deductive reasoning and transfer from concrete operational stage to formal operational stage (Piaget, 1985) according to Piaget's cognitive developmental phases.

The details of this study are reported in the following manner: In chapter 2, review of the literature about proof and argumentation in education is presented. In chapter 3, statement of the problem is explained. In chapter 4, the research questions are given, and operational definitions of the variables are provided. Methodology of the study is explained by informing about characteristics of the sample, details of the data collection instrument and the procedure about how the data is collected in chapter 5. In chapter 6, analysis of the data and the results of these analyses are shared. In chapter 7, conclusion, implications, limitations of the study and suggestions for further studies are presented. The instrument of the study, the rubrics of items, original student responses in written documents, correlation matrix for all variables, corresponding objectives of tasks in Turkish mathematics and science education curricula and contents related to proof and argumentation in Turkish mathematics education curriculum are placed in the appendices.

CHAPTER 2

LITERATURE REVIEW

In this section the focus is on the relevant literature about proof and argumentation skills and the relationship between them. The review of the literature covers definitions and functions of proof, proof schemes, proving abilities of students and their development, definitions and functions of arguments and argumentation, argumentation in different disciplines, structure and analysis of arguments, relationship between proving and argumentation and suggestions for developing proof and argumentation skills, respectively.

2.1 Definitions and functions of proof

Proof is an important component of mathematics and it is regarded as a vital skill for mathematics which distinguishes mathematics from other disciplines (Demiray & Işıksal - Bostan, 2017; Heinze & Reiss, 2003). Schoenfeld (1994) asserted that proof is not separable from mathematics (Knuth, Choppin, Slaughter, & Sutherland, 2002). Proofs are defined as conceptual syntactic derivations with specific technical approaches. Through application of logical inferences, each sentence is formed and demonstrated from previous axiom and the immediate consequences are obtained from preceding sentences while proving (Hanna, 2000). Accordingly, the structure of proofs is characterized by three variables: (1) the statement which will be proved, (2) axioms or previously proven statements that are used in the proving process, and (3) inference rules which are used in the process of proving (Csikos, 1999). Furthermore, proving ability is defined as the ability to make something evident and also to construct proofs (Csikos, 1999).

Proving can be conducted in two ways: First, the truth of a statement can be demonstrated. Second, the reason for the truth of a statement can be demonstrated to get insight why it is true. (Altıparmak & Öziş, 2005; Knipping, 2003; Reid & Knipping, 2010). Different disciplines have different criteria for accepting an explanation as a proof. So, what counts as proof differs in science, formal logic, mathematics etc. (Reid, 2005). Mathematics is perceived as a proving discipline (Heinze & Reiss, 2003). Formal proof is a widely accepted form of proof. Mathematical community uses a theoretical construct of a formal proof for evaluating a proof. They start from a 'real' proof and approach the formal proof by adding information from general knowledge until they are convinced that the real proof is correct (Heinze & Reiss, 2003).

The concept of proof begins in the preschool period (Aktaş, 2002; Altıparmak & Öziş, 2005) as classifying, matching, ordering and comparing activities, which create a foundation for proving skills. Piaget (1985) classifies this period as a transition period to logical thinking and named it as intuitive stage (Altıparmak & Öziş, 2005; Aylar, 2014). In preschool period, the proof concept is not understood as what we know as formal proof. The activities which promote informal proof skills which is important for building cause-effect chains in later years can be carried out during preschool ages.

In primary school period, concrete thinking is dominant for the first five years. In the following three years, abstract thinking begins to develop. In general, primary school period is the time for development of logical thinking (Altıparmak & Öziş, 2005; Piaget, 1985). According to NCTM (2000), by the end of the primary school students are expected to (1) develop and evaluate mathematical statements and proofs, and (2) select and use different logical thinking strategies and proof types. Students in 3-5 grades should know that a couple of examples is not enough to support a claim. Also, students in these ages should use counterexamples for falsifying a claim.

Students in 6-8 grades should be able to make generalizations from claims and evaluate claims. They can use deductive and inductive reasoning. It is stated that when the students in concrete operational stage at primary school period acquire the required skills, they won't be having trouble in proof producing in the formal operational stage. According to Altparmak and Öziş (2005), primary school is the period in which concept of mathematical argument forms. In middle school period, abstract thought develops, and students need to use induction and deduction methods to test mathematical arguments, they need to form examples for incorrect expressions-counterexamples. They need to be familiar to use symbolic language in mathematics and they need to be encouraged for the use of deduction in this period (Aylar, 2014).

By the time of high school, which is classified as formal operational stage, students can comprehend direct proof, contrapositive proof, proof by contradiction, induction and proof by geometrical shapes (Altıparmak & Öziş, 2005). It is stated that proof must be a central part of the curriculum for all grade levels. From preschool to grade 12 all children should develop and evaluate conjectures, arguments and proofs in mathematics (Ellis, Lockwood, Williams, Dogan, & Knuth, 2013).

In mathematics education, the role of proof is providing justifications and promoting mathematical understanding (Hanna, 2000). Proof has five main functions and goals (de Villiers, 1990; Hanna, 2000; Mejia-Ramos & Inglis, 2008):

- Verification: Proofs can establish the truth of a statement.
- Explanation: Proofs can show us why a statement is true.
- Systematization: Proofs can organize final statements in deductive system.
- Discovery: Proofs can provide us an opportunity to invent new knowledge.
- Communication: Proofs can establish the transmission of mathematical knowledge.

When considering the role of verification, Duval (2007) claims that a statement can be true or false logically. Psychologically, it may take on many values, which is described as its epistemic value. This term means a personal judgment of about how the proposition is believed. Mathematically "true" statements have to in a quite narrow range of epistemic values, whereas "true" scientific facts can fall into a wider range of epistemic values (Duval, 2007; Reid & Knipping, 2010)

For explanation role of proof, it is stated that not all proofs can fulfill the role of explanation even though all of them should meet the verification role (Reid& Knipping, 2010). Hanna (2000) stated that a proof that explains is precious. Middle school students seek for clear explanations when they meet proofs. They give importance to the explanation role of proofs (Bieda & Lepak, 2014).

The communication role of proof is explained as a vehicle for presentation of products. In other words, proving is defined as a communicative act (Carrascal, 2015). Also, proving is expressed as an interactive process in that students interact with their teachers (Ko, 2010; Sen & Güler, 2015).

Beside these roles, some other roles of proof are also explained as aesthetics, intellectual challenge, construction of empirical theory, clarifying a definition or some assumptions and incorporating a fact into a framework (Reid & Knipping, 2010).

In mathematics education, proof can be used in providing justifications and promoting mathematical understanding (Hanna, 2000). Proof facilitates conceptual knowledge construction and leads to meaningful learning. It prevents the memorization of mathematical facts (Aylar, 2014; Sen & Güler, 2015). Also, proof writing can foster comprehension of students and enhance the development of deductive reasoning, critical thinking (Cyr, 2011). Proving not only explains why a statement is true but also its product can be used for the further and following investigations (Bieda & Lepak, 2014).

Fawcett argued that the study of proof in mathematics has an effect on students' abilities in critical thinking for other domains (as cited in Reid, 2005, p.459). On the other hand, Healy and Hoyles (2000) claimed that it is not possible to transfer the method of proof learned in mathematics to other domains. The motivation for teaching proof is to understand the nature of mathematics in a better way (Reid, 2005, p.460). Reid (2005) suggests that showing students proving and the limits of mathematics can make them more critical in the use of numerical arguments in other domains.

Overall, it can be summarized as proof is an inseparable part of mathematics (Schoenfeld, 1994) and it should appear in mathematics education curricula from preschool to grade 12 level (Ellis et al., 2013). Proof serves a lot of roles but their applicability in other domains is controversial (Healy & Hoyles, 2000; Reid, 2005). In the next section, hierarchical proof schemes frameworks of students are presented.

2.2 Proof schemes

Proof skill refers to the ability to conduct and read proofs (Csikos, 1999). Students' proving abilities were studied and it was asserted that there are different kinds of proving abilities in which from intuitive to mathematically sophisticated hierarchical relationship between these abilities is present (Bieda & Lepak, 2014). Since, certain inference rules are harder for using and certain patterns of solutions are much more advanced, there exist a hierarchy among students' proving abilities (Bieda & Lepak, 2014; Csikos, 1999). The most commonly used proof scheme by researchers was developed by Harel and Sowder (1998). Most of the time, the word "proof" is used and accepted as it refers to arguments generated by the middle school students. That is why, Stylianides and Stylianides (2009) view these schemes as justification schemes rather than as proof schemes (Sen & Güler, 2015). These schemes are labeled as external conviction, empirical proof and analytical proof. Each scheme also involves subcategories. It is noted that these schemes are not mutually exclusive; people can have more than one kind of scheme (Harel & Sowder, 1998).

In external conviction, students only memorize prescriptions, no discovery or creativity is involved and most of the time an authority is present as the only source of knowledge. External conviction proof scheme involves three subcategories. The first one is ritual proof scheme in which judgements of an argument is based on its appearance rather than its content. The second one is authoritarian proof scheme in which proof is based on another student, the teacher or the textbook, namely on an authority. The third one is symbolic proof scheme in which symbolic representations are used without knowing the meaning and function of symbols (Harel & Sowder, 1998; Sen & Güler, 2015).

Empirical proofs are constructed by students' intuitions. Students form their arguments by appealing to some physical facts or sensory experiences (Harel & Sowder, 1998). It is divided into two subcategories as inductive and perceptual proof. In inductive proof scheme, it is observed that students get benefit from quantitative evaluations like examples and manipulations with trials. Through one or more example(s) generalizations are made (Harel & Sowder, 1998; Sen & Güler, 2015). In perceptual proof scheme, perceptional representations and rudimentary mental images are used. These representations lack the ability to transform the results and they are case dependent-unique to the context so that generalizations are made by a unique representation (Harel & Sowder, 1998; Sen & Güler, 2015).

Analytical proofs are conducted by the use of logical deductions. It has two subcategories: Transformational proof scheme and axiomatic proof scheme. In transformational proof scheme, there are some operations on objects and transformation of images by means of deduction. In axiomatic proof scheme, students comprehend that a mathematical justification must start with axioms and theorems (Harel & Sowder, 1998; Sen & Güler, 2015).

Harel and Sowder's proof scheme framework involves also some further subcategories of transformational and axiomatic proof schemes which are sometimes not included in the studies that analyze proving skills. They will not be included in this study either even though they appear on the original categorization developed by Harel and Sowder in Figure 1 .

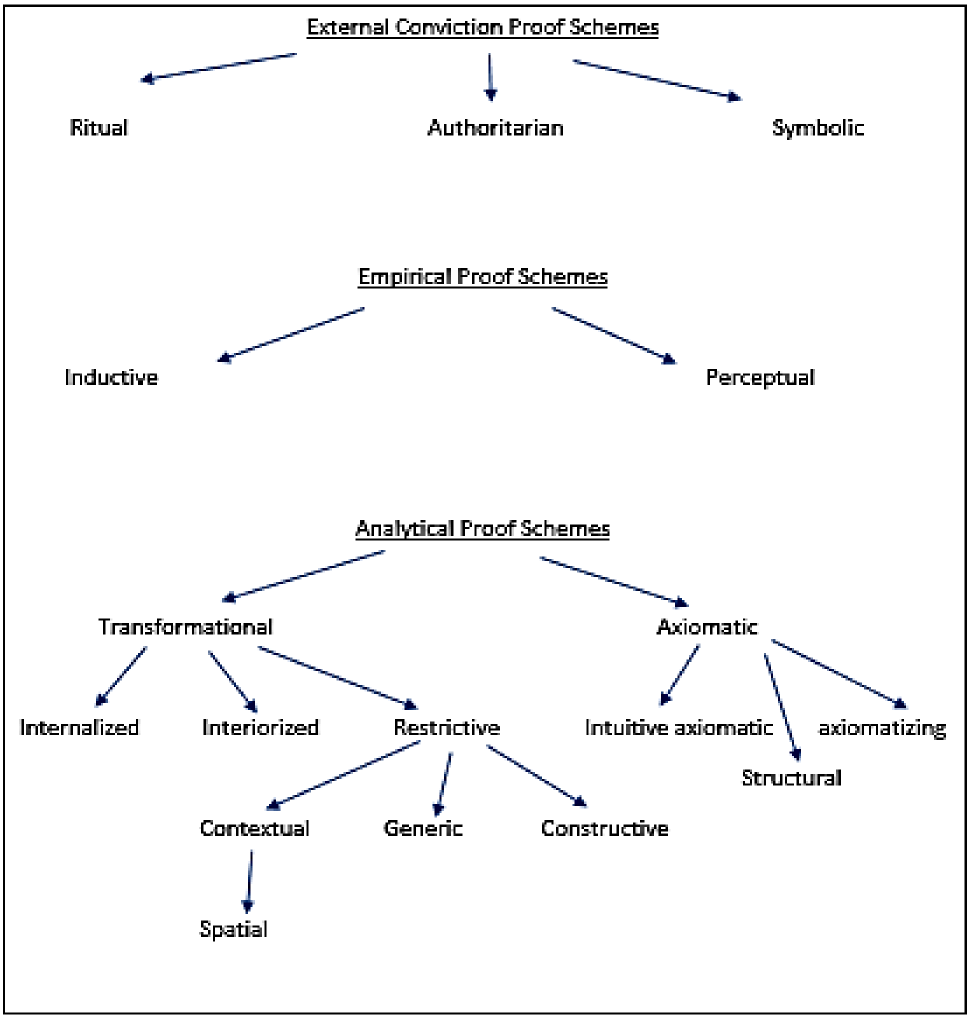


Fig. 1 Proof schemes framework by Harel and Sowder (1998)

In the literature, there are studies which aim to analyze proving ability by classification of proving skills. One of them was developed by Balacheff (1998) according to whom there exist two general categories of proofs generated by students: pragmatic proofs and conceptual proofs. Most commonly produced proofs by students fall into the category of pragmatic proofs, which may take two forms as naïve empiricism and crucial experiment. In naïve empiricism, students simply assert that a statement is true because it works with one or several examples without considering why the selected examples suggest that it holds for all possible members of the claim. In a crucial experiment, there is an intentional selection of a case which is accepted as a representative of other cases. Students' claim is that if this case can provide the truth of statement, then the statement is true for all other cases. Conceptual proofs, on the other hand, can take two different forms as generic example or demonstration. Use of a generic example is an empirically based approach in which operations on the examples can explain why the statement to be proven is true. In demonstrations, students apply a strategy which is more rigorous than generic example. Definitions, theorems and deductive rules are applied when proving statements (Balacheff, 1998; Bieda & Lepak, 2014).

Waring (2000) developed a proof concept development scheme that consists of six levels. In level 0, students do not care about the existence of and the need for proof. In level 1, students are aware the notion of proof, but they only check for a few examples. In level 2, students are aware of the fact that checking a few cases is not enough, so they can either check also extreme cases or use generic examples. In level 3, students are aware the need for a general argument, but they cannot produce them. They can follow a short chain of deductive reasoning. In level 4, students are aware of the need for a general statement and they can produce such arguments only in familiar contexts. In level 5, students are aware of the need for a general statement and they produce such arguments in both familiar and unfamiliar contexts (Knuth et al., 2002).

Overall, there exist frameworks for differentiating students' different proof skills in the literature (Bieda & Lepak, 2014). In this study, Harel and Sowder's (1998) proof scheme was used to detect students' proof skills. This framework was preferred for two reasons. First, it is the commonly used framework in the literature. Second, it is easier to detect students' proof schemes than other frameworks. In the next section, students' proving skills and their developments are explained.

2.3 Proving skills of students and their development

Proof is a central and important activity, but it is also viewed as a very difficult and complex practice (Mejía-Ramos & Inglis, 2009; Ubuz, Dincer & Bulbul, 2012). Proving depends on certain criteria and, students have difficulties in doing proofs (Demiray & Işıksal - Bostan, 2017). Many of them fail to understand what counts as verification and evidence (Cooper et al., 2011). Students from almost all grade levels have difficulties while generating valid proofs (Bieda & Lepak, 2014; Harel & Sowder, 1998; Weber, 2001). They need to have conceptual understanding of rules, theorems, techniques and the knowledge of the nature of proof (Gibson, 1998). There is a certain amount of required knowledge, beliefs, cognitive skills and social environment associated with reading and conducting proofs (Blanton, Stylianou, & David, 2003). Parallel to these views, when students' proofs were analyzed, it was observed that knowledge of concepts and theorems was not adequate for performing proofs in mathematics (Heinze & Reiss, 2003; Sen & Güler, 2015). Moore (1994) claimed that students do not possess the knowledge of definitions or they have difficulty when stating them in an appropriate manner. They have deficiencies in mathematical notations and mathematical language (Carrascal, 2015; Sen & Güler, 2015). They do not know how to start writing proofs. They cannot perform the transition from induction to deduction (Ellis et al., 2013). Another factor associated with difficulty in proving has been explained as emphasis on reasoning, critical thinking and problem solving rather than proof construction in the curriculum (MEB, 2013). Weber (2001) stated that students have inadequate strategic knowledge, so they cannot perform proofs. Also, it is observed that students do not know the procedures of proof. It is stated that declarative knowledge, methodological knowledge, and metacognition are important for proof competence. Methodological knowledge consists of three aspects: proof scheme, proof structure and chain of conclusions. For proof scheme, it is stated that, only deductive arguments are adequate for a mathematical proof. About proof structure, it is said that, a proof starts with premises and ends with a specific assertion. To prove this assertion, all arguments should be valid. For chain of conclusions, it is stated that, all steps in proof can be concluded from the previous ones (Heinze & Reiss, 2003).

The failure of the students in constructing and understanding proofs may be because proofs and proving processes are frequently regarded as isolated topics in mathematics courses (Reid, 2011). That is why, instead of a credible path to form reliable arguments, students perceive proof as a written work in a special form like two-column proofs (Chazan, 1993; Healy & Hoyles, 2000). The recent reform efforts emphasize conceptual understanding of the topics while paying less attention to the format of the proof in order to handle this issue (de Villiers, 1990; Hanna, 2000; Reid, 2011). Reasoning and proof cannot be taught in units; proof is a mathematical method that arises naturally from mathematical inquiry and the need to verify, explore and communicate (de Villiers, 1990; NCTM, 2000).

While analyzing 14- and 15- year-old students' decisions in proof evaluations, Healy and Hoyles (2000) found that students' performances were affected by their apprehension of the purpose of proof, their competencies in mathematics, the instruction that they were exposed to and their gender (de Villiers, 1990).

Students' construction and evaluation of proofs were found to be inconsistent across content areas. Students can produce or value a deductive proof in one area but prefer empirical evidences in another one (Harel & Sowder, 1998; Heally & Hoyles, 2000). These finding are compatible with the developmental models of proof understanding; understanding and producing deductive arguments is not actualized until students reach higher levels. However, relying on different schemes for different contexts is still unexplained (Tall, Yevdokimov, Koichu, Whiteley, Kondratieva, & Cheng, 2011; Waring, 2000).

Liu et al. (2016) asserted that while evaluating proofs of others, arguments which are based on empirical trials with examples, were found to be convincing by a lot of students. Students found numerical and narrative arguments rather than algebraic arguments easier to understand. Some of them evaluated algebraic arguments as clear while others perceived them as complex and confusing. In addition, the clarity of the presented explanation, students' familiarity with the context of the statement, and the complexity of arguments affected student judgments about the arguments.

Sometimes, students intentionally do not view proofs as justifications. Chazan (1993) reached the conclusion that there are some reasons for students' disbelief in deductive proofs as a way of verification: Counterexamples can still exist which are not covered in the proof. The proof might be proving a specific case. The assumptions used while proving can be incorrect (Reid & Knipping, 2010).

Proof-making and justification levels of students can increase in time as their grades increase (Sen & Güler, 2015). Through the years, students shift from visual and narrative methods to algebraic expressions (Cooper et al., 2011; Sen & Güler, 2015). Their understandings of mathematical justifications move from inductive to deductive (Knuth et al., 2002). Students begin to develop understanding of the benefits of proof in time (Cooper et al., 2011).

Teachers' actions to promote proving and justification skills are important. In order to provide students with more sophisticated and rigorous experiences, teachers are suggested to use more mathematically based rather than example-based explanations in their lectures. They should guide students on the use of deductive reasoning rather than caution them to not use examples (Bieda & Lepak, 2014). On the other hand, Reid and Knipping (2010) claim that as students may not be prepared to practice in the field, they should be encouraged to appreciate the products of the field. That is why, maybe at first proof reading rather than proof writing should be the focus of the curriculum.

Overall, students have difficulties while generating proofs (Bieda & Lepak, 2014; Harel & Sowder, 1998) and have deficiencies in mathematical notations and terminology (Carrascal, 2015). While analyzing proofs, they are mostly convinced by numerical arguments (Liu et al., 2016) and their proof constructions and evaluations were found to be inconsistent across content areas (Harel & Sowder, 1998; Healy & Hoyles, 2000). Studies show that proof skills of students can develop over time with proper interventions (Cooper et al., 2011; Sen & Güler, 2015). In the next section, the concepts of argument and argumentation are defined and their functions are explained.

