Constructing-Evaluating-Refining Mathematical Conjectures and Proofs

Zulfiye Zeybek Simsek
Tokat Gaziosmanpasa University
zulfiye.zeybek@gop.edu.tr

This study focused on investigating the ability of 58 pre-service mathematics teachers (PSMTs) to construct-evaluate-refine mathematical conjectures and proofs. The PSMTs enrolled in a three-credit mathematics education course that offered various opportunities to engage with mathematical activities including constructing-evaluating-refining proofs in various topics. The PSMTs' proof constructions were coded in three categories as: Type P_1, Type P_2 and Type P_3 in decreasing levels of sophistication (from a mathematical stand point) and the constructions of conjectures were coded in two categories as: Type C_1 (correct conjectures); and Type C_2 (incorrect conjectures). In addition to classifying the PSMTs' proof and conjecture constructions, how they reacted when they needed to refine conjectures and proofs were also classified. Samples of classroom episodes were provided to exemplify these different proof-conjecture constructions-evaluations-refining processes. The results of the study demonstrated that the combined construction-evaluation-refining activities of conjectures and proofs were not only helpful to better illuminate the PSMTs' understanding of mathematical proofs, but they were also an essential instructional tool to help PSMTs comprehend mathematical ideas and relations.

There is a wide recognition that proof should be an important educational goal in the mathematical experiences of all students (National Council of Teachers of Mathematics [NCTM], 2000; National Governors Association Center/Council of Chief State School Officers [NGA & CCSSO], 2010). Proof is viewed as an essential element for developing deep understanding of mathematics in all students (Hanna & Jahnke, 1996; Mariotti, 2006), even in elementary school (Ball & Bass, 2000; Ball, Hoyles, Jahnke, & Movshovitz-Hadar, 2002; NCTM, 2000; Stylianides, 2007b; Yackel & Hanna, 2003). The processes of conjecturing, generalizing, justifying and proving — components of reasoning-and-proving activities as captured by Stylianides (2008) — not only require students to think broadly and flexibly about mathematical ideas and relationships but also these practices support students as they seek to understand the mathematics they are learning and using (Lannin, Ellis, Elliott, & Zbiek, 2011).

Research indicates that students of all levels tend to have limited understanding of proof and struggle with constructing proofs (Balacheff, 1988; Bell, 1976; Coe & Ruthven, 1994; Harel & Sowder, 1998; Healy & Hoyles, 2000; Knuth, Choppin, & Bieda, 2009). Many researchers demonstrated that the empirical reasoning is pervasive among school students at all levels (e.g., Coe & Ruthven, 1994; Goetting, 1995; Healy & Hoyles, 2000) as well as prospective and in-service teachers (e.g., Morris, 2002; Simon & Blume, 1996). However, most of these studies focus on students' (and pre-service and in-service teachers') effort to prove a given statement or evaluate whether researcher-generated arguments presented constitute proofs. Stylianides and Stylianides (2009) argued that asking individuals to prove a given statement or evaluating researcher-generated arguments could only illuminate bits of their understanding of proofs. In this vein, we lack images of teachers (both in-service and pre-service) engaged in conjecturing, generalizing and justifying activities, and thus know little about their understanding of these practices. This study aimed to contribute to literature on pre-service teachers' understanding of proof by reporting on pre-service mathematics teachers' (PSMTs) processes of constructing-evaluating-refining mathematical conjectures and proofs during a course in which PSMTs had many opportunities to engage in proving tasks. More specifically, the purpose of this study was to examine how pre-service teachers formulated, justified, refuted, and refined mathematical claims as they worked on proof-related tasks in a mathematics education course. The following research questions guided the study:

1) What is the nature of PSMTs' proof-conjecture construction-evaluation-refining processes when they engaged in proof related tasks in a mathematics education course?
2) What types of arguments or counter examples do PSMTs construct?
3) How do PSMTs evaluate mathematical conjectures or arguments?
4) How do PSMTs overcome contradiction when they encounter a counterexample?

In this paper, hyphenated terms of constructing-evaluating-refining tasks are used to refer to the integrated mathematical activities of generating or working on conjectures, evaluating conjectures and developing arguments in support of these conjectures or refutations, dealing with counterexamples that contradict the conjecture or missing parts of the proof, and finally refining conjectures or proofs. It is hypothesized that the hyphenated activities of constructing-evaluating-refining conjectures and proofs will better illuminate parts of PSMTs' understanding of proof. Analysing proof comprehension is a complex process (Stylianides & Stylianides, 2009) and cannot be captured by solely asking individuals constructing arguments, or evaluating researcher-generated arguments, which seems to be the case for most of the existing studies.

The Meaning and Functions of Proof in Mathematics Classrooms

Proof can serve as a broad range of functions in the discipline of mathematics such as: (a) verification (concerned with the truth of a statement), (b) explanation (providing insight into why it is true), (c) systematization (the organization of various results into a deductive system of axioms, major concepts and theorems), (d) discovery (invention of new results), (e) communication (the transmission of mathematical knowledge), and (f) intellectual challenge (Bell, 1976; de Villiers, 1990, 1999; Hanna & Barbeau, 2008; Larsen & Zandieh, 2008; Stylianides, 2008; Weber, 2002, 2010). While some researchers define proof from a pure mathematical standpoint by associating it with logical deductions that link premises with conclusions (de Villiers, 2012; Griffiths, 2000; Healy & Hoyles, 2000; Mariotti, 2000), some others associate proof with one or more of its functions mentioned above.