2.4 Definitions and functions of arguments and argumentation

Argumentation is a reasoned discourse that may not be necessarily deductive (Reiss et al., 2008). Reid and Knipping (2010) reported in their study that Perelman mostly associates argumentation with convincing. Toulmin (1958) interprets argumentation as referring to the structure of the argument. Ducrot takes argumentation as the core activity of discourse on grammatical structures (Reid & Knipping, 2010). These different views have led to some possible classification about the meaning of argumentation as: (1) argumentation is what convinces other people, (2) it has a structure which is accepted by the community (3) it exists on grammatical elements and is present in discourses. Also, a lot different perspective shared by researchers about the meaning of argumentation: (1) it is kind of a reasoning, (2) it is a social behavior, (3) it is a process where a logical discourse is obtained at the end and (4) it is a process through which conjectures are given rise (Reid & Knipping, 2010).

Staples and Newton (2016) claim that there are two complementary purposes of argumentation practices: (1) concept development (2) developing proficiency with the practice of argumentation. Argumentation practices are important because they provide support for student thinking to analyze whether a proposed line of reasoning is a viable approach (Staples & Newton, 2016). Argumentation is significant for conceptual understanding because it provides acts of challenge and justification and mental processes are more involved for the resolution of conflicts (Wood, 1999).

Definition of argument also varies like the one for argumentation. An argument is defined as presenting reasons for or against a claim or progress of an event (Güneş, 2013). Bieda and Lepak (2014) define it as a sequence of statements constructed with the intention of convincing others about the validity of a claim. Also, it is defined as justifying a conclusion based on data (Mejia-Ramos & Inglis, 2008). An argument has been associated with argumentation in various ways: (1) arguments can give rise to argumentation, (2) arguments are the result of argumentation, (3) arguments are part of the argumentation, and (4) arguments are identical to argumentation (Reid & Knipping, 2010).

In argumentation studies the audience of persuasion create a difference in the type of argumentation (Cabassut, 2005; Conner et al., 2014). When a person tries to convince a particular audience in an environment in which there are a lot of participants who criticize, justify and evaluate concepts and develop a consensus after opposing perspective, the argumentation is labeled as collective or collaborative argumentation (Conner et al., 2014; Hunter, 2007; Rumsey & Langrall, 2016; Wood, 1999). In individual argumentation, students convince themselves for the truth of a claim. Through individual argumentation, they become intellectually autonomous individuals and develop their dispositions in a field (Yackel & Cobb, 1996).

Through engaging argumentation, students not only establish the truth of a mathematical claim, but they also have an opportunity to deepen their conceptual understanding in mathematics (Staples & Newton, 2016; Carrascal, 2015). Mathematical argumentation skills provide students with taking the ownership of mathematics that they are learning and promote conceptual understanding rather than procedural understanding (Ross, Fisher & Frey, 2009; Rumsey & Langrall, 2016). Conceptual understanding arises from cognitive conflicts and challenges that are the result of students’ distinct ideas. The resolution of conflict in ideas occurs through argumentation practices in mathematics lectures (Staples & Newton, 2016; Wood, 1999). Argumentation can develop competencies related to critical thinking and it may contribute to the intervention of progressive construction of mathematical concepts. Argumentation may be thus decreasing the cognitive load (Carrascal, 2015).

Overall, the concepts of argument and argumentation have varied definitions (Reid & Knipping, 2010) and argumentation provides establishing truthiness of a claim (Staples & Newton, 2016), provides conceptual understanding rather than procedural understanding (Carrascal, 2015; Rumsey & Langrall, 2016). In the next section, argumentation in the mathematics and science education disciplines is elaborated.

2.5 Argumentation in different disciplines

Argumentation practices contribute a lot to science education in which, argumentation studies have an important place. Argumentation in science includes presenting and responding to claims, looking for justifications, making a decision after analyzing all claims (Ross et al., 2009). It develops communicative competences and critical thinking. Scientific literacy- being able to write and talk science- can be achieved more easily. It helps the development of reasoning and rational thinking. Since science is viewed as a social construction of knowledge from inquiry processes and communication among scientific community, argumentation studies are seen as appropriate for science education (Erduran & Jiménez-Aleixandre, 2008). Argumentation is an important part of scientific inquiry. Maloney and Simon (2006) stated that students need to be aware of the tentativeness of scientific knowledge so that they could better cope with the uncertainties in the case of decision making (Ross et al., 2009).

Besides developing students' skills on nature of science and leading to deeper learning with higher order thinking, scientific argumentation can be in the form of socio-scientific argumentation in which its issues have a basis in science and they impact society (Christenson, 2015; Ratcliffe & Grace, 2003) Through socio-scientific argumentation, development of students' citizenship in the cases of socio-scientific issues is targeted (Christenson, 2015; Tiberghien, 2008). Although socio-scientific argumentation includes values, moral judgements and emotional reasoning together with more than one position on an issue (Acar, Türkmen & Roychoudhury, 2010; Sadler & Zeidler, 2005), they also involve decision making where citizens are required to have scientific literacy and the ability to process scientific knowledge and critical thinking (Erduran & Jiménez-Aleixandre, 2008; Norris & Phillips, 2003). Hence, engaging in scientific argumentation with socio-scientific issues provides opportunities for developing students to become more critically responsible citizens (Zeidler, 2014).

In mathematics education, there has been an increasing attention on argumentation. Two reasons for this increase are the recognition that natural languages rather than formal languages are the basis of human thinking and that communication and social processes are important in mathematics education (Reid & Knipping, 2010). Argumentation involves conjecturing, making hypotheses, representing mathematical ideas, taking others' point of views, and analyzing mathematical statements. Furthermore, argumentation exists in mathematical practices and not all mathematical activities are formal. In case of the application of problem solving strategies, argumentation practices can help one to solve problems, to resolve uncertainty, to formulate hypotheses, to produce explanations and to test one's understanding when it is considered as a critical and collaborative inquiry (Carrascal, 2015). Boero (1999) discusses argumentation in six phases of mathematical activity: (1) conjecture production, (2) formulation of a statement, (3) exploration of the content of the conjecture, (4) selecting and enchaining arguments into a deductive chain, (5) organization of the enchained arguments into a proof and (6) approaching a formal proof (Reid & Knipping, 2010). Krummheuer views argumentation as essential for learning mathematics (Reid & Knipping, 2010).

Overall, argumentation takes important place in science and mathematics education and contributes important skills when it is practiced. In the next section, structure and analysis of arguments are explained.

2.6 Structure and analysis of arguments

While generating arguments in class, many tools are established and used. Sometimes, formal logic is not adequate to analyze these arguments due to two reasons. First, students' thinking might include illogical elements which may be important for their development of thinking in the future. Second, use of natural language in proof generation may prevent arguments being captured by formal logic. That is why, researchers use a tool developed by Stephen E. Toulmin (1958) to analyze the arguments developed by students. The field-independent Toulmin model has made an important contribution to informal logic (Toulmin, 2003; Ubuz et al., 2012). Moreover, Toulmin's model-layout- is not only used in the analysis but also in assessment and construction of arguments (Banegas, 2013).

According to Toulmin, the core of an argument consists of a claim (C), data (D) which supports that claim and a warrant (W) in Figure 2. This core of argumentation is labeled as a three-part structure of argumentation (Cabassut, 2005). Warrants are the statements that connect data with claims. Warrants show us how one gets claims out of those data (Conner et al., 2014; Toulmin, 2003; Ubuz et al., 2012). Inglis, Mejia-Ramos and Simpson (2007) claim that there are two warrant types: Deductive and reference warrant. When students use reasoning like numerical computing, applying a rule or a theorem, constructing new ideas from the theorems or definitions, they use deductive warrants. In contrast, when they refer to a theorem, a rule or a definition they apply reference warrant (Ubuz et al., 2012).

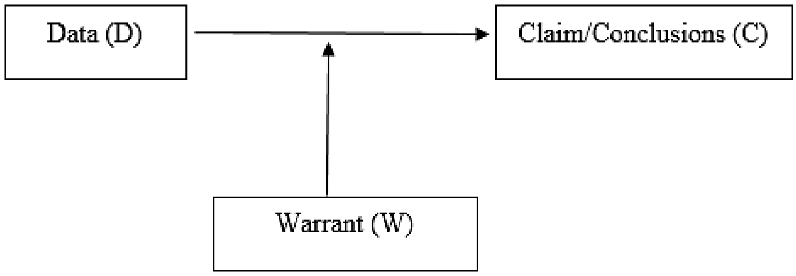


Fig. 2 Components of the core of an argument by Toulmin(1958)

Different kind of warrants may bring different degrees of force or confidence on conclusions. Qualifiers (Q) are parts of arguments which are statements of the certainty and level of confidence of claims. “Necessarily”, “Probably” or “Presumably” are examples of the adverbs used in qualifiers. Sometimes, there may be cases in which warrants may not support claims. There may be exceptional situations, which should be identified when argument is presented. This constitutes the Rebuttal (R) part of the arguments. Rebuttals are statements that describe circumstances under which warrants are invalid. For the general acceptability of warrants, there could be some other support from outside which may be from more reliable sources and authorities. These statements are called as Backing (B). All the components of Toulmin’s model/ Toulmin’s layout is shown in Figure 3.

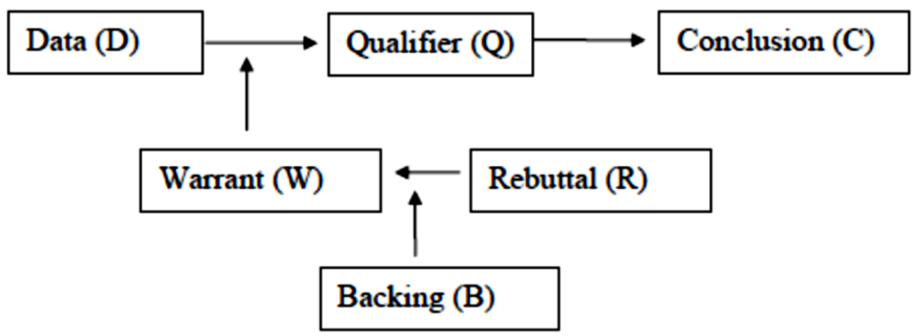


Fig. 3 Toulmin's model (1958)

There are two different forms of arguments based on their use of warrant. These are arguments of plausibility and arguments of necessity. Arguments of plausibility refer to arguments in which the warrants lead one to obtain tentative conclusions with qualifiers like 'probably' and exceptions or conditional circumstances. Arguments of necessity, however, refer to the arguments which contain warrants that clearly bring one to the conclusion. The 'modus ponens' is given as an example of this type of argument: A is observed, and also 'if A than B' is true, then B is true as well (Cabassut, 2005).

When students' argumentation skills are analyzed in the discourses or in the written surveys, Toulmin's layout is usually used as the core analysis framework. To categorize different kind of argumentation levels, Venville and Dawson (2010) developed a scheme based on the inclusion of parts of arguments as stated by Toulmin (1958). Importance was given on the presence of the claim, data, warrant, backing and qualifier in these schemes. Accordingly, four levels of argumentation were constructed. In level 1, there exists only a claim which is a statement, proposition or conclusion. In level 2, in addition to claim, data and or warrant is provided. In level 3, in addition to a claim, data or warrant and backing or qualifier is presented. In level 4, data or warrant with a backing and a qualifier is presented with a claim (Kaya, 2013; Venville & Dawson, 2010). These 4 levels of argumentation are shown in the Table 1.

Table 1. Argumentation Levels by Venville and Dawson (2010)

Levels	Description
Level 1	Claim (statement, conclusion, proposition only)
Level 2	Claim, data (evidence supporting the claim) and/or warrant (relationship between claim and data)
Level 3	Claim, data/warrant, backing (assumptions to support warrant) or qualifier (conditions under which claims are true)
Level 4	Claim, data/warrant, backing and qualifier

Quality of written arguments was measured through other frameworks by other researchers. Erduran, Osborne and Simon (2004) suggested a framework consisting of five levels for the identification of the quality of written arguments and of arguments constructed in argumentative discourses. Accordingly, the suggested five levels of arguments indicate the quality of arguments from the lowest quality, level 1, to the highest quality, level 5:

- Level 1: Arguments with a simple claim versus a counter-claim or a claim versus or a claim
- Level 2 : Arguments with a claim versus claim and with either data, warrants, or backings but not with any rebuttal
- Level 3: Arguments with a claim versus claim and with either data, warrants, or backings and an occasional weak rebuttal
- Level 4: Arguments with several claims or counter claims and a clearly identifiable rebuttal

- Level 5: Extended arguments with more than one rebuttal

This framework takes the rebuttal component of an arguments more into focus. The existence and quantity of rebuttal makes the differences in the quality of arguments. Since offering rebuttals in written arguments is difficult (Erduran, 2008), the highest levels of arguments, level 4 and level 5 arguments, are mostly observed in argumentative conversations (Evagorou & Osborne, 2013).

Osborne and his colleagues also developed a framework to analyze students' arguments and to follow their learning progression for argumentation in science (Osborne, Henderson, MacPherson, Szu & Wild, n.d.). They suggested three levels of arguments with sub-levels. According to their framework, students at level 0 can state a claim, identify a claim or provide evidence to support a claim. At level 1, students are able to construct a reasoning which links claim and evidence, identify the reasoning, construct a complete argument or provide an alternative counter argument. At level 2, students can provide a counter-critique, construct one-sided comparative argument, present two-sided comparative argument, or construct a counter claim with justifications. This framework also involves the components of Toulmin model.

Overall, components of the arguments identified by Toulmin (1958) have been widely used by researchers. Differences between students' arguments led some researchers to develop frameworks to differentiate students' argumentation skills by Venville and Dawson (2010), Erduran and her colleagues (2004) and by Osborne and his friends (n.d.). In this study, students' argumentation levels were detected through the framework developed by Venville and Dawson (2010) due to its easiness and applicability to eighth-grade students' argumentation skills. Other frameworks developed by Erduran and her friends and Osborne and his colleagues involved the formation of counter and comparative claims, which exist more collaborative argumentation. Since this study involved individual argumentation, Venville and Dawson's argumentation levels (2010) were used in this study. In the following section, relationship between proof and argumentation is explained.

2.7 Relationship between proving and argumentation

Every proof is an argument but not every argument is proof (Krabbe, 2013). To be a proof, an argument should fulfill certain conditions. Requirements of proofs were identified by Aristotle as: "They are true, indemonstrable, better knowable than the conclusion and gives the cause of the conclusion. The conclusion should be obtained from deductive argument.". Arguments that do not have these properties cannot be counted as proofs. It is added that, to be accepted as proof, an argument needs to be dialectically correct and must deal with all possible counterexamples, objections and potential cases (Krabbe, 2013). Dufour (2013) also supports the claim that a perfect proof is certainly different from argument. A proof raises no critical comment and no request for further explanations.

Argumentation and proof are distinguished as being not the same of nature by Balacheff (1998). The aim of argumentation is attaining agreement among partners in social interaction. The first aim of the argumentation is not to provide the truth of any statement. Since argumentation is a social activity and an open process, it can benefit from any kind of means. On the other hand, for constructing proofs one has to follow the requirements for the use of knowledge, which is taken from a body of knowledge on which mathematics authorities agree (Cabassut, 2005). Proof is seen as a combination of argumentation and reasoning since it involves both justifications and logical processes (Reiss et al., 2008).

The difficulty in proving skills of students mainly stem from the clear-cut distinctions between applications of validations among elementary school level and afterwards. Both argumentation and proof contain the acts of validation and justification. Argumentation is viewed as a precursor of proof. What usually takes place in mathematics classes is not a formal proof but a precursor of proof (Conner et al., 2014; Wood, 1999). In other words, students are engaged with argumentation activities that lay the foundation for formal proof. When writing proofs, students have difficulties in sequencing the inferences and deductive reasoning. They are expected to switch from a practical domain to a theoretical domain instantaneously. Students have trouble in understanding that practical validations by empirical observations are no longer acceptable as writing deductive proofs. The transition from practical to theoretical mathematics and geometry should be done in the elementary school curriculum. To do this, students should develop a degree of abstraction because mental constructs and objects in abstract space exist in the form of ideas in theory. Hence, it is important to establish cognitive unity among structures of arguments, which is a result of not being able to use deductive arguments due to the fact that inductive arguments are so dominant in thinking and reasoning. Deductive reasoning should be encouraged in elementary school for dealing with mathematical situations and for the smooth transition between practical mathematics to theoretical mathematics. Ability to reason deductively is a demanding process and requires extensive experiences and time for exercising proof properly (Cyr, 2011; Knipping, 2003).

Asking students to prove a statement can often lead to unsuccessful results. The reason for this situation is their lack of experience with argumentation tasks. Students who do not develop their own arguments and evaluate own and others' arguments and reconstruct new arguments experience difficulty when proving a statement. To enhance proving skills, it is necessary for students to familiarize and internalize deductive reasoning. But first, they should be working on argumentation exercises and activities (Güneş, 2013).

The van Hiele model indirectly points out the relationship between argumentation and proof. The van Hiele model was developed in the 1980s by Dina van Hiele-Geldof and Pierre van Hiele in order to understand children's level of geometric thinking. The model concerns how children's geometric thinking evolves progressively. According to the model, there exist five levels of geometric thinking (Breyfogle & Lynch, 2010; van Hiele, 1959):

Level 0: Recognition or Visualization

Level 1: Analysis

Level 2: Ordering or Informal Deductive

Level 3: Deduction or Formal Deductive

Level 4: Rigor

Students at level 0, visualization level, can sort shapes by looking at their similar appearances. At level 1, analysis level, children are able to list the properties of shapes, but they cannot comprehend the relationship between these properties and they cannot notice that some properties imply others. At level 2, ordering or informal deductive level, students can formulate meaningful definitions and produce informal deductive arguments. Students at level 3, deduction or formal deductive level, can understand relationships between properties of shapes and also comprehend relationships between definitions, theorems, axioms and postulates. They can learn how to do a formal proof and understand why proof is needed. At level 4, rigor level, children can think within an abstract mathematical system (van Hiele, 1959).

It was found that most elementary school students are at the visualization (level 0) or analysis level (level 1) and also some middle-school children are at the informal deduction level (level 2) (Abdullah & Zakaria, 2013; van Hiele, 1959). By the time a student finishes middle school, she or he is expected to be at least at the informal deductive level (level 2) (van Hiele, 1959).

In the van Hiele model, the levels are not age dependent. Instead they are related to the experiences that students have (Breyfogle & Lynch, 2010). The levels are sequential; children must pass through the levels in the given order as their understanding develops. In order to pass to the next level, students need a lot of experiences that involve exploration and communication about geometrical concepts (van Hiele, 1959).

The van Hiele model implies that students should develop informal reasoning and they should be introduced with deductive reasoning with various experiences to be able to improve their geometric thinking levels. Students should be directed to communicate with others through verbal and written approaches. They need to develop their geometrical thinking step by step (van Hiele, 1959).

The characteristics of the van Hiele model explained above resemble the practices of argumentation studies and support the claim that argumentation studies may be related with proving and can be precursor of proof. As students move from visualization level to rigor level, students learn how to produce formal proofs and use abstract notations in mathematics. They obtain inductive and deductive reasoning skills through experiences, which lay the foundation for producing formal proof.

Until reaching the top level of geometric thinking which is associated with proof, students need to be familiar with argumentative activities.

Overall, there exist different views on the relationship between proof and argumentation in the literature. Some researchers suggest no relationship between these two constructs (Balacheff, 1998; Dufour, 2013), however some others point out the possible relationship (Conner et al., 2014; Wood, 1999). The van Hiele model also indirectly indicates the relationship between proof and argumentation skills through presenting five geometric thinking levels with the shift from inductive reasoning to more formal proofs (van Hiele, 1959). In the following section, there are some suggestions for improving students' proof and argumentation skills.

2.8 Suggestions for developing proof and argumentation skills

Language frames are tools that help students to form arguments, which emphasize academic language and the syntax to develop and communicate arguments. These frames improve students' academic writing and reasoning skills (Ross et al., 2009). Teachers may present the language frames to students for asking them to produce arguments on their own or with their peers. On the other hand, they may use them when they are lecturing as if they are thinking aloud. Showing their cognitive processes can lead students to model the use of language frames and improve their reasoning skills (Ross et al., 2009).

Inductive reasoning can be used to form deductive arguments or proofs. However, not all inductive arguments have potential to construct more formal proofs. It is important to distinguish the inductive arguments that lay the foundations for formal arguments from those that do not. Middle school mathematics teachers should identify arguments with "key ideas" which includes an insight or understanding that form bases for more rigorous and formal arguments (Yopp, 2009). The notion of key idea was introduced by Raman (2003). She claimed that key ideas function like bridges between informal arguments with inductive reasoning and formal proofs. However, an inductive argument with a key idea is still away from a formal proof. Hence, the key idea should be expanded so that it is general and works for all cases through the uses of symbols, prose or algebraic representations (Yopp, 2009).

2.9 Summary of the literature review

Proof is an important component in mathematics education and it is viewed as inseparable from mathematics (Knuth et al., 2002). The concept of proof begins in preschool ages with classifying, matching, ordering and comparing activities. Logical thinking develops through the years from concrete to abstract thinking. Making generalizations and use of deductive reasoning can be seen in middle school ages and forms of proofs can be conducted by the time of high school period (Altıparmak & Öziş, 2005; Piaget, 1985). Since proof skills improve within years of experiences, the concept of proof should be included in the curricula of all grade levels (Ellis et al., 2013). Proof has a lot of functions and roles in mathematics education. It provides justification and promotes mathematical understanding (Hanna, 2000). It has five main functions and goals: verification, explanation, systematization, discovery and communication (de Villiers, 1990).

Proving skill refers to the ability to conducting and reading proofs (Csikos, 1999). Students' different proof skills were required to be compared through proof schemes. The most commonly used proof scheme was developed by Harel and

Sowder (1998). In this framework, there exist three proof schemes as external conviction proof scheme, empirical proof scheme and analytical proof scheme. In the external proof scheme, no discovery or creativity is involved in students' proofs and most of the time an authority is present as the only source of knowledge. In empirical proof scheme, individuals' proofs are formed based on quantitative evaluations and sensory experiences. Individuals produce analytical produce when they apply deductive reasoning and reach general statements through use of symbolic representations and mathematical theorems (Harel & Sowder, 1998).

Investigation of students' proof skills reveals that they have difficulties in constructing proofs (Bieda & Lepak, 2014; Harel & Sowder, 1998; Weber, 2001). Knowledge of concepts and definitions are not adequate to form proofs (Heinze & Reiss, 2003; Sen & Güler, 2015). Students have trouble while appropriately stating mathematical theorems and notations (Carrascal, 2015; Sen & Güler, 2015). Students' failure in constructing and evaluating proofs can be due to considering proof as an isolated topic in mathematics (Reid, 2011). Students' proof evaluations were found to be affected by their competencies in mathematics, the instruction that they take on proof and their beliefs about the purpose of the proof (Healy & Hoyles, 2000; de Villiers, 1990). Students mostly rely on empirical evidences while evaluating proofs (Liu et al., 2016). Many inconsistencies among proof construction and proof evaluation skills had been observed in students' proof schemes. Students may have different proof schemes in different contexts (Harel & Sowder, 1998; Heally & Hoyles, 2000). Students' proof skills increased in some studies, leading to the conclusion that exposure to proof instruction can bring about positive changes in students' proof schemes (Cooper et al., 2011; Sen & Güler, 2015).

Argumentation is a reasoned discourse and is associated with convincing (Reiss et al., 2008). It involves presenting and evaluating arguments. Through argumentation, students establish truth of a claim and deepen their conceptual understanding (Carrascal, 2015; Staples & Newton, 2016). Argumentation has been studied in science education and mathematics education. Scientific literacy, understanding the nature of scientific knowledge and development of social responsibility in socio-scientific issues can be achieved with scientific argumentation studies (Erduran & Jiménez-Aleixandre, 2008; Norris & Phillips, 2003). Mathematics argumentation studies provides opportunities for conjecturing, making hypotheses, representing mathematical ideas, taking others' point of views, and analyzing mathematical statements (Carrascal, 2015; Reid & Knipping, 2010).

The field-independent Toulmin model was formed to analyze arguments constructed by students (Banegas, 2013; Toulmin, 1958). This model suggest that an argument may include different components as claim, data, warrant, qualifier, backing and rebuttal. Claims are statements; data refer to the evidence for these statements. A warrant connects data to the claim. Qualifier is statement of the certainty and level of confidence of claims. Rebuttal is a statement that describes circumstances under which warrants are invalid. Backing is a statement from more reliable sources, which supports the warrant in the arguments (Toulmin, 1958). As students' arguments differ in the amount of these components, there formed frameworks for distinguishing different levels of arguments. Venville and Dawson (2010) developed a scheme based on the inclusion of components of arguments as stated by Toulmin (1958). Accordingly, four levels of argumentation were constructed. Level 1 arguments consist of only a claim which is a statement, proposition or a conclusion. In level 2, in addition to a claim, data and or warrant is provided. In level 3, in addition to a claim, data or warrant and backing or qualifier is presented. In level 4, data or warrant with a backing and a qualifier is presented with a claim (Venville & Dawson, 2010).

The relationship between proof and argumentation skills has been studied by many researchers (Cabassut, 2005; Conner et al., 2014; Dufour 2013; Güneş, 2013; Krabbe, 2013; Reiss et al., 2008; Wood, 1999). There exist clashing views about whether these constructs share similar characteristics and structure and whether these skills have relationships. There have been studies which defend that argumentation and proof skills are related in their findings (Cabassut, 2005; Güneş, 2013; Reiss et al., 2008). The van Hiele model indirectly points out that there might be a relationship between these two skills.

To develop students' proof and argumentation skills researchers provided some suggestions. Beginning with inductive reasoning, presenting tasks with key ideas, asking students to change their statements so that it encompasses all the possible cases were offered for leading students to move from inductive reasoning to deductive reasoning and for improving students' proving skills (Yopp, 2009). To enhance argumentation skills, teaching and use of language frames in lectures was proposed. These language frames can help students to communicate their arguments in a more efficient ways and to improve their reasoning skills (Ross et al., 2009).

Based on the literature review, analyses of students' skills in proof and argumentation tasks are needed with empirical studies. Considering the importance of proof and argumentation for education and the probable relationship between proof and argumentation skills, researcher should consider investigating students' proof and argumentation skills before implementing any intervention. This study can cover the need for empirical studies on the relationship between students' proof and argumentation skills, which were presented only theoretically in the literature.

CHAPTER 3

STATEMENT OF THE PROBLEM

Proof and reasoning are perceived as important processes for mathematics educators. Argumentation is a concept related to students' reasoning and justification abilities. Middle school students are expected to be competent in constructing and evaluating proofs and developing good arguments. The aim of this study is to investigate these competencies in middle school ages. Specifically, the researcher aims to explore the following:

- Middle school students' proof construction skills
- Middle school students' proof evaluation skills.
- Middle school students' argumentation skills and their relationship with proof construction and evaluation skills.

The target population of the study consists of eight-grade students. The reasons for studying with this group of students is that, they are in the final grade in the middle school and they are in the transition period to high school where proof is much more emphasized in mathematics lessons. It could be beneficial to study students' skills in proving and argumentation in middle school period since it is reported in several studies that high school students have difficulties in producing formal proofs and sound mathematical arguments (Bieda & Lepak, 2014). The students in these age groups are considered to be in the stages of concrete operational and formal operational and sometimes in between (Piaget, 1985). It could make sense to observe whether eight-grade students are able to use and appreciate the use of symbolic representations and abstract notations so that they are more likely in formal operational stage. This study aims to reveal information about eight-grade students' cognitive skills, potentials and readiness in proving and argumentation contexts before becoming high school students.

CHAPTER 4

RESEARCH QUESTIONS AND OPERATIONAL DEFINITIONS

This chapter presents information about research questions and operational definitions for the variables that were used in this study.

4.1 Research questions

This study aims to investigate proof and argumentation skills of eight-grade students and the relationship between them. The research questions were formulated as to describe and detail students' proof and argumentation skills.

1. What are the students' proof construction levels?
2. What are the students' proof evaluation levels?
3. Are there any statistically significant differences between students' proof construction and proof evaluation levels for algebra and geometry proof tasks?
4. Are there any gender differences in students' mathematics achievement, proof construction levels and proof evaluation levels?
5. What are the students' argumentation levels?
6. Are there any gender differences in students' argumentation levels?
7. Are there any statistically significant relationships between mathematics achievement, proof construction levels, proof evaluation levels and argumentation levels of students for each task?
8. How are the performances of students in proof construction tasks? What are the characteristics of students' proof schemes?

9. What are the factors which make students convinced in proof evaluation tasks?
10. How do the students perform in argumentation tasks? What are the characteristics of students' argumentation levels?

4.2 Operational definitions

In this study, the term proof refers to the conceptual syntactic derivations with specific technical approaches (Hanna, 2000). Proving skill refers to proof construction and proof evaluation levels as assessed by the students' performances on the assessment tool by means of Harel and Sowder's (1998) framework for proof schemes. Argumentation skill refers to the argumentation level similarly assessed by the students' level based on the argumentation task in the assessment tool by means of Venville and Dawson's (2010) framework for argumentation levels. The variables and the operational definition that are used in this study are as follows:

Mathematics achievement: Mathematics achievement of students refers to the final score (out of 100) obtained from the judgements of students' mathematics teachers about their mathematics performances in their report cards.

Proof Construction (PC) Skill: Proof construction skill refers to the category of the students' response in Harel and Sowder's (1998) proof scheme in the proof construction tasks. This category could be external conviction proof scheme if it is based on an external authority and does not involve a cognitive effort, empirical proof scheme if it is based on empirical evidences or analytical proof scheme if it is based on deductive reasoning and use of algebraic representations. Each of them is scored as 1, 2 and 3, respectively for both algebra and geometry tasks. Unanswered tasks and irrational answers were scored as 0. (Appendix B)

Proof Evaluation (PE) Skill: Proof evaluation skill refers to the category of the students' response in Harel and Sowder's Proof Scheme (1998) in proof evaluation tasks. This category could be external conviction proof scheme if it is based on an external authority and does not involve a cognitive effort, empirical proof scheme if it is based on empirical evidences or analytical proof scheme if it is based on deductive reasoning and use of algebraic representations. Each of them was scored as 1, 2 and 3, respectively. Unanswered tasks and irrational answers were scored as 0. (Appendix B)

Argumentation Skill: Argumentation skill refers to level of the students' response in Venville and Dawson (2010)'s scheme in argumentation tasks. This level could be Level 1 (only claim), Level 2 (a claim and data and/or a warrant), Level 3 (a claim, data and/or a warrant, a backing and/or a qualifier), and Level 4 Level 3 (a claim, data and/or a warrant, a backing and a qualifier). Each category was scored as 1, 2, 3 and 4. Unanswered tasks and irrational answers were scored as 0. (Appendix B) (Appendix C)

Students proof and argumentation skills were scored for both algebra and geometry tasks. In this study, proof construction skill was obtained through the average of proof construction skills in algebra and geometry. Likewise, proof evaluation skill was obtained through the average of proof evaluation skills in algebra and geometry. The argumentation skill was acquired through the average of argumentation skills for three argumentation tasks.

CHAPTER 5

METHODOLOGY

This chapter presents the details of the methodology of this study. The following sections cover the design of the study, sample, instruments, procedure and data analysis, respectively.

5.1 Research design

The aim of this sequential explanatory mixed methods design is to examine eighth-grade students' proof and argumentation skills. In this design, quantitative and qualitative methods were used together to get benefit from strengths of each and to compensate weaknesses of both methods. Quantitative data provided general information about eighth-grade students' proof and argumentation skills and their relationships. Qualitative data was also used to explain these quantitative findings (Creswell, 2014). While collecting data, both quantitative and qualitative data sets were taken into consideration. The instrument of the study (see Appendix D and Appendix E) was used to detect students' proof and argumentation skills by their proof schemes and argumentation levels. At the same time, students' explanations and rationales in the written survey were acquired. The reason for gathering both quantitative and qualitative data was to embody elements of both approaches to provide more in-depth understanding of research questions than only one approach (Creswell, 2014).

5.1.1 Quantitative design of the study

A survey with open-ended questions was used as a quantitative data source for this sequential explanatory design study (see Appendix D and Appendix E). The rationale of using this form was to present categorical and numeric descriptions of students' proof and argumentation skills and their mathematics achievement scores. The identification of scores led to determine eighth-grade students' proof schemes in four categories, argumentation levels in five categories and mathematics achievement scores out of one-hundred points. After the analysis of the data from this survey, quantitative findings were supported and detailed through analyses of students' responses in a more elaborated manner (Creswell, 2014).

5.1.2 Qualitative design of the study

The survey with open-ended questions for the quantitative data sources was used for qualitative design of this study as well. The qualitative data findings were used to determine characteristics of students' proof schemes and argumentation levels and to detect factors that led students' proof schemes while evaluating proofs. Qualitative data provides opportunity for understanding why participants perform in certain ways in natural setting with more in-depth manner (Creswell, 2014). Thus, examination and interpretation of students' proof and argumentation skills can become more persuasive with the use of qualitative design.

5.2 Sample

In this study, the sample consisted of 242 middle school students from four different public schools. There were 106 male students and 136 female students from eighth-grade level. The sample included students from four schools in three different districts in İstanbul: Beşiktaş, Ümraniye and Esenler. These districts represent three different socioeconomic conditions. It was tried to establish a representative sample of the students in İstanbul in terms of socioeconomic status. Convenience sampling was used for both school and classroom selections since the schools and teachers gave permission for data collection on a voluntary basis.

The reasons for selecting eight-grade students for this study were that they are in the phase of the transition from middle school to high school. Proof practices are mostly encountered and given much more importance in the high school period. Eight- grade students' performances in proof tasks have a potential to reveal their proving skills which will be transferred to high school (Piaget, 1985). Furthermore, eight-grade students are in the transition phase from the concrete operational stage to the formal operational stage according to Piaget's cognitive developmental stages (1985). In concrete operational stage, students rely on concrete entities and think in concrete ways whereas in formal operational stage, students are capable of thinking abstractly.

Two hundred and forty-two students in the sample were also formed the sample for qualitative analysis. Responses of 42 students were used to support qualitative findings.

5.3 Instruments

The assessment tool contained two proof construction tasks, two proof evaluation tasks, three argumentation tasks, and demographic items as gender and school code. Also, students' mathematics achievement scores were asked to be filled in in the assessment tool. The items of the whole instrument were selected from different sources and combined as a survey by the researcher in order to get information on students' proof construction, proof evaluation and argumentation skills together. The instrument of this study is given in Appendix D and in Appendix E .

Mathematics achievement score of students was the final score that they got (out of one hundred points) from their mathematics teachers' evaluations in the previous term. Students recorded their own mathematics achievement score on top right-hand corner of the assessment tool.

In the proof construction tasks, students were expected to generate proofs for an algebraic and a geometric statement. Proof construction tasks were taken and adapted from mathematics textbooks and related literature sources (Bieda & Lepak, 2014). In proof evaluation tasks, students were given exactly the same statements in the proof construction tasks, in which three imaginary students proved each statement. Students also were required to state the most convincing response, reasons for selecting that answer as more convincing and suggestions for others to be more convincing in last two sub-questions. In the proof evaluation part of the instrument, same tasks in the proof construction part were used and the structure of the task was adapted from a study conducted by Aylar (2014).

The contents of the proof tasks were selected for the eighth-grade level taking into consideration that students in that grade level have already learned the content in the previous years. In the algebraic proof task, students need to know the odd and even numbers and the properties about the addition of add and even numbers. These objectives are attained first in the third grade and students have continuously exposed them in Turkish mathematics education curriculum. In the geometric proof task, students ought to know the properties of the angles which are formed by the intersection of a transversal and two parallel lines. This objective is attained in the seventh grade in Turkish mathematics curriculum. The decision of having tasks from third and seventh grade was taken because students are familiar with these objectives. Having both algebra and geometry tasks together as proof tasks, it was aimed to provide diversity in the domains of mathematics education. The objectives of the proof tasks are given in Appendix F.

In the argumentation tasks, students were asked to present their arguments with their rationales. Argumentation tasks comprised one mathematics tasks and two science tasks. Two science argumentation tasks were asked because the first science task was about only mixtures, but the other task involved a socio-scientific decision-making component. Presenting mathematics and science argumentation tasks together was done due to the desire of having information about students' argumentation skills from different contexts. In mathematics task, students were asked to write an argument about the truth of a given statement which involves inequality and exponential expressions. The content of the task was selected to be proper to students' level. Students learn the inequality signs in the third grade and exponential numbers in the sixth grade. In the seventh grade they learn finding powers of integers and rational numbers in the Turkish mathematics curriculum. In the first science task, students were asked to construct an argument for defending one of the two friends who have opposing views about the existence of sugar after mixing it with hot water. Students were given evidences and information to present their arguments. In the second science task, students were expected to give advice to a governor for selecting the best method among two for obtaining filtered water. Students were given the procedures and disadvantages of each method in the instrument to defend their positions. These science tasks are corresponding to the topics of mixtures and their decomposition. The decision of having science argumentation tasks in these topics were taken because the researcher has more confidence in the topics of mixtures and their decompositions when compared to other science argumentation tasks. It was regarded that students' argumentation skills can be best detected when the researcher has more comprehensive knowledge about the content. Students first meet with these topics in the fourth grade in Turkish science curriculum. Curricular objectives of the argumentation tasks are given in Appendix F.

The mathematics argumentation task was adapted from a study conducted by Nardi, Biza and Watson (2014). Also, it was adapted from a science argumentation task developed by Kaya (2013) so that the structure of the task was maintained but the content was modified into a mathematics task. The science argumentation tasks were taken, translated, shortened and adapted from the tasks developed from a project titled "Assessment of Argumentation in Science Beyond Multiple Choice" which was carried out by Stanford University. The tasks, which are called "Desalination" and "Mixing Sugar and Water", were taken from the assessment items of this project (http://scientificargumentation.stanford.edu/assessments/) (Osborne et al., n.d).

Reliability of the instrument was tested by looking at inter-rater agreement between the scorings of two raters. Another researcher for inter-rater agreement analyzed approximately ten percent of the data, twenty-five students' responses. The correlation between the evaluations of two researchers for proof construction tasks was calculated as 0.74. For argumentation tasks, the correlation was calculated as 0.73. To ensure the content validity of the instrument, several mathematics and science education experts examined it. Improvements and changes for the instrument were done through their suggestions before conducting the study.

Qualitative data source of the study consisted of students' written responses and expressions in the tasks. Whole instrument together with all sub-sections comprised the data source for qualitative analyses. Hence the tasks in the instrument functioned as data sources for both quantitative and qualitative data analyses.

5.4 Procedure

After developing an instrument for proof and argumentation skills and getting expert opinions, a pilot study with twenty eighth-grade students was conducted. According to the results of the analysis of the pilot study, the number of the items were decreased (from nine to seven items) to achieve two purposes: Overlapping objectives of two tasks were considered and one of the tasks discarded because the students' responses did not differ in these two tasks. In the pilot study, students could complete the instrument in forty-minutes but in a rush. Hence, they provided short responses and explanations for the items. To finish answering the tasks in one-lesson-hour duration, to allocate necessary time for each item and to get more detailed and long explanations, the number of items decreased. Also, the sequences of the tasks were changed and revised because it was observed that a task could direct students' responses and may become a confounding variable for the latter task and for the whole instrument.

With the latest revision, the permission was taken from Provincial Directorate of National Education of İstanbul (İstanbul İl Milli Eğitim Müdürlüğü) and four public schools were selected for the data collection. Students from the selected schools were given the instrument in two phases. In the first phase, students were given proof construction tasks. Approximately fifteen minutes were given to students to complete these tasks and then these forms were collected. In the second phase, the remaining tasks were given in an attached form and the overall instrument completed in one-lesson hour. These two different forms were reattached by matching the papers according to students' school codes.

Students' responses were first analyzed with quantitative data analysis approaches. Then students' proof schemes and argumentation levels formed the base for the follow-up qualitative data analyses. Students' responses in proof construction, proof evaluation and argumentation tasks were coded and categorized then recorded by using their written responses in the survey. Research findings were presented with both quantitative and qualitative data analyses to reveal a broader picture for students' proof and argumentation skills.

5.5 Data analysis

In this section, the analyses of the quantitative and qualitative data are clarified. The data analysis procedures for quantitative and qualitative data are elaborated in the following sub-sections (5.5.1 and 5.5.2), respectively.

5.5.1 Quantitative data analysis

The main focus of the analysis was describing and detailing eight-grade students' proof and argumentation skill in general. That is why; the average values for students' performances were calculated from two proof construction tasks, two proof evaluation tasks and three argumentation tasks. Hence, proof construction, proof evaluation and argumentation skills were obtained respectively. To obtain proof construction and proof evaluation skills, Harel and Sowder's (1998) proof schemes framework was used. Accordingly, students' responses in proof construction parts were scored as "0" when there existed no proof. They were scored as "1" when students assigned to "External conviction proof scheme" since they did not have cognitive effort on their proof, they only transferred the information that is presented to them, or they repeated what an authority said before, they used the symbolic representations without knowing their meaning and functions. Students' responses were scored as "2" when they were assigned to "Empirical proof scheme" since they gave examples, applied trial and error and representations that lack generalization power. Students' responses were scored as "3" when students were assigned to "Analytical proof scheme" since they used deductive reasoning, reached a general judgement through symbols or algebraic expressions. Furthermore, students' responses in proof evaluation parts were scored as "0" when they did not select any option as the most convincing. They were scored as "1" when they selected the "External conviction proof scheme" as the most convincing option. Students' responses were scored as "2" when they selected "Empirical proof scheme" as the most convincing option. They were scored as "3" when they selected "Analytical proof scheme" as the most convincing option. Also, argumentation skills of students were obtained by detecting students' argumentation levels through the framework developed by Venville and Dawson (2010). Accordingly, students' responses were scored as "0" when there was no argument. Their responses were scored as "1" when there existed a level 1 argument (only a claim). They were scored as "2" when there existed a level 2 argument (a claim, data and or warrant). They were scored as "3" when there existed a level 3 argument (a claim, data and or warrant, and qualifier and or backing). Students' responses were scored as "4" when there existed a level 4 argument (a claim, data and or warrant, qualifier and backing).

The obtained data was analyzed calculating the Spearman rho correlation coefficient, using Wilcoxon signed-ranks test and Mann-Whitney U test of the SPSS program and the level of significance was specified as 0.05. The decisions of the use of these non-parametric tests were made after having non-normal distribution of the variables with Shapiro-Wilk normality tests.

The Spearman's rho coefficients were obtained to reveal information about the relationship between variables in the study. Spearman rho was computed because of the non-normal distribution of the variables. Wilcoxon-signed ranks test was used to compare the proof construction and proof evaluation skills. According to Huck (2012), Wilcoxon-signed ranks test is reasonable when the compared variables are from paired samples and when the data show non-normal distribution. Mann-Whitney U test was conducted to see whether there exist gender differences in students' mathematics achievement, proof construction, proof evaluation and argumentation skills.

5.5.2 Qualitative data analysis

The analysis of the survey findings became more comprehensive and meaningful through including the analysis of students’ responses in an in-depth manner. All tasks in the instrument were analyzed for qualitative analysis. While all tasks used in quantitative data analysis were used in the qualitative data analysis too, the latter two sub-sections of the third and fourth tasks were used only for the qualitative data analysis. Themes, codes and categories were formed after analyzing all student responses for each task. The use of categorizations and coding was helpful for attaining shared and distinctive characteristics and for revealing patterns in students’ responses (Creswell, 2014).

Students’ proof schemes and argumentation levels were taken as bases for the coding of their responses. Codes and themes were attained for each proof scheme and argumentation level of students separately through using students’ own responses and expressions. Literature was also used as coding source for providing coherence between literature review and research findings.

CHAPTER 6
RESULTS

Since this is a mixed-methods study intended to investigate proving skills, argumentation skills and examine the relationship among these skills for eight-grade students, the results will be presented in three parts. The quantitative component of this study first set out to describe the proof and argumentation skills of students then aimed to find correlations among proof items, argumentation items and mathematics achievement. The qualitative component of the study was conducted to observe students' written responses to each item in detail.

Descriptive statistics of the quantitative data is presented in the first section. In Section 6.2, findings of the statistical analyses are presented. In section 6.3, the focus will be on the qualitative findings from the written data.

6.1 Descriptive statistics

In this section, students' mathematics achievement, proof construction skills, proof evaluation skills and argumentation skills are summarized and described. As seen in Table 2, 103 male and 135 female students' mathematics achievement scores, their central tendency and dispersion measures are summarized. Four students' mathematics achievement scores were missing therefore they were not included into the analysis. According to the descriptive analyses of mathematics achievement, male students' scores ranged from 30 to 100, a median of 87.5 and a mean of 82.29, SD=16.15. The distribution was skewed to the left (skewness=-1.087, kurtosis=0.544). Female students' scores ranged from 45 to 100, a median of 83 and a mean of 79.97, SD=16.56. The distribution of mathematics achievement scores was skewed to the left (skewness=-0.574, kurtosis=-1.055). In total, the mathematics achievement score of the sample ranged from 30 to 100, a median of 85 and a mean of 80.98, SD=16.39. The distribution was skewed to the left (skewness=-0.781, kurtosis=-0.480).

Table 2. Descriptive Statistics of the Sample by Mathematics Achievement

Gender	N	Mean	SD
Male	103	82.29	16.15
Female	135	79.97	16.56
Total	238	80.98	16.39

Frequencies and percentages of proof construction skill in algebra task were summarized in Table 3 to respond the first research question: “What are the students’ proof construction levels?” As seen, most of the students (82.2 %) produced proofs in “Empirical” proof scheme, in which students tried to reach generalizations through examples. “External Conviction” proof scheme for this task was the least used proof scheme (5.4%), where students had no cognitive effort and relied on an authority. The other proof schemes were also show low frequency when compared to Empirical proof scheme.

Table 3. Frequencies and Percentages of Proof Construction Skill in the Algebra Task

Category	Frequency	Percent
No proof	20	8.3
External Conviction	10	4.1
Empirical	199	82.2
Analytical	13	5.4
Total	242	100

For the proof construction task in geometry, most of the students performed in the no proof category (48.3 %) as seen in the Table 4 as a response for the first research question: "What are the students' proof construction levels?" 73 students (30.2 %) provided "External Conviction" proof scheme, where students had no cognitive effort and relied on an authority and 35 students (14.5 %) constructed "Analytical" proof, in which students had symbolic representations with deductive reasoning. Also, 7 percent of the students provided "Empirical" proofs, where students tried to reach generalizations through examples.

Table 4. Frequencies and Percentages in Proof Construction Skill in the Geometry Task

Category	Frequency	Percent
No proof	117	48.3
External Conviction	73	30.2
Empirical	17	7.0
Analytical	35	14.5
Total	242	100

In this part, the second research question "What are the students' proof evaluation levels?" was targeted. In the proof evaluation task for algebra, students' responses indicated that empirical proof example was the most convincing proof scheme as

50.4 % of the students preferred “Empirical” proof example, where students tried to reach generalizations through examples, as most convincing. “Analytical” proof scheme, in which students had symbolic representations with deductive reasoning, was stated as the most convincing proof by 26.4 % of the students. “External Conviction” proof scheme, where students had no cognitive effort and relied on an authority, was also identified as the convincing proof by 22.3 % of the students. Only 0.8 % of the students did not state their preferences in Table 5.

Table 5. Frequencies and Percentages in Proof Evaluation in the Algebra Task

Category	Frequency	Percent
No proof	2	0.8
External Conviction	54	22.3
Empirical	122	50.4
Analytical	64	26.4
Total	242	100

For the proof evaluation skill in geometry task, most students indicated that they view “Analytical” proof example, in which students had symbolic representations with deductive reasoning, as the most convincing one (56.2%) as seen in Table 6. Only 2.1 % of the students did not report their preferences. Students who preferred “External Conviction” where students had no cognitive effort and relied on an authority and “Empirical” proof scheme where students tried to reach generalizations through examples, was quite close in frequency (45 and 56) and percentage (18.6 and 23.1 %) as a response for the second research question: “What are the students’ proof evaluation levels?”

Table 6. Frequencies and Percentages in Proof Evaluation Skill in the Geometry Task

Category	Frequency	Percent
No proof	5	2.1
External Conviction	45	18.6
Empirical	56	23.1
Analytical	136	56.2
Total	242	100

In Table 7, findings for the fifth research question: “What are the students’ argumentation levels?” was presented. Accordingly, produced arguments show quite similar percentages for mathematics argumentation task. The most important finding in this table is that the least produced argument level was Level 3 (14 %) in which an argument consists of claim, data and/ or warrant and qualifier or backing components. The most produced argument level was level 2 (32.6 %) in which an argument consists of claim and data and/ or warrant.

Table 7. Frequencies and Percentages in Argumentation Skill in the Mathematics Task

Category	Frequency	Percent
No argument	64	26.4
Level 1	65	26.9
Level 2	79	32.6
Level 3	34	14.0
Total	242	100

Note. Level 1= Claim, Level 2= Claim+ (Data/ Warrant), Level 3= Claim+ (Data/ Warrant) + (Qualifier/ Backing)

In Table 8, the fifth research question: “What are the students’ argumentation levels?” was responded. Accordingly, majority of the students (73.6 %) produced Level 2 arguments (Claim+ (Data/ Warrant)) for the first science argumentation task.

19 % of the students presented only their claim. 9 students (3.7 %) produced no argument. 8 students (3.3 %) constructed Level 3 arguments (Claim+ (Data/ Warrant) + (Qualifier/ Backing)). Only 1 student (0.4 %) provided Level 4 argument in which the argument consists of "Claim+ (Data/ Warrant) + Qualifier+ Backing" components.

Table 8. Frequencies and Percentages in Argumentation Skill in the First Science Task

Category	Frequency	Percent
No argument	9	3.7
Level 1	46	19.0
Level 2	178	73.6
Level 3	8	3.3
Level 4	1	0.4
Total	242	100

Note. Level 1= Claim, Level 2= Claim+ (Data/ Warrant), Level 3= Claim+ (Data/ Warrant) + (Qualifier/ Backing), Level 4= Claim+ (Data/ Warrant) + Qualifier+ Backing

In Table 9, students' argumentation levels in the second science task are summarized to respond the fifth research question as "What are the students' argumentation levels?" Accordingly, majority of the students (46.3 %) provided arguments in level 2, in which an argument consists of claim and data/warrant. Percentages of arguments in Level 1 (only claim; 22.7 %) and Level 3 (Claim+ (Data/ Warrant) + (Qualifier/ Backing); 523.6 %) were quite similar. Only 1 student (0.4 %) provided in Level 4 (Claim+ (Data/ Warrant) + Qualifier+ Backing).

Table 9. Frequencies and Percentages in Argumentation Skill in the Second Science Task

Category	Frequency	Percent
No argument	17	7.0
Level 1	55	22.7
Level 2	112	46.3
Level 3	57	23.6
Level 4	1	0.4
Total	242	100

Note. Level 1= Claim, Level 2= Claim+ (Data/ Warrant), Level 3= Claim+ (Data/ Warrant) + (Qualifier/ Backing), Level 4= Claim+ (Data/ Warrant) + Qualifier+ Backing

6.2 Findings of the statistical data analyses

The data of this study were analyzed by calculating the correlation coefficients, Wilcoxon signed ranks test, and Mann-Whitney U tests. These analyses were conducted in this order. The decisions of conducting non-parametric tests were taken after computing Shapiro-Wilk Normality tests as seen in Table 10. Data from the mathematics achievement, argumentation, proof evaluation and proof construction levels were not normally distributed.

Table 10. Shapiro-Wilk Normality Test Results or the Performances of Students in Mathematics Achievement, and on Argumentation, Proof Evaluation and Proof Construction Tasks

Variables	Statistic	*df*	*p*
Math. Achievement	.893	238	.001
Argumentation	.959	242	.001
Proof Evaluation	.908	242	.001
Proof Construction	.891	242	.001

6.2.1 Findings of the correlational analyses

Correlation coefficients were calculated to investigate the following question: Are there any statistically significant relationships between mathematics achievement, proof construction levels, proof evaluation levels and argumentation levels of students for each task? Since the variables were not distributed normally, Spearman's rho was computed.

Mathematics achievement was found to be statistically significantly correlated with the proof construction skill in geometry task (r_s=0.58), proof construction skill in average (r_s =0.61), argumentation skill in mathematics (r_s =0.51) and argumentation skill in average (r_s =0.53). Also, mathematics achievement was found statistically significantly weakly correlated with proof construction skill in algebra task (r_s =0.26), proof evaluation skill in algebra task (r_s =0.24), proof evaluation skill in geometry task (r_s =0.40), proof evaluation skill in average(r_s =0.42) and with two argumentation skill tasks in science (r_s =0.32, r_s =0.29) respectively.

Other significant correlational findings of this study were found as there exists a statistically significant correlation between proof construction skill and argumentation skill (r_s =0.38). A weak correlation was found between proof construction skill and proof evaluation skill (r_s =0.25) as well. Proof evaluation skill and argumentation skill was found to be statistically significantly correlated (r_s =0.31). All other correlations between variables of this study can be found in Appendix G.

6.2.2 Findings of the Wilcoxon signed ranks tests

Two Wilcoxon signed ranks tests were conducted to answer the following question: Are there any statistically significant differences between students' proof construction and proof evaluation levels for algebra and geometry proof tasks? The comparisons of students' proof schemes in proof construction and proof evaluation tasks with frequencies and percentages gave the impression of that students' proof evaluation skills were better than their proof construction skills. Hence, these two Wilcoxon signed ranks test were conducted to see whether this impression was statistically significant.

Students' performances while constructing and evaluating proofs were compared by Wilcoxon signed ranks tests for both algebra and geometry tasks. For the first proof task (algebra), students performed better in the proof evaluation part (M=2.02) than in the proof construction part (M=1.85), (Z= -2.847, $p < .05$). For the second proof task (geometry), students performed better in the proof evaluation part (M=2.33) than in the proof construction part (M=.88), (Z= -11.671, $p < .05$) (Table 13). Hence there exist statistically significant differences between proof evaluation and proof construction performances of participants as favoring the former.

Table 11. Wilcoxon Signed Ranks Tests for Two Proof Tasks

	Z	*df*	*p*
PC1-PE1	-2.847	241	.004
PC2-PE2	-11.671	241	.001

Note. PC1, PC2: the first and the second proof construction tasks; PE1, PE2: The first and the second proof evaluation tasks; $p<.05$

6.2.3 Findings of the gender differences

Mann-Whitney U tests were performed to see whether there exists a statistically significant difference between female and male students in mathematics achievement, proof construction, proof evaluation and argumentation scores to respond the fourth and sixth research questions: " Are there any gender differences in students' mathematics achievement, proof construction levels and proof evaluation levels?" and " Are there any gender differences in students' argumentation levels?" According to the results, no statistically significant gender differences were found on mathematics achievement, U= 6384, p=.279, with a mean rank score of 125.02 for male students and of 115.29 for female students. No statistically significant gender differences were found on proof construction scores, U= 6711, p=.333, with a mean rank score of 116.81 for male students and of 125.15 for female students. No statistically significant gender differences were found on proof evaluation scores U = 6449, p=.148, with a mean rank score of 128.66 for male students and of 115.92 for female students. No statistically significant gender differences were found on argumentation scores U =6398, p=.128, with a mean rank score of 113.85 for male students and of 127.46 for female students.

6.3 Findings of the qualitative data analyses

Qualitative analyses were conducted to elaborate on the characteristics of students' proofs and arguments and to reveal information on factors that might affect students' proof evaluations. Characteristics of students' proofs and arguments were obtained through looking into themes and codes in each proof scheme and argumentation level separately for each task. Common and distinctive characteristics were reported for both proof schemes and argumentation levels. Factors leading to students' proof schemes in proof evaluation tasks were attained through simply recording students' reasoning. These factors were put in a table with their frequencies.

Qualitative data analyses of this study consisted of examination of students' responses to seven items in detail. Students' performances in proof construction tasks, proof evaluation tasks and argumentation tasks were elaborated one by one. In the first sub-section (6.3.1) students' performances in proof construction tasks were analyzed to reveal information on characteristics of students' proof schemes. In the second sub-section (6.3.2) factors leading to students' proof evaluation schemes were detailed. In the final sub-section (6.3.3) students' performances in argumentation tasks were elaborated for acquiring information about characteristics of argumentation levels of the sample.

6.3.1 Characteristics of students' proofs in the proof construction tasks

In this section, students' proof schemes were reanalyzed to reveal information about the characteristics of proofs in both algebra and geometry proof construction tasks for each proof scheme to respond the following research question: "How are the performances of students in proof construction tasks? What are the characteristics of students' proof schemes?". It was intended to observe whether there exists a similarity between approaches towards algebra and geometry proof construction tasks for each proof scheme. Accordingly, students' performances were detailed by their proof schemes and by quotes from their original responses in the following paragraphs. Original student responses were selected to be the best representation of each characteristic and they were given in Appendix H.

In “No Proof” category, students’ answers presented some hints about why they could not produce any kind of proof scheme in both proof construction tasks. The most frequently observed characteristics of these responses were lack of knowledge or misconceptions about the content of the task. This was apparent in the responses of fourteen students out of twenty-one in the algebra task and seventy-four students out of one hundred and seventeen students in the geometry task. Inadequate knowledge and misconceptions involved use of addition and multiplication interchangeably for the algebra task as in this example: “I think the sum of the two odd numbers can be equal to an even number or it may not be equal to an even number. For example: 3x2=6 and 3x5=15 (Appendix H, 1) and use of unrepresentative real-life examples for even and odd numbers for the algebra task as in “When singles are come together they make a couple [there are drawings of stickman: a boy and a girl]” (Appendix H, 2), “First of all, suppose you divide an apple into two pieces. Then, take one half into your right hand and the other into your left hand. Think this half of apples as odd numbers. Now, we have odd numbers in our both hands. When we combine them, we construct a whole. If you think the whole as an even number, you would prove it.” (Appendix H, 3).

On the other side, assigning mathematically meaningless numbers to angles as in “If 1+2+3+4=20 and 5+6+7+8=20… then each number is calculated out of ten. All are equal.” (Appendix H, 4), having incomplete information about the properties of angles as in “Since exterior angles are equal...” (Appendix H, 5), considering the parallelism between the lines to assert that the given angles are congruent as in “Since line m is parallel to line n, obtuse angle is equal to the obtuse angle and acute angle is equal to the acute angle.” (Appendix H, 6) were the main evidences for students’ inadequate knowledge in the content and terminology of the geometry task.

Another characteristic of “No proof” category responses was the lack of explanation or justification. Students either accepted the given statements or rejected them without presenting any reason. This happened in two cases in the algebra task; in eleven cases in the geometry task as in “…No need to prove, actually. Science is always right.” (Appendix H, 7)

A lot of students who were assigned to “No proof” category provided no answer, they just left the blank space empty. Five students out of twenty-one had these kinds of answers in the algebra task; twenty-three students out of one hundred and seventeen for the geometry task.

External conviction proof scheme was observed in ten students’ responses for the algebra task and in seventy-three students’ responses for the geometry task. The main and the most obvious characteristics of these responses were dependence on the rules and on the appearance. While mentioning the rules to prove the given statement, students provided different approaches for algebra and geometry tasks. They either simply stated the rules as: “In mathematics, there are formulas. With their help, this result has been reached. This question is one of the basic points of the mathematics (I think.)” (Appendix H, 8), “… (example) …This is either a mathematical rule or completely a coincidence.” (Appendix H, 9), “…They are equal to each other, and they are alternate angles…We learned that, and it is known.” (Appendix H, 10), “Interior and exterior alternate angles are equal…Certain rules cannot be changed.” (Appendix H, 11), or they referred to the other rules in the algebra task as “This is a rule. For example, … and there exist other rules like this in mathematics.” (Appendix H, 12), “This is exactly the same as (-)x(-) = (+)” (Appendix H, 13). On the other hand, students named the rules in incomplete or incorrect ways or they came up with new names for expressing the importance of rules as in “According to the rule of Z, alternate exterior angles are equal to each other.” (Appendix H, 14) and in “Since, in parallel lines, the angles of corresponding places are mathematically equal.” (Appendix H, 15)

In empirical proof scheme responses, the dominant characteristics were the use of examples and the dependence on measurement in both algebra and geometry tasks. 82.2 % of the students constructed proofs in Empirical Proof scheme in the algebra task whereas only 7% of students provided proofs in this scheme in the geometry task. Hence, responses in the algebra task brought about more information about the characteristics of empirical proof scheme responses.

The examples provided in the algebra task showed variety in the number of examples and in use of various digit numbers. In other words, some of the empirical proofs had few examples with small digit numbers while some empirical proofs had many examples with multi-digit numbers in. Figure 4 represents a student answer in which there existed only one numerical example “ 1 (odd) + 1 (odd) = 2 (even) :

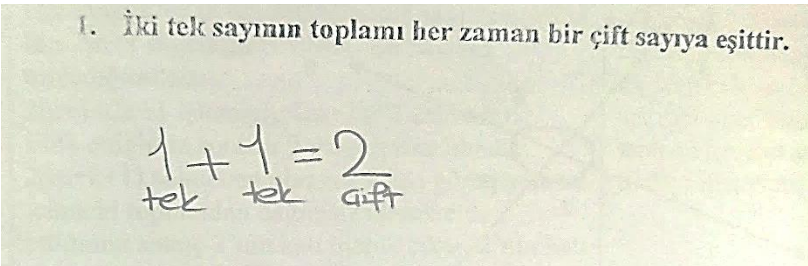


Fig. 4 Empirical proof done with one example

On the other hand, there were beliefs about increasing the number of examples for presenting more comprehensive explanations. In some of them, use of a lot of numbers with multiple-digits was emphasized as in the following excerpt:

> …We can give more than one example about this topic. And these examples prove the correctness of this information. For example, …These examples

could be millions because we have a lot of odd numbers …If we explain by examples we can make sure that it remains better in one's mind. (Appendix H, 16)

In some cases, students did not hesitate to give examples with various numbers and various digits as in Figure 5, the student provided examples with one, two and three-digit numbers and commented that no matter how many digits there are the result is always an even number when we add two odd numbers:

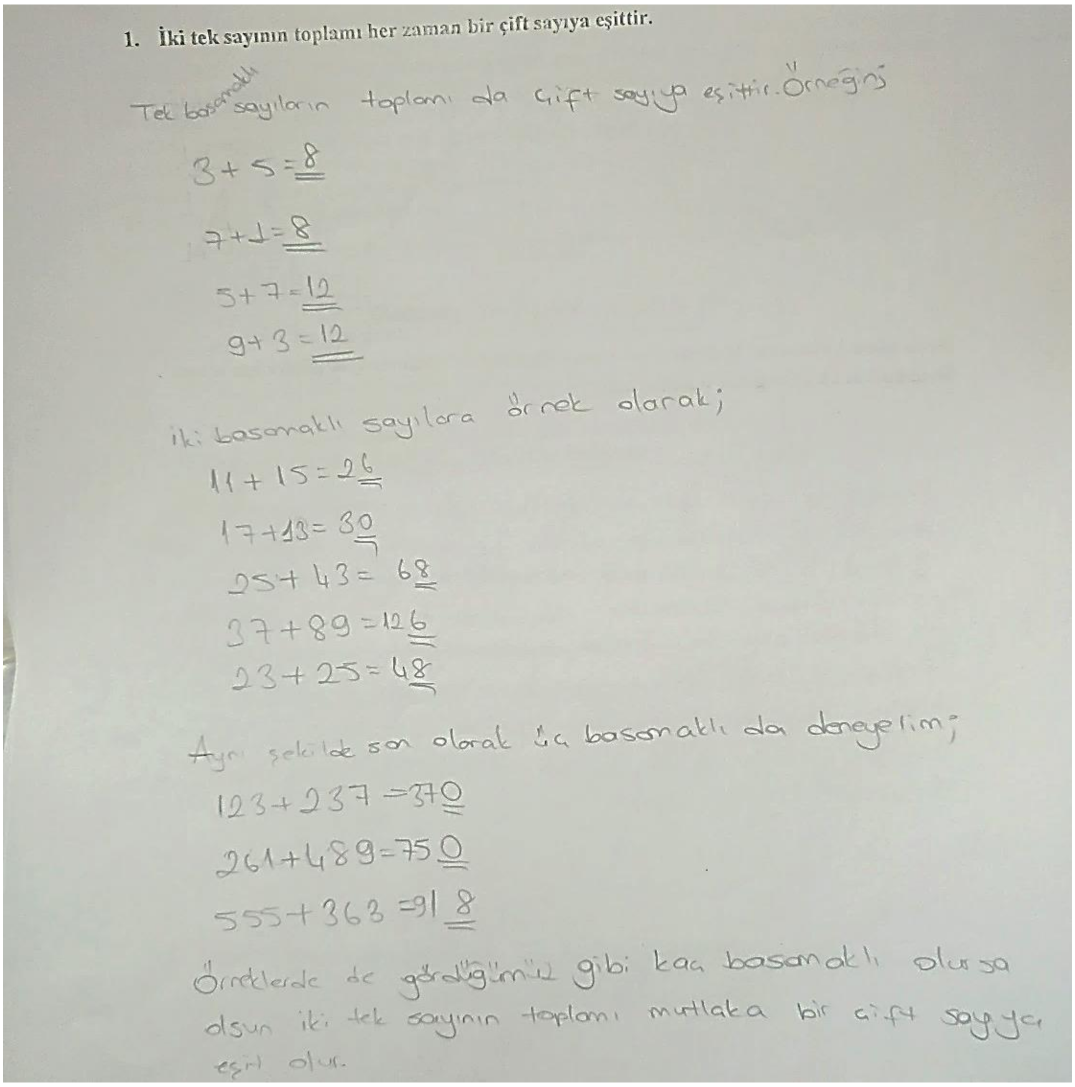
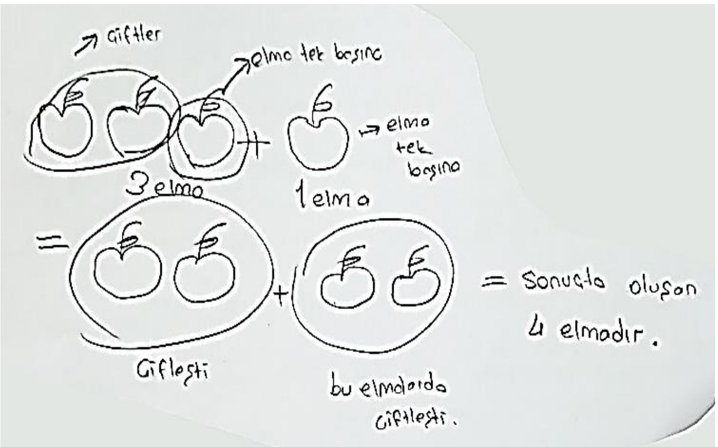


Fig. 5 Empirical proof done with a variety of examples

Another important characteristic of empirical proof scheme responses was the emphasis on the use of examples to prove because it was the only way they knew: "I can prove this only through examples." (Appendix H, 17), "…In any case, there exists a formula, but I do not know." (Appendix H, 18), "I am not good at explaining a topic, therefore I will give an example." (Appendix H, 19)

In analytical proof scheme responses, the main characteristic was the use of deductive reasoning to reach a general claim both in algebra and in geometry tasks. The other characteristics of the responses were the use of representations on concrete objects or benefiting from visual representations, appealing to verbal expressions rather than symbolic representations, providing numerical validations after presenting deductive reasoning and inadequate information about terminology. In the following excerpt, a student in the sample tried to explain his proof through concrete objects, apples, to show grouping of apples which were 'left alone' before:

> It is correct because 3+1=4. The sum of two odd numbers is always an even number…Its proof is: I have three apples. When I have one more apple, both pair with each other and it becomes four. I mean: (Appendix H, 20)



(Appendix H, 20)

Inadequate symbolic representations and appealing to verbal explanations can be observed in the following student response. Here, the student could not express the even numbers as multiples of two, in an algebraic form. Rather, she verbally explained that the odd numbers are in the form of 2k+1:

> All even numbers are divisible by two. In other words, they can be written in groups of two. For example; Four can be divided in groups like 2+2. Odd numbers are not divisible by two and cannot be written in groups of two. For example; five can be written as 2+2+1. When we want to divide them into such groups, 1 remains…Now let's add two odd numbers: 7 (2+2+2+1) + 9 (2+2+2+2+1) = 2+2+2+2+2+2+2+ (1+1) =16 is an even number. (Appendix H, 21)

Inadequate terminology about the content of the geometry task was observed in many student responses for the analytical proof scheme. The following excerpt could be shown as an example in which a student showed a deductive approach without using proper terminology, she used "opposite angles" instead of alternate angles. She also expressed "in the same direction" for corresponding angles :

> In the parallel line segment, opposite angles are equal. For example, 6 and 7, 2 and 3. Also, the angles in the same direction are equal: 2 and 6, 3 and 7. According to these, 2 and 7 becomes equal (Appendix H, 22)

In summary, students presented no proof if they did not have adequate knowledge or had misconceptions about the content of the task. Many students could not provide any justification or explanations. Students who provided proofs in external conviction proof scheme depended on rules or appearance. They referred to the rules without explanations or they emphasized other rules to show that what matter there were the rules. While expressing the importance of rules, most of the time, students had deficiencies in the use of proper terminology of the rules. Students who presented proofs in empirical proof scheme mainly used examples. The differences between empirical proofs were observed in the number and in the kind of examples that students provided. All proofs in analytical proof scheme had deductive reasoning. In most of them, verbal expressions were salient rather than symbolic ones.

6.3.2 Factors leading to students' proof schemes when evaluating proofs

In this section, students' responses about the reasons for selecting the most convincing proof were analyzed to respond the following research question: "What are the factors which make students convinced in proof evaluation tasks?". The factors which led students to prefer a proof among three options were summarized for each proof scheme and for two tasks with their frequencies in Table 12.

According to the Table 12, there were seventeen factors expressed by students as their rationales for selecting a proof among three options. These factors were labeled from F1 to F17 as:

- F1: Closer to what I think, the way I did
- F2: Explanatory and or Understandable
- F3: Simple/ Easy/ Not Complicated
- F4: More logical
- F5: There exists a rule/ Use of rule-formula
- F6: There are evidences/ experiments/ examples
- F7: The way we were taught
- F8: Scientific
- F9: Detailed
- F10: Short/ practical
- F11: Long
- F12: There is a reasoning/ no need for memorization

Table 12. Factors That Convinced Students for Selecting the Best Proof

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17
1A	9	15	10	5	13	2	2	1	1	1	0	0	0	0	0	0	0
1G	10	12	8	3	0	0	0	1	0	10	0	0	0	0	0	0	0
2A	24	43	49	8	0	58	0	0	0	3	0	0	0	0	0	0	0
2G	5	19	2	1	0	36	0	2	0	4	0	0	0	1	0	3	0
3A	2	10	0	7	6	7	0	13	2	1	1	5	9	12	2	0	0
3G	17	46	0	14	32	2	5	17	4	0	5	16	2	10	8	5	4
Tot.	67	145	69	38	51	105	7	34	7	19	6	21	11	23	10	8	4

Note. 1A: External Conviction Proof Scheme- Algebra; 1G: External Conviction Proof Scheme- Geometry; 2A: Empirical Proof Scheme- Algebra; 2G: Empirical Proof Scheme- Geometry; 3A: Analytical Proof Scheme- Algebra; 3G: Analytical Proof Scheme- Geometry;

F1: Closer to what I think, the way I did; F2: Explanatory and or Understandable; F3: Simple/ Easy/ Not Complicated; F4: More logical; F5: There exists a rule/ Use of rule-formula; F6: There are evidences/ experiments/ examples; F7: The way we were taught; F8: Scientific; F9: Detailed; F10: Short/ practical; F11: Long; F12: There are reasonings/ no need for memorization; F13: Comprehensive/ general; F14: Mathematical; F15: There are more & stronger information; F16: Certain/ Objective; F17: Step-by-step

- F13: Comprehensive/ general
- F14: Mathematical
- F15: There is more & stronger information
- F16: Certain/ Objective
- F17: Step-by-step

There were some factors which were stated mostly as criteria for selection of the most convincing proof. Being explanatory or understandable was the most commonly used reason for selection, as observed in one hundred and forty-five cases. The existence of concrete evidences and examples from operations and experiments was another factor which was indicated in one hundred and five cases. Being closer to what s(he) think or his/her method and being simple and easy were the other factors which were used in sixty-seven and sixty-nine students' responses respectively.

The least expressed factor for proof evaluation was being presented step-by-step (in four cases). This was followed by being long. Only in six cases, length was given as a factor for convincing proofs. Being detailed was also one of the least indicated factors for convincing arguments. This was observed only in seven cases.

As in the Table 12, some factors appeared in certain proof schemes. For instance, factors which were labeled from F11 to F17 (long, use of reasoning, comprehensiveness, being mathematical, having more information, being certain and being presented as step-by-step) were mainly dominant in Analytical proof schemes. On the other hand, there were three factors being stated in every proof scheme: being closer to their method or their idea, being explanatory and being more logical.

Table 12 also provides information for comparing whether the factors asserted by students for each proof scheme had differences between the kind of task; algebra or geometry. For example, even though both were classified as external conviction proof scheme, students' criteria for selecting them differed: Use of rules or formulas were presented as a reason to select the external conviction proof scheme response in thirteen cases of algebra task; however, no one reported that as a factor in geometry task. Likewise, shortness or practicality was indicated as a factor in ten cases of external conviction proof scheme geometry task, but it appeared in only one case in external conviction proof scheme algebra task. Also, simplicity and easiness were expressed in forty-nine cases in empirical proof scheme in the algebra task but appeared in only two cases in the geometry task. These findings show that the domain of the task (algebra or geometry) brought about different rationales with different frequencies while evaluating proofs.

In summary, seventeen factors were detected from students' responses. Being explanatory and understandable was the most stated reason for selection of proof schemes. The existence of evidences and examples, being logical and closeness to their idea or their answer were the other main factors viewed in their responses. Also, students' criteria differed in geometry and algebra tasks. There were some factors associated specifically with some proof schemes whereas some of them were observed in every proof scheme.

6.3.3 Characteristics of students' arguments in the argumentation tasks

Students' responses in the argumentation tasks were initially analyzed quantitatively through using a scheme developed by Venville and Dawson (2010) which takes the components of arguments into its focus. Accordingly, students may have four different argumentation levels. In the first level, there exits only a claim in the argument. In the second level, evidence(s) is also included on top of a claim. In the third level, a claim, a data/ a warrant as evidence and a qualifier or a backing constitutes the argument. In the fourth-level, a claim, a data/ a warrant, a qualifier and a backing are included altogether. In this section, students' arguments in each argumentation task were reanalyzed by taking argumentation levels into consideration. In other words, characteristics of level 1, level 2 and level 3 arguments for each task were detailed in this part to respond the following research question: "How do the students perform in argumentation tasks? What are the characteristics of students' argumentation levels?"

The level 1 arguments consisted of only a relevant claim. In these arguments, students either did not present any evidence to support their claims or they tried to provide data or a warrant which were not supporting the claim. The excerpts "It may not be correct. It depends on the value that we assign for x." (Appendix H, 23), "Meltem. Sugar is in the water, but it is invisible." (Appendix H, 24), and "Making clean water through heating the water vaporizing it then passing over a cold object." (Appendix H, 25) could be shown as examples for arguments that includes only a relevant claim. On the other hand, the excerpts like "x^2 is not always greater than x. It could be x^2=70 and x=71" (Appendix H, 26), "If we assign values with minus sign to this equation it becomes correct. For example, let x be -5.-5^2<-5=-25<-5. As you see, it becomes correct." (Appendix H, 27), "Meltem is right… The reason why we don't see the sugar is that the amount of water is much more than when compared to the amount of sugar." (Appendix H, 28) , and "The first method because in the membrane method the salt remains even though it passes through the filter since the salt is dissolved" (Appendix H, 29) were examples of level 1 arguments whose evidences did not support the claims.

Level 1 arguments with unsupportive evidences were formed because students had little or no information about the content of the tasks. In the excerpt "x^2 is not always greater than x. It could be $x^2=70$ and $x=71$" (Appendix H, 26), the student took the x's as two distinct variables and assigned them different values to support her claim. She apparently had difficulties in algebraic expressions. The excerpt "If we assign values with minus sign to this equation it becomes correct. For example, let x be $-5.-5^2<-5=-25<-5$. As you see, it becomes correct." (Appendix H, 27) also involved misconceptions about taking power of negative numbers when unknowns were included. The student might not evaluate the negative numbers as numbers with negative values but as minus-signed positive numbers.

The science level 1 arguments also had irrelevant justifications for the claims. In the excerpt of "Meltem is right… The reason of why we don't see the sugar is that the amount of water is much more than when compared to the amount of sugar." (Appendix H, 28), the student attributed the invisibility of sugar in hot water to the excessive amount of water when compared to the amount of sugar, which was not the correct reason. Moreover, the excerpt of "The first method because in the membrane method the salt remains even though it passes through the filter since the salt is dissolved" (Appendix H, 29) provided us information about how a misinterpretation can lead someone's arguments to be affected negatively. In this case, the student preferred one method to another just because he believed that one is not effective even though it was not represented like that in the given text. Overall, lack of evidences, having irrelevant or irrational evidences were the main characteristics of level 1 arguments.

Level 2 arguments involve a claim and supporting evidence, which is named as data or warrant, only (Venville & Dawson, 2010). In mathematics argumentation task, the evidences were mainly the numerical examples. In science argumentation tasks, most of the evidences were given in the tasks and students had an opportunity to connect them to their claims. Having already set-up evidences contributed easiness to form level 2 arguments in science argumentation tasks. On the other hand, students were required to connect their existing knowledge to the given information in the mathematics task. This situation could be observed in the percentages of students in constructing level 2 arguments.

Quantity of the evidences and source of the evidences were the distinctive features for argumentation tasks where evidence is presented explicitly or implicitly. In most of the level 2 arguments, students presented their data explicitly as can be seen in these mathematics excerpts: “Let x be 2; $2^2<2$ $\cancel{4<2}$. Let x be 3; $3^2<3$ $\cancel{9<3}$. Let x be 4; $4^2<4$ $\cancel{16<4}$. As you see they are all wrong.” (Appendix H, 30) and in science excerpts “The evaporation should be done. They don’t have a right to harm any living thing.” (Appendix H, 31). On the other hand, some students preferred to justify their claims through implicit explanations: “Meltem. What Leyla said and what their teacher said contradicts.” (Appendix H, 32), “The person who wrote the point that I marked is right. Meltem is right.” (Appendix H, 33)

One of the most salient features of level 2 arguments was about the differences in the quantity of the evidences that students used. More explicitly, while some students relied on only one evidence like in the following excerpts; “I don’t find it correct. For example, let x be 6; x^2=6x6=36. How is x^2 smaller than x?” (Appendix H, 34) , “Meltem because the sugar is still in the water. Sugar and water are intertwined so the sugar may be seen as invisible and sugar gave its taste to water." (Appendix H, 35), some others used variety of evidences to support their claim: "To me, Meltem is right because it is like what their teachers' explained. That matter cannot be disappeared. When we mixed sugar and water, the water just absorbs the sugar and we can't see the sugar" (Appendix H, 36).

Another distinctive characteristics between level 2 arguments were about the source of evidence that had been used to support their claims. By source of evidence, it is meant that whether the information existed already in the task or it was provided as student's self-knowledge. Majority of the evidences were in science tasks, whereas some students preferred to back up their claims using their self-knowledge.

Level 3 arguments were encountered when students had their claims, data or warrant, backing or qualifier in their arguments (C+(D/W) +(B/Q)). In other words, on top of their claim and evidences, students who produced level 3 arguments either presented their certainty through qualifiers or they fed their evidences with additional evidences called as backing. The percentages of level 3 arguments constituted 14%, 3.3 % and 23.6% of the responses in mathematics and science argumentation tasks respectively.

Students' responses differed in the use of qualifiers or backings for level 3 arguments. In mathematics tasks, for instance, there were mainly qualifiers in these arguments. In science tasks, however, backings were more dominant in level 3 arguments. In the cases where students involved qualifiers in their arguments, there were explicit or implicit indicators of certainty.

> This expression is invalid for natural numbers. For fractions it may be possible, but its truth is questionable. But it is sometimes correct: If we let x be -1; 1<-1 but this statement is wrong. If x=1/2, then 1/4<1/2 and this becomes correct. So, it is not always valid. (Appendix H, 37)

The quote above was one of the level 3 arguments where the qualifier was explicitly indicated in the last line of the explanation. The expression " It is not always valid" served both as a claim and a qualifier. The student explained his idea in the first sentences superficially. Then he added numerical evidences and involved adverbs of frequency (sometimes and not always) to support his claim. He finalized and made clear his argument with repeating his claim and qualifier.

> This statement is a true rule for some numbers, wrong for others. For instance, if x=2; since 4<2 this rule will not be met. But a rational number can meet this rule. For instance, if x=1/3, $(1/3)^2=1/9$. Since 1/9<1/3, it can meet this rule. (Appendix H, 38)

The explanation above was level 3 argument, too. It involved numerical data however the claim and the qualifier were not directly stated. In other words, they were absent by appearance, but they were available in the meaning. "True for some numbers, wrong for others" can be thought as sometimes true or sometimes wrong, where "true" or "wrong" is a claim and "sometimes" is a qualifier.

Some of the level 3 arguments involved backing rather than qualifiers. Majority of the arguments for level 3 science tasks had backings on top of claims and warrants. These backings were mainly additional support or further suggestions to their evidences.

> I think Meltem is right because when the sugar is added into hot water, it slowly melts, and it seems united with water but after a while the sugar can reappear when they are evaporated. As the teacher says, matter cannot be created or destroyed. (Appendix H, 39)

The quote above served as a level 3 argument which has a backing. The student claimed that Meltem is right and he supported this claim through a scientific fact "Matter cannot be created or destroyed" , a warrant. He also presented his own explanation that the sugar was melted in the hot water that is why it is invisible. The backing of his argument was "…but after a while the sugar can reappear when they are evaporated.", which were supporting the scientific fact that the teacher provided.

> The second one; because in the evaporation method it is done without hurting animals. About energy, renewable energy fuels can be used. Why not the first one: Here, animals get hurt and this can result in extinction of their species. The second one is better. (Appendix H, 40)

The excerpt above was another level 3 argument that contained a backing. This student took advantages and disadvantages of two desalination methods into consideration: The membrane method was harmful for ocean animals whereas the evaporation method was too expensive for California. The expression "The animals can get hurt and this result in extinction of their species" involved both data and an interpretation. In the task, there was no statement for the animal extinction, but some students like him, asserted that. This interpretation was evaluated as backing since these expressions were involved more than data or warrant. "I would pick the second method because the membrane method harms ocean animals. If animals get harmed, the balance of the nature becomes destroyed. If the balance of the nature is destroyed, the balance of the earth is destroyed" (Appendix H, 41). In this excerpt the student not only used the information of "The ocean animals get harmed", he also elaborated it with possible consequences to support her evidence.

In mathematics argumentation task, there were level 3 arguments that contained backing, too. "Let x be ½. In this case, the number becomes smaller as the denominator gets bigger. The reason is that things get smaller when we divide them into much more pieces." (Appendix H, 42) In this quote, the data and the claim of the argument were implicit since the student indicated that this statement is true for unit fractions by only giving example of a unit fraction as ½. The student did not even show the square of ½ as ¼. But s(he) pointed out that ¼ is smaller than ½ through a warrant: "…the number becomes smaller as the denominator gets bigger.". In this expression, the student used a rule for comparisons of fractions with the same nominator but different denominators: When two fractions share the same nominator but have different denominators, the one with the smallest denominator is greater than the other fraction. The student continued his/her argument with a rationale that explains this warrant: Division into more pieces of an entity makes the parts smaller in size or in amount. Hence, the rationale was counted as a backing.

Overall, arguments in mathematics and science tasks showed similarities and differences in characteristics. In level 1 arguments, it was seen that students either did not have any evidence or their evidences could not support their claims. This happened due to students' lack of information, their misconceptions and their irrelevant or irrational evidences. In level 2 arguments, there were explicit and implicit evidences. Students' responses in level 2 arguments varied based on the number of evidences that they used. The sources of evidences differed in the students' responses in level 2 arguments. Level 3 arguments had a claim, evidences and backing or qualifier. These arguments differed in terms of whether they had a backing or a qualifier within. For all argumentation levels, it was observed that having difficulties and misconceptions about the content and having insufficient or wrong information led the level of arguments to decrease.

CHAPTER 7

DISCUSSION, IMPLICATIONS AND FUTURE RESEARCH

7.1 Summary of research findings

Research findings of this study were based on descriptive, statistical and qualitative analyses. It was found that students performed and favored empirical proofs for algebra task. For the geometry proof task, most of the students produced no proof and there were more students who presented analytical proof rather than empirical proof. Most of the students preferred the empirical proof scheme to be more convincing when they were given three proof schemes in the algebra task; whereas analytical proof scheme was the most frequent option in the geometry task.

Students' arguments were mostly in the form of Level 2 arguments, arguments consisting of a claim and a supporting evidence, data or warrant. There was only one student in the sample who provided the highest level of arguments (level 4).

Correlational analyses revealed that mathematics achievement is statistically significantly related with all variables in the study. Also, statistically significant correlations were found among proof construction, proof evaluation and argumentation skills.

Comparison of the proof evaluation tasks and proof construction tasks revealed that students' performances in proof evaluation tasks were statistically significantly better than in proof construction tasks. Also, gender differences were not statistically significant for any of the variables.

In qualitative analyses it was observed that students' proofs had some shared and distinctive characteristics for each proof scheme. Accordingly, students presented no proof if they did not have adequate information or had misconceptions about the content of the task. Many students could not provide any justification or explanations. Students who provided proofs in external conviction proof scheme depended on rules or appearance. They referred to the rules without explanations or they emphasized other rules to show that what matter there were the rules. While expressing the importance of rules, most of the time, students had deficiencies in the use of proper terminology of the rules. Students who presented proofs in empirical proof scheme mainly used examples or depended on measurements. Some of these students explicitly indicated that their responses were formed just because it is the only way they have known. The differences between empirical proofs were observed in the number and in the kind of examples that students provided. All proofs in analytical proof scheme had deductive reasoning. In most of them, verbal expressions were salient rather than symbolic ones. Use of concrete objects as representations, providing cross-checks with numerical data and lack of information about terminology of the topic of tasks was some of the characteristics of proofs in analytical proof scheme.

Examination of proof schemes while students evaluating proofs revealed that there were seventeen main factors leading students' proof schemes. Being explanatory and understandable was the most stated reason for selection of proof schemes. Also, the existence of (concrete) evidences and examples was indicated in a lot of student explanations. Being "logical" and closeness to their idea or their answer were the other main factors viewed in their responses. It was observed that the topic of the task brought about different reasons or factors for selecting the most convincing proofs. In other words, students' criteria differed in geometry and algebra tasks. There were some factors associated specifically with some proof schemes whereas some of them were observed in every proof scheme. Being closer to their method or their idea, being explanatory, being simple and having evidences were the most dominant factors expressed by the students for the most convincing arguments.

In the analyses of students' performances in argumentation tasks it was observed that students' argumentation levels for each task had common and distinct aspects. In other words, arguments in mathematics and science tasks showed similarities and differences in characteristics. In level 1 arguments, it was seen that students either did not have any evidence or they their evidences could not support their claims. The possible reasons could be students' lack of information, their misconceptions and their irrelevant or irrational evidences.

In level 2 arguments, there were evidences but some of them were explicit and the others were implicit. Students' responses in level 2 arguments varied based on the number of evidences that they used. For instance, some students confined themselves to only one evidence whereas some students made use of more than one evidence. The sources of evidences differed in the students' responses. In the mathematics argumentation task, students had to use their self-knowledge as evidence, however in science tasks students made use of both their current knowledge and the given information in the tasks.

Level 3 arguments had a claim, evidences and backing or qualifier. These arguments differed in terms of whether they had a backing or a qualifier within. In majority of the mathematics argument, there was a qualifier rather than a backing. In science tasks, on the other hand, backing was the most frequently observed component of an argument.

For all argumentation levels, it was observed that having difficulties and misconceptions about the content and having insufficient or wrong information led the level of arguments to decrease. Consideration of various kinds of evidences led students to establish more rigorous arguments. For instance, students who provided examples from only one number set (e.g., natural numbers, whole numbers) had lower level of arguments because the lack of holistic examination of all possible number sets caused students to produce arguments without qualifiers most of the time.

7.2 Discussion

In the relevant literature about students' proof construction skills, it was stated that students mostly produced empirical proofs (Harel & Sowder, 1998; Healy & Hoyles, 2000). This situation was observed in this study for the algebra task. However, in the geometry task, empirical proof scheme was the least used proof scheme. Majority of the students presented no proof for the geometry task. These findings could be interpreted as the type of the task could make a difference in students' proof schemes. Students might have different proof schemes for different tasks (Tall et al., 2011; Waring, 2000). The reason that most of the students failed to construct any kind of proof in geometry task could be that they had less experience with the content of the geometry task. The property of angles appears in the seventh-grade mathematics curriculum; students may have had little practice about it or they could not internalize these properties well enough. The content of the algebra task (even and odd numbers) is more known and internalized by students.

Students' performances in proof construction tasks demonstrated that their conceptions, misconceptions, lack of knowledge, and previous experiences had impact on their proof schemes as indicated in the literature (de Villiers, 1990; Healy & Hoyles, 2000). Students who presented empirical proofs in the proof algebra construction task had different explanations with a diverse number of examples and different digit numbers. Students who provided a lot of examples with a variety of numbers could be viewed as having an understanding that a proof should encompasses all possible cases. These students differ from their peers in terms of perception of proof. They might have believed that a proof should be general and be valid for even extreme cases. Some students who produced empirical proofs indicated that their responses were not proof but they could construct proofs only through examples. This showed that these students might have an insight about what should a proof be, but they did not have methodological knowledge to present analytical proofs since they had not had any experience in proof construction.

It was observed that students who presented analytical proofs used verbal expressions rather than symbolic expressions. They utilized concrete representations in their explanations. Some of them used numerical examples to crosscheck their arguments and had difficulties in use of correct terminology. These findings are compatible with what Piaget (1985) suggests for students in middle school ages: They are at the transition phase from concrete operational stage to formal operational stage. Hence, they may show characteristics of both cognitive developmental stages. Having difficulties in symbolic representation and appealing numerical validations after their explanations reflected the characteristics of concrete operational stage. On the other hand, reaching general judgements through verbal expressions which resemble algebraic expressions can be evaluated as these students show characteristics of formal operational stage as well.

Students' proof schemes while constructing and evaluating proofs were found to be different from each other as suggested in the literature (Harel & Sowder, 1998; Healy & Hoyles, 2000). This was observed in geometry task more clearly. While the majority had no proof in the construction task, most of them evaluated the analytical proof as the most convincing in proof evaluation task. This finding could be interpreted as students may have conceptions or ideas about what the proof or the convincing argument should be, but they might have no experience for proof construction, hence they could not produce any kind of proof scheme.

Students' rationales for selecting the most convincing proof in the proof evaluation task match with the literature findings. Students stated that being explanatory and understandable is important for an argument to be convincing. In some of the proof studies, this finding was emphasized as the role of explanation is perceived as the most important role for proofs (Bieda & Lepak, 2014; Hanna, 2000). Seeking for concrete evidences and examples is quite expected reaction from students since they may show characteristics of concrete operational stages (Piaget, 1985). Students who stated their reasons for selecting a proof as being close to their idea or their method showed that they did not expose various kinds of arguments hence they stick with their own methods.

Students' reasoning in the proof evaluation part was found to be compatible with their proof schemes in the proof evaluation task. For instance, factors of being comprehensive or general were found only in students' explanations in analytical proof scheme. The students who claimed that a proof is convincing when it is based on concrete data or examples, showed that they had empirical proof scheme. This finding shows that the proof schemes framework developed by Harel and Sowder (1998) reflects students' proof skills in an efficient way.

Findings about students' level of arguments were parallel with the literature as majority of the students presented level 2 arguments, arguments that contain a claim and an evidence such as data and/or warrant (Kaya, 2013; Venville & Dawson, 2010). In some cases, students presented their arguments with implicit components. These could be stem from the fact that these students did not get any instruction on argumentation practices. Some students might think that the evidences are so obvious hence they did not need to express them again. Or this could be due to students' lack of experiences in mathematical writing. Moreover, the content of the task, its complexity, and students' effort to express them made differences in students' argumentation skills (Liu et al., 2016).

Students' mathematics achievement, proof skills and argumentation skills were found to be statistically significantly correlated. The finding about relationship between proof and argumentation skills supports some of the literature which suggests a probable relationship between these skills (Reid & Knipping, 2010) and contradicts with the researchers who claim that there cannot exist relationship between them (Carrascal, 2015). This statistically significant correlation could be supported with the view that both of these constructs include justifications and logical processes, hence they have shared characteristics (Reiss et al., 2008).

In some studies, it was asserted that the differences between students' proof schemes can be due to students' gender differences (de Villiers, 1990; Healy &

Hoyles, 2000). Findings of this study did not support this assertion; gender differences were found to be ineffective for the differences in students' mathematics achievement, proof and argumentation skills.

7.3 Implications

This study can provide important implications for learning, teaching, academic research studies and for curriculum purposes.

Analyses of this study revealed that some students could produce analytical proofs without taking any former instruction about proof. More students evaluated the analytical proof to be the most convincing one in the geometry task. These results show that, even if they were not exposed to any instruction on proofs they could somehow appreciate the accepted form of proof. If students had opportunity for proof instruction they may have even constructed analytical proofs.

In the analyses, it was observed that students performed better in geometry proof tasks than algebra tasks in terms of proof evaluation skills. Because students conducted and favored analytical proofs more in geometry tasks and empirical proofs can be constructed more easily in algebra tasks, proof instruction can be best initiated with geometry tasks.

Analyses about argumentation skills indicated that a lot of students produced arguments with claim and evidence only. If the fact that they were not practiced with argumentation tasks and instructions is considered, the performances of these students can be evaluated as expected. So, argumentation practices given in the literature can be implemented in the lessons to see whether they can make a difference in students' argumentation levels.

Finding correlations between proof and argumentation skills can direct researchers to take these constructs together in their studies. To obtain cause and effect relations between them, comparisons of these skills through experimental designs can be done.

In the analyses, it was attained that students' proofs and arguments were affected by their conceptions, misconceptions and current knowledge. Consideration of instruction on proof and/or argumentation studies should supervene conceptual knowledge of the contents.

By considering the importance of proof and argumentation skills and students' skills in proof and argumentation, mathematics teachers should involve practices of proof and argumentation as integrated to other curricular objectives. Hence, proof and argumentation should not be considered as isolated topics. Mathematics teachers should include proof and argumentation activities in their lesson plans and hidden curriculums.

7.4 Limitations and future work

The number of items was limited both for proof and argumentation tasks. This limitation happened due to the time restrictions that the schools impose. Most of the schools were not voluntary to devote school time for research purposes. The stressful environment of high school entrance exam and the rush for teaching all the learning objectives in time led schools to permit only one lesson hour for data collection. Therefore, the argumentation and proof tasks were designed so that they can be completed in one-lesson hour. More thorough investigation could be done with more tasks.

The content of the instrument was limited for the same reasons of limited number of items. The instrument includes five objectives from middle school mathematics and science curriculums. To ensure content validity and comprehensiveness, the number of items could be increased, or the instrument could have focused on a specific topic/objective.

Different schools and different mathematics teachers and their judgements brought limitations for this study in terms of mathematics achievement scores. Having no standardized test for mathematics achievement led the researcher to obtain mathematics achievement scores from their mathematics teachers' evaluations. The lack of same measurement tool for students' mathematics achievement could have brought about subjectivity in their evaluations.

The performances of students in the proof and argumentation tasks might be affected from the fact that they have not taken any instruction on proof or argumentation. That might have produced a limitation for the levels of constructed proofs or arguments as majority of the students produced low level of arguments or proofs.

Studying this issue with all middle school levels can be good to observe whether students are capable in producing deductive proofs and/ or high-quality arguments. The analyses of this study revealed that some of the eight grade students can produce good proofs and arguments. If these students can achieve those in previous years, then an instruction for proof or argumentation could be initiated in earlier years. Therefore, analyses beginning from the early years of middle school can promote new insights about abilities and inabilities of middle school students. These analyses can even reshape the objectives of the lesson plans and curriculums.

An intervention program can be designed by taking proof or argumentation into the focus: A researcher/teacher who wants her students to be good at in producing high-quality arguments through allocating time for proof instructions may observe students' performances before and after an intervention by comparing them. These comparisons can be achieved through having pre-tests and post-tests and generating control groups and experimental groups by conducting an experimental research. Another idea for future work could be studying the effects of the instructions of argumentation onto proof skills. Since argumentation studies can be merged to any subject more easily than proof studies, it could be seen easier to study argumentation to see whether it has any effect on proof skill, through experimental designs. Therefore, obtaining statistically significant correlations among proof and argumentation skills can open doors for further statistical analyses.

Mathematics teachers can also be included as participants into the study. Their ability to prove and produce arguments could be correlated or compared with their students to observe whether there exists a relationship or effect among these. Since students lack the opportunity to produce proofs and mathematical arguments in the middle school, it would be meaningful to learn mathematics teachers' attitudes towards these concepts and teaching them, their self-efficacy about them and their willingness and readiness to teach them. Even though students can be cognitively adequate to learn producing deductive proofs and high-quality arguments, their teachers will be one of the most effective factors to leading these skills. By merging them into their hidden curricula or lesson plans, mathematics teachers can provide students with opportunities to construct proofs and "good" arguments.

Proof tasks had construction and evaluation parts, but argumentation tasks did not in this study. It would be nicer to see students' performances when evaluating others' arguments and the comparisons of constructed and evaluated arguments. Although it does not exist in the literature, the distinction in constructing and evaluating arguments, as in proof tasks, could be a beneficial contribution to the literature.

APPENDIX A

CONTENTS RELATED TO PROOF, ARGUMENTATION AND REASONING IN THE TURKISH MIDDLE SCHOOL MATHEMATICS CURRICULUM

"Bu öğretim programı matematik öğrenmeyi etkin bir süreç olarak ele almakta, öğrencilerin öğrenme sürecinde aktif katılımcı olmalarını vurgulamakta ve dolayısıyla kendi öğrenme süreçlerinin öznesi olmalarını öngörmektedir. Bu bağlamda öğrencilerin araştırma ve sorgulama yapabilecekleri, iletişim kurabilecekleri, eleştirel düşünebilecekleri, gerekçelendirme yapabilecekleri, fikirlerini rahatlıkla paylaşabilecekleri ve farklı çözüm yöntemlerini sunabilecekleri sınıf ortamları oluşturulmalıdır." (MEB, 2013, p. 3)

"… Bu teknolojiler yardımıyla, öğrencilerin modelleme yaparak problem çözme, iletişim kurma, akıl yürütme gibi becerilerinin geliştirilmesine yönelik ortamlar hazırlanmalıdır." (MEB, 2013, p. 3)

"Matematik Eğitiminin Genel Amaçları

Öğrenci,

…

3. Problem çözme sürecinde kendi düşünce ve akıl yürütmelerini ifade edebilecektir.
4. Matematiksel düşüncelerini mantıklı bir şekilde açıklamak ve paylaşmak için matematiksel terminoloji ve dili doğru kullanabilecektir…" (MEB, 2013, p. 4)

“PROGRAMDA KAZANDIRILMASI ÖNGÖRÜLEN TEMEL BECERİLER

Ortaokul matematik öğretim programında matematiksel kavramların kazandırılmasının yanı sıra, matematiği etkili öğrenmeye ve kullanmaya yönelik bazı temel becerilerin geliştirilmesi de hedeflenmektedir. Bu beceriler şöyle sıralanmaktadır:

…

• Matematiksel süreç becerileri:

- İletişim
- Akıl yürütme
- İlişkilendirme

…” (MEB, 2013, p. 5)

“Bu programda, öğrencilerin iletişim becerilerinin gelişimine önem verilmektedir. Bunun için dikkate alınması gereken bazı göstergeler şunlardır:

…

• Matematiğin sembol ve terimlerini etkili ve doğru kullanma

…

• Somut model, şekil, resim, grafik, tablo, sembol vb. farklı temsil biçimlerini kullanarak matematiksel düşünceleri ifade etme

• Matematiksel düşünceleri sözlü ve yazılı ifade etme

…

• Matematiksel düşüncelerin doğruluğunu ve anlamını yorumlama” (MEB, 2013, p. 7)

“Akıl Yürütme: Akıl yürütme (muhakeme), eldeki bilgilerden hareketle matematiğin kendine özgü araç (semboller, tanımlar, ilişkiler, vb.) ve düşünme tekniklerini (tümevarım, tümdengelim, karşılaştırma, genelleme, vb.) kullanarak yeni bilgiler elde etme süreci olarak tanımlanabilir. Akıl yürütme becerisinin okul ve okul dışı hayatı kolaylaştırmadaki etkisi de dikkate alındığında matematik öğretim sürecinde bu becerinin geliştirilmesi için ortamlar hazırlanmasının gerekliliği ortaya çıkmaktadır. Bu nedenle, öğretim programında öğrencilere akıl yürütme becerilerinin kazandırılması için dikkate alınması gereken bazı göstergeler şunlardır:

• Çıkarımların doğruluğunu ve geçerliliğini savunma

• Mantıklı genellemelerde ve çıkarımlarda bulunma

• Bir matematiksel durumu analiz ederken matematiksel örüntü ve ilişkileri açıklama ve kullanma

….” (MEB, 2013, p. 7)

APPENDIX B

RUBRICS FOR THE TASKS

Rubric for Proof Construction Tasks		
Score	Criteria	Example(s)
0	There exists no proof or the proof is constructed wrongly.	“The multiples of zero and two are called as even numbers.”
1	External Conviction proof: The student has no cognitive effort on his/her proof. S(he) only transferred the information that is presented to her/him. S(he) repeats what an authority (teacher, book, friend) said before. S(he) uses the symbolic representations without knowing their meaning and functions.	“This is a rule. For example, … and there exist other rules like this in mathematics.” known.” “Interior and exterior alternate angles are equal…Certain rules cannot be changed.”
2	Empirical Proof: The student gives examples, applies trial and error and representations that lack generalization power.	“I am not good at in explaining a topic, therefore I will give an example.” “The easiest solution is folding. Or use protractor.”
3	Analytical Proof: The student uses deductive reasoning, reaches a general judgement through symbols or algebraic expressions.	“Exterior alternate angles are equal and the angles 2 and 7 are exterior alternate angles, hence they are equal to each other.”

Rubric for Proof Evaluation Tasks	
Score	Criteria
0	The student does not select any option as the most convincing.
1	The student selects the External Conviction proof scheme as the most convincing option (Gaye/ Pınar)
2	The student selects the Empirical proof scheme as the most convincing option (Harun/Sinan)
3	The student selects the Analytical proof scheme as the most convincing option (Fatma/ Ömer)

Rubric for Argumentation Tasks		
Score	Criteria	Example(s)
0	There exists no argument.	-
1	There exists a level 1 argument; there is a claim only (C).	“x^2 is not always greater than x. It could be x^2 =70 and x=71”
2	There exists a Level 2 argument; there are claim, data and or warrant (C+D/W).	“I don’t find it correct. For example, let x be 6; x^2 =6x6=36. How does x^2 smaller than x?”
3	There exists a Level 3 argument; there are claim, data and or warrant, and qualifier and or backing (C+D/W+ Q/B).	“Definitely, Meltem is right. Because matter cannot be created or destroyed. This shows that the matter was not destroyed. It just dissolved. It was divided into small pieces”.
4	There exists a Level 4 argument; there are claim, data and or warrant, qualifier and backing (C+D/W+Q+B).	Of course, Meltem. Because we learned ionization in eight-grade and accordingly the sugar dissociated into ions and drops a level so that it is invisible. If we still do not believe, we can look at it through microscope. I picked Meltem for these reasons.

APPENDIX C

SAMPLE ARGUMENTS AND THEIR COMPONENTS

Sample Level 1 Arguments and Their Components

1. This statement is wrong. Because no number can be greater than its square; $x^2>x$.

Claim: Wrong

2. I would recommend the membrane method to the governor. Because at the end, if we do not produce messy and sludgy environment, it won't hurt animals. When we use the membrane with tiny holes, the salt remains inside, and we use the water in an unsalted and potable way.

Claim: The membrane method is better.

Sample Level 2 Arguments and Their Components

1. I don't find it correct. For example, let x be 6; x^2 =6x6=36. How does x^2 smaller than x?

Claim: Not correct.

Data: $x=6$ and $x^2=36$; x^2 is not smaller than x

2. Meltem because the sugar is still in the water. Sugar and water are intertwined so the sugar may be seen as invisible and sugar gave its taste to water.

Claim: Meltem is right.

Data/ Warrant: Sugar gave its taste to the water.

3. The evaporation should be done. They don't have a right to harm any living thing.

Claim: The evaporation method is better.

Data/ Warrant: Animals get hurt with the membrane method (the method is implicit here.)

Sample Level 3 Arguments and Their Components

1. For same cases like when we talk about for natural numbers it is wrong, for example 2x2<2 4<2 is a wrong statement. In some cases, for simple fractions, it is correct:1/9x1/9<1/9 1/81<1/9 is a correct statement.

Claim and Qualifier: Sometimes wrong sometimes correct (Qualifier is implicit here.)

Data: 2x2<2 4<2 (for the wrong case), 1/9x1/9<1/9 1/81<1/9 (for the correct case)

2. I think Meltem is right because when the sugar is added into hot water, it slowly melts and it seems united with water but after a while the sugar can reappear when they are evaporated. As the teacher says, matter cannot be created or destroyed.

Claim: Meltem is right.

Data/ Warrant: Matter cannot be created or destroyed, When the sugar is added into hot water, it slowly melts

Backing: Sugar can reappear when they are evaporated.

3. I would pick the second method because the membrane method harms ocean animals. If animals get harmed, the balance of the nature becomes deteriorated. If the balance of the nature deteriorated, the balance of the world become deteriorated.

Claim: The evaporation method is better (in the student argument second method refers to the evaporation method)

Data: The membrane method harms ocean animals.

Backing: If animals get harmed, the balance of the nature becomes deteriorated. If the balance of the nature deteriorated, the balance of the world become deteriorated.

Sample Level 4 Arguments and Their Components

1. Of course, Meltem. Because we learned ionization in eight-grade and accordingly the sugar dissociated into ions and drops a level so that it is invisible. If we still do not believe, we can look at it through microscope. I picked Meltem for these reasons.

Claim: Meltem

Data/ Warrant: Sugar dissociated into ions and drops a level so that it is invisible.

Qualifier: Of course

Backing: We can look it through microscope.

2. Of course, the second one. Because the extinction of generation of animals can result in the deaths of humans by deteriorating the balance of the world.

But what is money, it can come back whereas we cannot bring back what the God (Allah) created.

Claim: The evaporation method is better (the second one refers to the evaporation method in the student response)

Data: The animals get hurt with the membrane method (this is gotten implicitly from the student argument).

Qualifier: Of course

Backing: Because the extinction of generation of animals can result in the deaths of humans by deteriorating the balance of the world, we cannot bring back what the God (Allah) created.

APPENDIX D

THE INSTRUMENT (TURKISH)

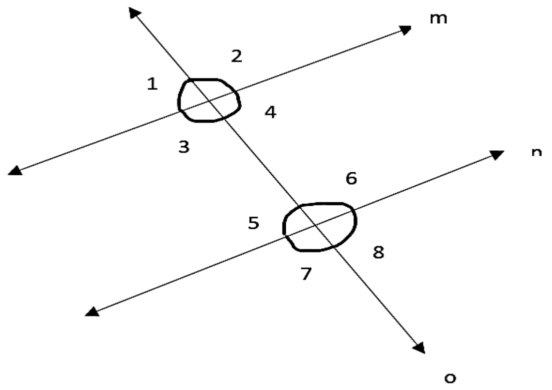
Cinsiyet: K / E	Geçen seneki matematik notu:
Okul kodu:	

İSPAT TESTİ

Aşağıdaki ifadeleri ispatlayınız.

1. İki tek sayının toplamı her zaman bir çift sayıya eşittir.

2.



m doğrusu n doğrusuna paraleldir.

Buna göre 2 numarayla gösterilen açının ölçüsünün 7 numarayla gösterilen açının ölçüsüne eşit olduğunu ispatlayınız.

Aşağıda bazı ifadeler ve bu ifadelerin doğruluğunu veya yanlışlığını ispatlayan bazı öğrenci açıklamaları bulunmaktadır. Bu açıklamaları verilen sorulara göre değerlendiriniz.

1. İki tek sayının toplamı her zaman bir çift sayıya eşittir.

FATMA	GAYE
Herhangi k ve m sayıları için, tek sayılarımızdan biri 2m+1 diğeri 2k+1 olsun. Bu sayıları topladığımızda; 2m+1+2k+1 işlemiyle 2m+2k+2 elde ederiz. Elde ettiğimiz sonucu 2 parantezine alırsak; 2(m+k+1) sonucuna ulaşırız. Buna göre parantez içindeki toplamdan bağımsız olarak elde ettiğimiz sonuç 2'nin katı olarak çıkar. 2'nin katı olan sayılar çifttir. Böylelikle bu ifadeyi ispatlamış oluruz.	Matematik öğretmenimizden bir kural öğrendik. İki çift sayının toplamı çift sayı, iki tek sayının toplamı çift sayı, bir çift bir tek sayının toplamı tek sayı olur. Bu kurala göre iki tek sayının toplamı her zaman çift olur. Böylelikle bu ifadeyi ispatlamış oluruz.

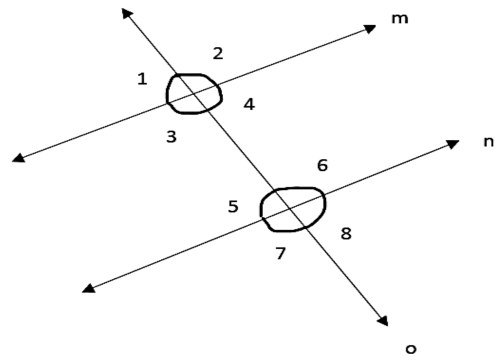
HARUN
İki tane tek sayı seçelim. Mesela 7 ve 59 seçtiğimiz sayılar olsun. Bu sayıların toplamı 66 çıkar. Elde ettiğimiz bu sayı çift bir sayıdır. Başka iki tane tek sayı seçelim. Bunlar 11 ve 83 olsun. Bunların toplamı 94'e yani çift bir sayıya eşittir. Bu durum seçebileceğimiz bütün tek sayılarda böyle olur. Böylelikle bu ifadeyi ispatlamış oluruz.

1. Hangi öğrencinin cevabı sizi bu ifadenin doğru olduğuna ikna etti?

2. Seçtiğiniz cevap neden diğerlerinden daha ikna edici?

3. Seçmediğiniz yanıtı veren öğrencilere daha ikna edici olmaları için ne önerirdiniz?

2.



m doğrusu n doğrusuna paraleldir.

Buna göre 2 numarayla gösterilen açının ölçüsünün 7 numarayla gösterilen açının ölçüsüne eşit olduğunu ispatlayınız.

PINAR	SİNAN
Şekle baktığımızda 2 numaralı açıyla 7 numaralı açının ölçülerinin birbirine eşit olduğu görülebilir. Böylelikle bu ifadeyi ispatlamış oluruz.	2 numaralı açıyı açıölçer ile ölçtüm ve 105° buldum. Aynı şekilde 7 numaralı açıyı da ölçtüm ve 105° olduğunu gözlemledim. Bundan yola çıkarak bu iki açının ölçülerinin birbirine eşit olduğunu ispatlamış oldum.

ÖMER
2 numaralı açının ölçüsüyle 6 numaralı açının ölçüsü birbirine eşittir çünkü bu iki açı yöndeştir. Yöndeş açıların ölçüsü birbirine denktir. 6 numaralı açı ile 7 numaralı açı ters açılardır ve ölçüleri eşittir. Çünkü ters açıların ölçüleri birbirine eşittir. 6 numaralı açı hem 2 numaralı açıya hem de 7 numaralı açıya eşit olduğundan ve eşitliğin geçişsel özelliğine göre 2 numaralı açıyla 7 numaralı açının ölçüsünün birbirine eşit olduğunu ispatlamış oluruz.

1. Hangi öğrencinin cevabı sizi bu ifadenin doğru olduğuna ikna etti?
2. Seçtiğiniz cevap neden diğerlerinden daha ikna edici?
3. Seçmediğiniz yanıtı veren öğrencilere daha ikna edici olmaları için ne önerirdiniz?

ARGÜMANTASYON SORULARI:

1. Diyelim ki x herhangi bir sayı olsun.

$x^2 < x$ ifadesinin doğruluğu için ne söylersiniz? Nedeniyle birlikte açıklayınız.

2. ŞEKERLE SUYU KARIŞTIRMAK

İki öğrenci bir miktar şekeri bir bardak sıcak suyun içine döküp gözlem yapıyorlar ve üç bilgi ediniyorlar.

1. Şekeri suya ilave ettikten sonra ikisi karıştırılıyor ve şeker artık görünmüyor.
2. Karıştırdıktan sonra her öğrenci suyun tadına bakıyor ve ikisi de suyun tatlı olduğunu belirtiyor.
3. Karıştırmadan önceki suyun, bardağın ve şekerin ağırlıklarının toplamı, karıştırdıktan sonraki bardaktaki su ve şeker karışımın ağırlıklarına eşit çıkıyor.

O zaman bu iki öğrenci neden şekeri göremiyor? *Leyla'ya göre* şeker artık yok oldu bu yüzden onu göremiyoruz. *Meltem'e göre* ise şeker hala suyun içerisinde bulunuyor.

Bu soruya çözüm bulmak için öğretmenlerine danışan Leyla ve Meltem, öğretmenlerinden aşağıdaki bilgileri ediniyorlar:

- Bir madde yoktan var edilemez, vardan yok edilemez.
- Bazen bir madde başka bir maddeyle karıştırıldığında oldukça küçük parçacıklara bölünebilir.

Bütün bu bilgilere dayanarak sizce kim haklı? Leyla mı? Meltem mi? Lütfen düşüncenizi sebepleriyle birlikte belirtiniz.

3. TUZDAN ARINDIRMA

Amerika'nın California eyaletinde su kıtlığı yaşanmaktaydı. Eyaletin valisi bu soruna çözüm aramaktaydı. Yardımcıları dünyanın %97' sinin okyanus sularından oluştuğunu ama bu suların tuzlu ve içilemez durumda olduğunu bildirdiler. Suları içilebilir hale getirmek için "Tuzdan Arındırma" işlemi uygulanmalıydı. Vali yardımcıları bu işlemi yapabilmek için iki yöntemin bilgilerini valiyle paylaştılar. Buna göre:

Birinci yöntem *"Buharlaştırma Yöntemi"*. Bu yöntemle okyanus suyu ısıtılır ve buhara dönüştürülür. Geride tuz kalır. Daha sonra buhar soğuk bir objenin üzerinden geçirilir ve temiz içme suyu elde edilir.

İkinci yöntem *"İnce Zar Yöntemi"*. Bu yöntemle üzerinde oldukça küçük delikler olan ince bir zardan okyanus suyu geçirtilir. Bu zardan yalnızca su molekülleri geçebilir. Tuz molekülleri geçemez. Böylelikle temiz ime suyu elde edilir.

Vali ayrıca bu iki yöntem hakkında iki bilgi daha ediniyor:

1. İnce zar yönteminin sonucunda kirli ve çamurlu bir ortam oluşuyor ve bu ortam okyanusta yaşayan canlılara zarar veriyor.
2. Buharlaştırma yöntemi İnce zar yönteminden çok daha fazla enerji kullanımı gerektiriyor. Bu enerji tüketimi California için pahalı gelebilir.

Bu iki tuzdan arındırma yönteminden hangisini valiye önerirsiniz? İki yöntemi karşılaştırın ve hangi yöntemin daha iyi olduğuna karar verin.

APPENDIX E

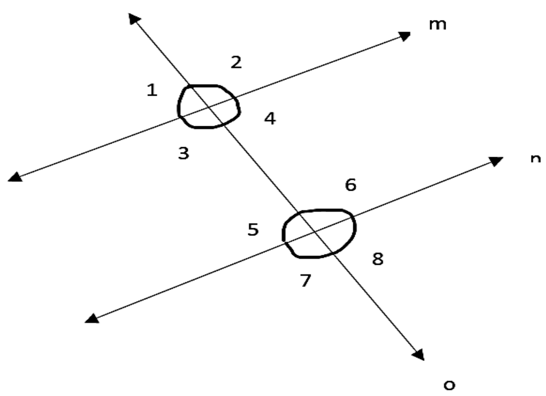
THE INSTRUMENT (ENGLISH)

Gender: F / M	Math. grade from previous term:
School Code:	

PROOF TEST

Prove the statements below, please.

1. The sum of two odd numbers is always equal to an even number.
2.



Line m is parallel to line n.

Accordingly, prove that the angles with number 2 and number 7 are congruent.

Below, there are some statements and students' proofs about the correctness of these statements. Evaluate students' proofs according to the given questions.

1. The sum of two odd numbers is always equal to an even number.

FATMA	GAYE
For any number k and m, let our odd numbers be 2k+1 and 2m+1. When we add these numbers, we get 2m+2k+2. If we take the two out of the parenthesis, we find 2(m+k+1). Accordingly, we get a number which is a multiple of two, independent from the sum of the numbers in the parenthesis. Multiples of two are even numbers. The statement is proved.	We learned a rule from our mathematics teacher. The sum of two even numbers is an even number, the sum of one odd and one even number is an odd number and the sum of two odd numbers is an even number. According to this rule, the sum of two odd numbers becomes an even number. The statement is proved.

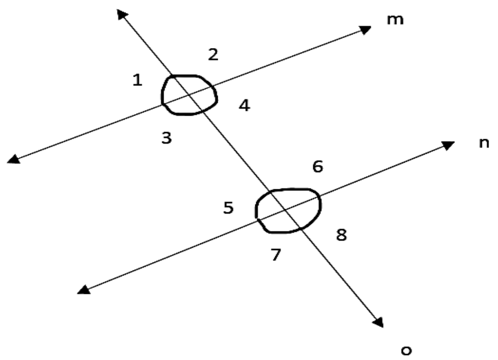
HARUN
Let's pick two odd numbers. For instance, 7 and 59. When we add them we get 66, which is an even number. Let's pick other two odd numbers, 11 and 83. Their sum is 94, an even number. This situation is similar for any two odd numbers we select. The statement is proved.

1. Which student response convinced you most about the correctness of the statement?

2. Why that response is more convincing than others?

3. What would you suggest others to become more convincing?

2.



Line m is parallel to line n.

Accordingly, prove that the angles with number 2 and number 7 are congruent.

PINAR	SİNAN
When wo look at the picture, it can be viewed clearly that these the angles with number 2 and 7 are congruent. The statement is proved.	I measured the angle with number 2 and I found that it is 105° by protractor. I found that the angle with number 7 is also 105°, with the same way. Accordingly, I proved that these angles are congruent.

ÖMER
The angles with number 2 and 6 are congruent since they are corresponding angles. Corresponding angles are congruent. The angles with number 6 and 7 are vertical angles and their degrees are equal because vertcal angles are congruent. Since the angle with number 6 is congruent to both the angles 2 and 7, the angles with numbers 2 and 7 are congruent. The statement is proved.

1. Which student response convinced you most about the correctness of the statement?

2. Why that response is more convincing than others?

3. What would you suggest others to become more convincing?

ARGUMENTATION TASKS

1. Let x be any number.

What can you say about the correctness of this statement: $x^2 < x$. Explain with your reasoning.

2. Two students pour sugar grains into a glass of hot water. They make three observations.

    1. Once the sugar is poured into the water, it is stirred. After stirring, the sugar can no longer be seen

    2. Also, after stirring, each student tastes the water. They both agree that the water tastes sweet.
    3. The weight of the water + glass + sugar is the same as the weight of the glass containing the mixture after the sugar was stirred in.

    Why can we no longer see the sugar?
    Leyla thinks that the sugar is disappeared. That is why, we cannot see it anymore. Meltem thinks that sugar is still inside the water.
    To find an answer, they conducted their teacher and got the following information:

    - Matter cannot be created or destroyed.
    - Sometimes a substance breaks into very small pieces when mixed with another substance.

    According to this information, who is right? Leyla or Meltem? Please indicate your answer with your reasoning.

3. The governor has declared that removing salt from ocean water can help California's water shortage. The governor's advisors point out that 97% of the Earth's water is in the oceans. However, ocean water is salty and cannot be drunk. If salt can be removed from ocean water with a process called desalination, a lot of previously undrinkable ocean water can be used to help the water shortage. The governor has been informed that one method of desalination is to heat ocean water so that it turns into water vapor. This leaves the salt behind. The water vapor is then passed over a cold object, causing it turn back into clean liquid water. This is called the evaporation method.

    The governor has also been made aware of another method of desalination that involves pushing ocean water through a membrane, which is a sheet with tiny holes in it. Only the water particles pass through the membrane, leaving all the salt behind. This is called the membrane method.

    The governor has learned that the membrane method produces a messy sludge called brine that can be harmful to ocean animals when dumped back into the ocean. However, the governor is also told that heating water with the evaporation method requires more energy than the membrane method. Increased energy consumption can be expensive for California.

Which desalination method would you recommend to the governor? Compare the two methods and determine which one is better.

APPENDIX F

CORRESPONDING OBJECTIVES OF PROOF AND ARGUMENTATION TASKS IN THE TURKISH MATHEMATICS AND SCIENCE CURRICULA

Objectives of Proof Tasks:

M.3.1.1.8. Tek ve çift doğal sayıları kavrar. Tek ve çift doğal sayılarla çalışılırken gerçek nesneler kullanılır.
M.3.1.1.9. Tek ve çift doğal sayıların toplamlarını model üzerinde inceleyerek toplamların tek mi çift mi olduğunu ifade eder.

M.7.3.1.2. İki paralel doğruyla bir keseninin oluşturduğu yöndeş, ters, iç ters, dış ters açıları belirleyerek özelliklerini inceler; oluşan açıların eş veya bütünler olanlarını belirler; ilgili problemleri çözer. a) Aynı düzlemde olan üç doğrunun birbirine göre durumları ele alınır. b) İki doğrunun birbirine paralel olup olmadığına karar vermeye yönelik çalışmalara da yer verilir. Bunu yaparken doğruların ortak kesenle yaptığı açıların eş olma durumlarından yararlanılabilir.

Objectives of Argumentation Tasks:

M.6.1.1.1. Bir doğal sayının kendisiyle tekrarlı çarpımını üslü ifade olarak yazar ve değerini hesaplar.
M.7.1.1.4. Tam sayıların kendileri ile tekrarlı çarpımını üslü nicelik olarak ifade eder.
M.7.1.3.4. Rasyonel sayıların kare ve küplerini hesaplar.

F.4.4.5.1. Günlük yaşamında sıklıkla kullandığı maddeleri saf madde ve karışım şeklinde sınıflandırarak aralarındaki farkları açıklar
F.4.4.5.2. Günlük yaşamda karşılaştığı karışımların ayrılmasında kullanılabilecek yöntemlerden uygun olanı seçer.
Eleme, süzme ve mıknatısla ayırma yöntemleri üzerinde durulur.

F.4.4.5.3. Karışımların ayrılmasını, ülke ekonomisine katkısı ve kaynakların etkili kullanımı bakımından tartışır.

APPENDIX G

CORRELATION MATRIX FOR ALL VARIABLES

Table 13. Spearman's Rho Correlations Coefficients

	M. ach	PC1	PC2	PC	PE1	PE2	PE	ARG1	ARG2	ARG3	ARG
M. ach	1.00										
PC1	.26**	1.00									
PC2	.58**	.11	1.00								
PC	.61**	.49**	.90**	1.00							
PE1	.24**	.02	.12	.12	1.00						
PE2	.40**	.03	.27**	.26**	.25**	1.00					
PE	.42**	.35	.26**	.25**	.75**	.81**	1.00				
ARG1	.51**	.17**	.35**	.37**	.16*	.30**	.30**	1.00			
ARG2	.32**	.17**	.21**	.25**	.07	.18**	.17**	.34**	1.00		
ARG3	.29**	.11	.20**	.23**	.10	.22**	.22**	.34**	.34**	1.00	
ARG	.53**	.19**	.35**	.38**	.15*	.32**	.31**	.83**	.60**	.74**	1.00

Note.

** $p < .01$

* $p < 0.5$

APPENDIX H

ORIGINAL STUDENT RESPONSES (TURKISH)

No Proof Responses in Proof Construction Tasks

1) Bence iki tek sayının toplamı bir çift sayıya eşit olabilir, olmaya dabilir. Örnek: 3x2=6 ve 3x5=15.

2) Tekler bir araya gelirse çift olur [Çöp adam çizimleri var; bir kız ve bir erkekten oluşan…]

3) Öncelikle bir elmayı ikiye böldüğümüzü düşünün. Sonra bu iki yarım elmanın bir yarısını sağ elimize diğer yarısını sol elimize alalım Bu yarım elmaların her birini tek sayı olarak düşünelim. Artık iki elimizde de bir adet tek sayı var. Bu iki tek sayıyı birleştirdiğimizde bir bütün oluşturuyoruz. Bütünü de çift sayı olarak düşünürsek ispatlamış olursunuz.

4) Eğer 1+2+3+4=20 ve 5+6+7+8=20 ise her sayı on üzerinden hesaplanır. Hepsi eşittir.

5) Dış açılar eşit olduğu için …

6) m doğrusu ile n doğrusu paralel olduğu için geniş açıyla geniş açı, dar açıyla dar açı eşittir.

7) ... ispatlamaya gerek yok aslında bilim her zaman haklıdır.

External Conviction Proof Responses in Proof Construction Tasks

8) Matematik dersinde formüller vardır. Bu formüller sayesinde sonuca ulaşılmıştır. Bu soru da matematiğin temel noktalarından bir tanesidir(bence)

9) …(örnek)… Bu ya matematiğin bir kuralı ya da tamamen tesadüf olabilir.

10) ...Birbirine eşittir, ters açıdır.. Bunu biz öğrendik ve biliniyor.

11) İç ve dış ters açılar eşittir… Belli kurallar değiştirilemez

12) Bu matematikte bir kuraldır. Örneğin… Bunun gibi başka kurallar da vardır.

13) Bu aynen -x-= + gibi bir şeydir.

14) Z kuralına göre dıştaki çapraz açılar birbirine eşittir.

15) Paralel doğrularda karşılıklı yerlerin ölçüsü matematiksel olarak dereceleri eşit olduğundan dolayı

Empirical Proof Responses in Proof Construction Tasks

16) …Bu konuyla ilgili birden fazla örnek verebiliriz. Bu örnekler de bu bilginin doğruluğunu kanıtlar. Mesela… Bu örnekler milyonlarca kadar olabilir. Çünkü çok tek sayımız vardır…Örnek vererek anlatırsak anlatılan kişinin aklında daha iyi kalmasını sağlayabiliriz.

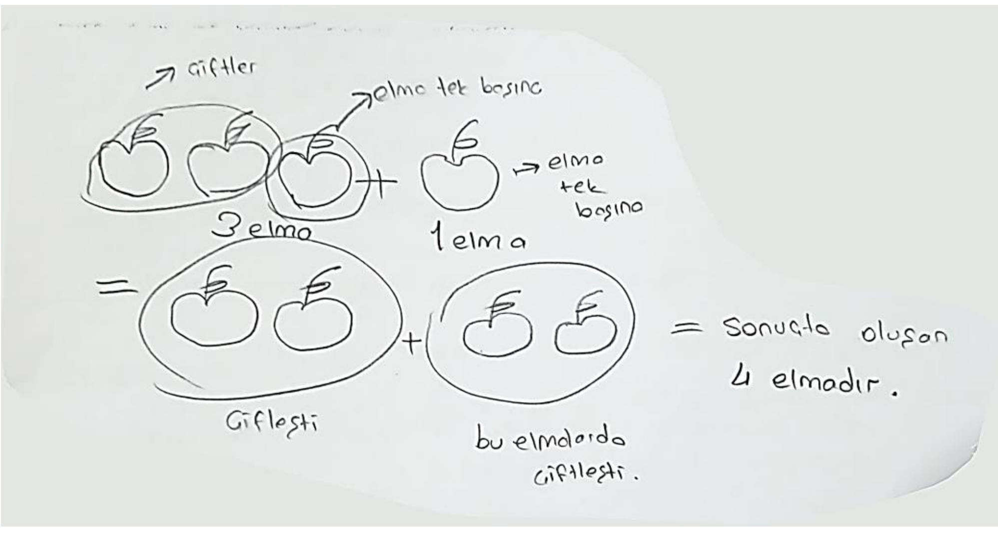
17) Bunun ispatını sadece örneklerle yapabilirim.

18) …İllaki bir formülü vardır ancak bilmiyorum.

19) Bir konuyu açıklamada iyi değilim, bu yüzden örnek vereceğim.

Analytical Proof Responses in Proof Construction Tasks

20) Doğru çünkü 3+1=4 bir tek sayı ile yine bir tek sayının toplamı her zaman çifttir…Bunun ispatı 3 elmam var 1 tane elma daha verirsem her ikisi de biririni eşler ve 4 olur. Yani şöyle:



21) Çift sayıların tamamı 2'ye tam bölünebilir. Yani, ikili gruplar halinde yazılabilir. Örneğin; 4 sayısı, 2+2 şeklinde gruplara ayrılabilir. Tek sayılar 2'ye bölünemez. Ve ikili gruplar halinde yazılamaz. Örneğin; 5 sayısı, 2+2+1 şeklinde yazılır. İkili gruplara ayırmak istediğimizde 1 kalır…Şimdi, iki tek sayıyı toplayalım: 7 (2+2+2+1) + 9 (2+2+2+2+1) = 2+2+2+2+2+2+2+ (1+1)= 16. Çift sayıdır.

22) Paralel doğrularda karşılıklı yerlerin ölçüsü eşittir. Mesela 6 ve 7, 2 ve 3. Ayrıca, aynı doğrultudaki açıların ölçüsü de eşittir: 2 ve 6, 3 ve 7. Bu bilgilere göre 2 ve 7 eşit olur.

Level 1 Arguments

23) Doğru olmaya da bilir. x'e verdiğimiz sayıya göre değişir.

24) Meltem. Şeker suyun içerisinde ama görünemiyor.

25) Suyu ısıtarak buhar yapıp sonra o buharı soğuk yerden geçirerek temiz suyu oluşturması

26) Her zaman x^2 x'ten büyük olmaz x^2=70, x=71 olabilir.

27) Bu denkleme eksili değer verirsek doğru çıkıyor. Mesela -5 sayısını verirsek - 5^2<-5= -25<-5 yani gördüğünüz gibi doğru çıkıyor.

28) Meltem haklı çünkü … suyu görememememizin nedeni suyun şekere oranla daha fazla olması.

29) Birinci yöntem çünkü ince zarda süzgeçten geçse bile tuzu gitmez çünkü tuz çözünmüştür ve tuzu hala kalır.

Level 2 Arguments

30) x iki olsun 2^2<2 ~~4<2~~ x üç olsun 3^2<3 ~~9<3~~ x dört olsun 4^2<4 ~~16<4~~ Gördüğünüz gibi hepsi yanlış çıkıyor.

31) Buharlaştırma yapılmalı. Hiçbir canlıya zarar vermeye hakları yok.

32) Meltem. Leyla'nın dediğiyle hocasının dediği zaten ters düşüyor.

33) İşaretlediğim maddeyi yazan kişi haklıdır. Meltem haklı.

34) Ben doğru bulmuyorum mesela x'e 6 diyelim x^2=6x6=36 oluyorsa x'den nasıl küçük oluyor?

35) Meltem çünkü şeker hala suyun içindedir. Sadece suyla iç içedir bu yüzden görünmez gibi görünebilir ve tadını suya vermiştir.

36) Bana göre Meltem haklı çünkü öğretmenin de açıkladığı gibidir. O madde yok olmaz. Sadece su ile karıştırdığımızda su şekeri emer ve biz şekeri göremeyiz.

Level 3 Arguments

37) Bu ifade doğal sayılar için biraz geçersiz. Kesirli sayılar için olanaklı olması ile beraber doğruluğu tartışılır. Ama bazen doğrudur ve şuna geliyor: x'e -1 versek= 1<-1 olur ama bu ifade yanlıştır ama x'e ½ versek= 1/4<1/2 olur ve bu doğru olur. Yani her zaman geçerli değil.

38) Bu ifade bazı sayılar için doğru bazı sayılar için de yanlış bir kuraldır. Mesela x=2 olursa 4<2 olduğu için bu kralı karşılayamayacaktır. Ama bir rasyonel sayı olursa bu kuralı karşılayabilir. Mesela x=1/3 ise $(1/3)^2$=1/9'a eşittir. 1/9<1/3 olduğu için bu kuralı karşılayabilir.

39) Bence Meltem haklı çünkü şeker sıcak suyun içine atıldığında yavaşça eriyerek suyu ile bir görünür ancak bir süre sonra o suyu buharlaştırdığımızda şekerin tekrar ortaya çıktığı görülecektir. Hocanın da dediği gibi bir madde yoktan var vardan da yok edilemez.

40) 2.çünkü buharlaştırma yönteminde canlılara zarar verilmeden yapılıyor. Enerji konusunda ise yenilenebilir enerji yakıtları kullanılabilir. 1. Neden değil burada canlılara zarar veriliyor ve bu canlı türlerinin yok olmasına sebep olabilir. 2.yöntem daha iyi.

41) 2.yöntemi seçerdim çünkü ince zar yöntemi deniz canlılarına zarar vermektedir. Eğer bir canlıya zarar verirse doğanın dengesini bozmuş oluruz. Doğanın dengesini bozarsak da dünyadaki denge bozulmuş olur.

42) x’e ½ verelim. Bu durumda payda ne kadar büyük olursa sayı o kadar küçük olur. Bunun sebebi bir şeyi ne kadar küçük parçalara bölersek o kadar az olur

REFERENCES

Abdullah, A. H., & Zakaria, E. (2013). Enhancing students' level of geometric thinking through van Hiele's phase-based learning. *Indian Journal of Science and Technology*, *6*(5), 4432-4446.

Acar, O., Türkmen, L., & Roychoudhury, A. (2010). Student difficulties in socio-scientific argumentation and decision-making research findings: Crossing the borders of two research lines. *International Journal of Science Education, 32*, 1191-1206.

Aktaş, Y. (2002). *Okul öncesi dönemde matematik eğitimi*. Adana: Nobel Kitabevi

Altıparmak, K., & Öziş, T. (2005). Matematiksel ispat ve matematiksel muhakemenin gelişimi üzerine bir inceleme. *Ege Eğitim Dergisi*, *6*(1), 25-37

Aylar, E. (2014). A Study on the forms of perception of 7th grade students towards the concept of proof. *Journal of Education and Future*, (5), 39-56

Balacheff, N. (1988). Aspects of proof in pupils' practice of school mathematics. In D. Pimm (Ed.), *Mathematics, teachers and children* (pp. 216–235). London, UK: Hodder & Stoughton.

Banegas, J. A. (2013). Argumentation in mathematics. In A. Aberdein & I.J. Dove (Eds.), *The argument of mathematics* (pp. 47-60). Dordrecht, The Netherlands: Springer.

Bieda, K. N., & Lepak, J. (2014). Are you convinced? Middle-grade students' evaluations of mathematical arguments. *School Science and Mathematics*, *114*(4), 166-177.

Blanton, M., Stylianou, D. A., & David, M. (2003). The nature of scaffolding in undergraduate students' transition to mathematical proof. In N.A. Pateman, B.J. Dougherty & J. Zilliox (Eds.), *Proceedings of the 27th PME International Conference* (pp. 113–120). Honolulu, Hawaii: ERIC.

Boero, P. (1999). Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education. *International Newsletter on the Teaching and Learning of Mathematical Proof, 7*(8), 7-8

Breyfogle, M., & Lynch, C. M. (2010). Van Hiele revisited. *Mathematics Teaching in the Middle School*, *16*(4), 232-238.

Cabassut, R. (2005). Argumentation and proof in examples taken from French and German textbooks. In M. Bosh (Ed.) *Proceedings of the Fourth Congress of the European Society for Research in Mathematics Education* (pp. 391-400). Sant Feliu de Guíxols, Spain: FUNDEMI IQS, Universitat Ramon Llull

Carrascal, B. (2015). Proofs, mathematical practice and argumentation. *Argumentation*, *29* (3), 305-324.

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics*, *24*(4), 359-387.

Christenson, N. (2015). *Socioscientific argumentation: Aspects of content and structure* (Unpublished PhD thesis). Karlstad University, Sweden

Conner, A., Singletary, L. M., Smith, R. C., Wagner, P. A., & Francisco, R. T. (2014). Teacher support for collective argumentation: A framework for examining how teachers support students’ engagement in mathematical activities. *Educational Studies in Mathematics*, *86*(3), 401-429.

Cooper, J., Walkington, C., Williams, C., Akinsiku, O., Kalish, C., Ellis, A., & Knuth, E. (2011). Adolescent reasoning in mathematics: Exploring middle school students' strategic approaches in empirical justifications. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2188-2193). Austin, TX: Cognitive Science Society.

Creswell, J. W. (2014). *Research design qualitative, quantitative, and mixed methods approaches (*4th ed*.).* Thousand Oaks, CA: SAGE Publications.

Csíkos, C. A. (1999). Measuring students' proving ability by means of Harel and Sowder's proof categorization, In O. Zaslavsky (Ed.), *Proceedings of the 23rd Conference of the International Group for* the *Psycgology of Mathematics Education* (pp. 233-240). Haifa, Israel: Israel Institute of Technology.

Cyr, S. (2011). Development of beginning skills in proving and proof writing by elementary school students, In M. Pytlak, T. Rowland, & E. Swoboda (Eds.), *Proceedings of the Seventh Congress of the European Society for Research in Mathematics Education* (pp. 263-272). Rzeszów, Poland: University of Rzeszów

Dawson, V., & Venville, G. J. (2009). High-school students’ informal reasoning and argumentation about biotechnology: An indicator of scientific literacy? *International Journal of Science Education*, *31*(11), 1421-1445.

Demiray, E., & Işıksal - Bostan, M. (2017). An investigation of pre-service middle school mathematics teachers’ ability to conduct valid proofs, methods used, and reasons for invalid arguments. *International Journal of Science and Mathematics Education*, *15*(1), 109-130.

de Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras 24*, 17-24.

Dufour, M. (2013). Arguing around mathematical proofs. In A. Aberdein & I.J. Dove (Eds.), *The argument of mathematics* (pp. 61-76). Dordrecht, The Netherlands: Springer.

Duval, R. (2007). Cognitive functioning and the understanding of mathematical processes of proof. In P. Boero (Ed.), *Theorems in school: From historic, epistemology, and cognition to classroom practice* (pp. 163-181). Rotterdam: Sense Publishers.

Ellis, A. E., Lockwood, E., Williams, C. C. W., Dogan, M. F., & Knuth, E. (2013). Middle school students' example use in conjecture exploration and justification. In L. R. Van Zoest, J.J. Lo, & J. L. Kratky (Eds.), *Proceedings of the 34th Annual Meeting of the North American Chapter of the Psychology of Mathematics Education* (pp. 135-142). Kalamazoo, MI: Western Michigan University.

Erduran, S. (2008). 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. 47–70). Dordrecht, The Netherlands: Springer.

Erduran, S., & Jiménez-Aleixandre, M. P. (2008). *Argumentation in science education: Perspectives from classroom-based research.* Dordrecht, The Netherlands: Springer.

Erduran, S., Osborne, J., & Simon, S. (2004). TAPping into argumentation: Developments in the application of Toulmin. *Science Education*, *88*(6), 915–933.

Evagorou, M., & Osborne, J. (2013). Exploring young students' collaborative argumentation within a socioscientific issue. *Journal of Research in Science Teaching*, *50*(2), 209-237.

Gibson, D. (1998). Students' use of diagrams to develop proofs in an introductory analysis course. In E. Dubinsky, A. Schoenfeld, & J. Kaput (Eds.), *Research in collegiate mathematics education. III.* (pp.284–307). Providence, RI: American Mathematical Society.

Güneş, S. (2013). *Matematik eğitiminde argümantasyon ve kanıt süreçlerinin analizi ve karşılaştırılması* (Unpublished Master's thesis). Hacettepe University, Ankara, Turkey.

Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational studies in mathematics, 44*(1-2), 5-23.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. *American Mathematical* Society, *7*, 234-283.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education*, *31*(4), 396-428.

Heinze, A., & Reiss, K. (2003). Reasoning and proof: Methodological knowledge as a component of proof competence. In M. A. Mariotti (Ed.), *Proceedings of the Third Conference of the European Society for Research in Mathematics Education* (pp. 4-6). Bellaria, Italy: University of Pisa.

Huck, S. W. (2012). *Reading statistics and research* (6th ed.). Boston, MA: Pearson.

Hunter, R. (2007). Can you convince me: Learning to use mathematical argumentation. In J. Woo, H. Lew, K. Park, D. Seo (Eds.), *Proceedings of the 31st Annual Conference of the International Group for the Psychology of Mathematics Education,* (Vol. 3, pp. 81-88). Seoul, South Korea: The Korea Society of Educational Studies in Mathematics.

Inglis, M., Mejia-Ramos, J. P., & Simpson, A. (2007). Modelling mathematical argumentation: The importance of qualification. *Educational Studies in Mathematics*, *66*(1), 3-21.

Kaya, E. (2013). Argumentation practices in classroom: Pre-service teachers' conceptual understanding of chemical equilibrium. *International Journal of Science Education*, *35*(7), 1139-1158.

Knipping, C. (2003). Argumentation structures in classroom proving situations. In M. A. Mariotti (Ed.), *Proceedings of the Third Congress of the European Society for Research in Mathematics Education*. Bellaria, Italy: University of Pisa. Retrieved from: http://ermeweb.free.fr/CERME3/Groups/TG4/TG4_Knipping_cerme3.pdf

Knuth, E. J., Choppin, J., Slaughter, M., & Sutherland, J. (2002). Mapping the conceptual terrain of middle school students' competencies in justifying and proving. In D. S. Mewborn, P. Sztajn, D. Y. White, H. G. Weigel, R. L. Bryant, & K. Nooney (Eds.), *Proceedings of the 24th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, Vol. 4 (pp. 1693-1670). Athens, GA: Clearinghouse for Science, Mathematics, and Environmental Education.

Ko, Y-Y. (2010). Mathematics teachers' conceptions of proof: Implications for educational research. *International Journal of Science and Mathematics Education*, *8*, 1109-1129.

Krabbe, E. C. (2013). Arguments, proofs, and dialogues In A. Aberdein & I.J. Dove (Eds.), *The argument of mathematics* (pp. 31- 45). Dordrecht, The Netherlands: Springer.

Liu, Y., Tague, J., & Somayajulu, R. (2016). What do eighth grade students look for when determining if a mathematical argument is convincing. *International Electronic Journal of Mathematics Education*, *11*(7), 2373-2401.

Maloney, J. F., & Simon, S. (2006). Mapping children's discussions of evidence in science to assess collaboration and argumentation. *International Journal of Science Education*, *28*(15), 1817–1841.

Mejía-Ramos, J. P., & Inglis, M. (2009). What are the argumentative activities associated with proof? *Research in Mathematics Education*, *11*(1), 77-78.

Milli Egitim Bakanligi (2013). Ortaokul Matematik Dersi (5-8. Sınıflar) Öğretim Programı. Ankara: Talim ve Terbiye Kurulu Başkanlığı Yayınları. Retrieved from https://ttkb.meb.gov.tr

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics*, *27*(3), 249– 266.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 International Results in Mathematics.* Retrieved from: https://timssandpirls.bc.edu/timss2015/international-results/

Nardi, E., Biza, I., & Watson, S. (2014). What makes a claim an acceptable mathematical argument in the secondary classroom? A preliminary analysis of teachers' warrants in the context of an Algebra Task. In S. Pope (Ed.), *Proceedings of the 8th British Congress of Mathematics Education* (pp. 247-254). Nottingham, UK: University of Nottingham

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM

Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education, 87*, 224-240.

Osborne, J., Henderson, B., MacPherson, A., Szu, E., & Wild, A. (n.d.). *Assessments of argumentation in science-Beyond multiple choice, Stanford University,* Retrieved from http://scientificargumentation.stanford.edu/assessments/

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development* (T. Brown & K. J. Thampy, Trans.). Chicago, IL: University of Chicago Press.

Polat, M., Gönen, E., Parlak, B., Yıldırım, A., & Özgürlük, B. (2016). *TIMSS 2015 ulusal matematik ve fen bilimleri ön raporu: 4. ve 8. Sınıflar.* Ankara: T.C. Milli Eğitim Bakanlığı Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü

Raman, M. (2003). Key ideas: What are they and how can they help us understand how people view proof? *Educational Studies in Mathematics*, *52*(3), 319-325.

Ratcliffe, M., & Grace, M. (2003). *Science education for citizenship: Teaching socioscientific issues.* Buckingham,UK: Open University Press

Reid, D. A. (2005). The meaning of proof in mathematics education. In M. Bosh (Ed.) *Proceedings of the Fourth Congress of the European Society for Research in Mathematics Education* (pp. 458-468). Sant Feliu de Guíxols, Spain: FUNDEMI IQS, Universitat Ramon Llull

Reid, D. A., & Knipping, C. (2010). *Proof in mathematics education: Research, learning and teaching*. Rotterdam, The Netherlands: Sense Publishers.

Reid, D. A. (2011). Understanding proof and transforming teaching. In L. R. Wiest, & Lamberg, T. (Eds.), *Proceedings of the 33rd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 15-30). Reno, NV: University of Nevada, Reno.

Reiss, K. M., Heinze, A., Renkl, A., & Groß, C. (2008). Reasoning and proof in geometry: Effects of a learning environment based on heuristic worked-out examples. *ZDM- International Journal of Mathematics Education*, *40*(3), 455-467.

Ross, D., Fisher, D., & Frey, N. (2009). The art of argumentation. *Science and Children*, *47*(3), 28-31.

Rumsey, C., & Langrall, C. W. (2016). Promoting mathematical argumentation. *Teaching Children Mathematics*, *22*(7), 412-419.

Sadler, T. D., & Zeidler, D. L. (2005). Patterns of informal reasoning in the context of socioscientific decision making. *Journal of Research in Science Teaching, 42*, 112-138.

Schoenfeld, A. (1994). What do we know about mathematics curricula? *Journal of Mathematical Behavior*, *13*(1), 55-80.

Sen, C., & Guler, G. (2015). Examination of secondary school seventh graders' proof skills and proof schemes. *Universal Journal of Educational Research*, *3*(9), 617-631.

Staples, M., & Newton, J. (2016). Teachers' contextualization of argumentation in the mathematics classroom. *Theory into Practice*, *55*(4), 294-301.

Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education*, *40*(3), 314-352

Tall, D., Yevdokimov, O., Koichu, B., Whitley, W., Kondratieva, M., & Cheng, Y-H., (2011). Cognitive development of proof. In G. Hanna, & M. de Villiers (Eds.), *Proof and proving in mathematics education* (pp. 13-49). Dordrecht, The Netherlands: Springer.

Taş, U. E., Arıcı, Ö., Ozarkan, H. B., & Özgürlük, B. (2016). *PISA 2015 ulusal raporu*, Ankara: T.C. Millî Eğitim Bakanlığı Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü.

Tiberghien, A. (2008). Foreword. In S. Erduran & M. P. Jiménez- Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. ix-xv). Dordrecht, The Netherlands: Springer.

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

Toulmin, S.E. (2003). *The uses of argument*. Cambridge, UK: Cambridge University Press.

Ubuz, B., Dincer, S., & Bülbül, A. (2012). Argumentation in undergraduate math courses: A study on proof generation. In T.-Y. Tso (Ed.), *Proceedings of the 36th Conference of the International Group for the Psychology of Mathematics Education* (Vol 4., pp. 163- 170). Taipeh, Taiwan: PME van Hiele, P. H. (1959). *Levels of mental development in geometry*. Retrieved from http://www.math.uiuc.edu/~castelln/VanHiele.pdf

Venville, G. J., & Dawson, V. M. (2010). The impact of a classroom intervention on grade 10 students' argumentation skills, informal reasoning, and conceptual understanding of science. *Journal of Research in Science Teaching*, *47*(8), 952–977.

Waring, S. (2000). *Can you prove it? Developing concepts of proof in primary and secondary schools*. Leicester, UK: The Mathematical Association.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics, 48*(1), 101-119.

Wood, T. (1999). Creating a context for argument in mathematics class. *Journal for Research in Mathematics Education*, *30*(2), 171-191.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458-477.

Yopp, D. A. (2009). From inductive reasoning to proof. *Mathematics Teaching in the Middle School*, *15*(5), 286-291.

Zeidler, D. L. (2014). Socioscientific issues as a curriculum emphasis: Theory, research, and practice. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of Research on Science Education*. (Vol. 2, pp. 697- 726). New York, NY: Routledge.