Traditionally the function of proof has been seen almost exclusively as being to verify or justify the correctness of mathematical statements (e.g., Ball & Bass, 2000). The "verification" function of proof is often interpreted in subjective terms, establishing the truth of a statement with an individual's belief in the truth of a statement and thus allocating proof a role in the subjective acquisition of such belief. Harel and Sowder (2007) use the term proof as "what establishes truth for a person or a community" (p. 806). Harel and Sowder (2007) conceptualize proof as the process that an individual or a community employs to remove doubts about the truth of a statement. This is consistent with Duval's (2002) view about proofs. Duval (2002) argues that a proof can change the logical value as well as epistemic value of a statement. That is, not only a proof may logically validate a statement, but it also affects the belief of the cognizing subject as to the truth of the statement. Bell (1976), however, argues that proof is not necessarily a prerequisite for conviction; proof is essentially a social activity of validation or establishing results, which follows reaching a conviction. Although the debates about whether proof is a prerequisite for conviction continue, researchers agree more on its social aspect. Balacheff (1988) claims that an argument becomes a proof after the social act of accepting. He defines proof as "an explanation which is accepted by a community at a given time," (Balacheff, 1988, p. 285). These two functions of proof—to convince individuals and to establish results in the field—are by no means the only functions of proof in mathematical activity.

Hersh (1993) argues that the role of proof in the classroom and the role of proof in mathematical discipline could be different, stating that the purpose of proof in mathematical discipline to be to convince, while in a classroom it should be to explain. Hanna (1990) distinguishes between "proofs that prove" and "proofs that explain". Knuth (2002) has echoed this theme by defining proof as "a deductive argument that shows why a statement is true by utilizing other mathematical results and/or insight into the mathematical structure involved in a statement" (p. 86). Even though there are diverse definitions to proof in mathematics classrooms, one crucial principle underlines all of them—the intellectual-honesty principle.

Stylianides (2007a) proposes a term—the intellectual-honesty principle, which "states that the notion of proof in school mathematics should be conceptualized so that it is, at once, honest to mathematics as a discipline and honouring of students as mathematical learners" (p. 3). Thus, proof in school mathematics should not only be considered from a mathematical point of view, which encapsulates that arguments accepted as proofs use true statements, valid forms of reasoning, and appropriate forms of expression, but also from a social point of view, which requires the statements and forms of reasoning used in an argument should be part of the classroom community's shared knowledge or accessible to the community. Stylianides (2007b) defines proof as a mathematical argument, a connected sequence of assertions for or against a mathematical claim, with the following characteristics:

1. It uses statements accepted by the classroom community (set of accepted statements) that are true and available without further justification;
2. It employs forms of reasoning (modes of argumentation) that are valid and known to, or within the conceptual reach of, the classroom community; and
3. It is communicated with forms of expression (modes of argument representation) that are appropriate and known to, or within the conceptual reach of, the classroom community (p. 291).

According to this conceptualization, proofs do not have to be restricted to formal proofs. Rather, both mathematical authenticity and students' (or classroom communities') development need to be taken into consideration, and informal proofs should be recognized as legitimate ones especially in early grade levels. Consistently, conceptualisation of proof in the course where the study took place was by no means restricted to formal proofs. Stylianides (2007b) argues that the definition can be used to describe students' conceptions of proof with respect to each of the three components of an argument identified in the definition: set of accepted statements, modes of argumentation, and modes of argument representation. As Stylianides proposed, the definition was used to judge whether PSMTs' arguments meet the standard of proof and, if not, to decide about which specific components of their arguments required development in this study. PSMTs' proof constructions were coded in three categories as follows:

- Type P_1: valid deductive arguments that justifies an assertion was true or false by showing how the statement of a theorem coheres and connects with the key properties of the concepts involved in the proof,
- Type P_2: general argument that fall short of being acceptable proofs, and

- Type P_3: unsuccessful attempt for a valid general argument that might include unfinished or irrelevant responses (or potentially relevant response but the relevance was not made evident) or use invalid modes of representations, argumentations, or false set of statements.

While Type P_1 arguments meet all three components of the definition mentioned above, Type P_3 arguments fail to meet one or more of the three components such as employing invalid modes of argumentation or using false set of statements. More about coding PSMTs' proof and conjecture constructions will be discussed next.

Methods

In this section, the context in which the research reported here took place, the research participants and the data collection and analysis processes is described.

Participants

Participants of the study were 58 pre-service mathematics teachers (PSMTs) studying to teach mathematics in grades 5 through 8. The participants were sophomores at a public university in Turkey when the study took place. The PSMTs enrolled in a mathematics education course during Spring semester (2016), which was taught by the author of this study. The class met three hours per week during the semester. The course was designed to cover a wide range of mathematical topics in three major mathematical domains (algebra, geometry, and number theory). The PSMTs were offered various opportunities to engage with mathematical proofs including constructing-evaluating-refining proofs, representing them in different ways (using everyday language, algebra, or pictures), and examining the correspondences among different representations.

Tasks

A sample of proof tasks in which the PSMTs were engaged in during the semester is presented here for readers to better conceptualize PSMTs' conjecture/proof construction-evaluation-refining processes (see Table 1).

Table 1
Sample Tasks

Task A: A fast-food restaurant sells chicken nuggets in packs of 4 and 7. What is the largest number of nuggets you cannot buy? How do you know this is the largest number you cannot buy? (adapted from Wilburne, 2014)	Task B: For every odd integer n, n^2 - 1 is divisible by 8 (adapted from Weber, 2003).
Task C: Justify that the area formula of a kite is $\frac{d_1 d_2}{2}$, where d_1 and d_2 are the diagonals of the kite.	Task D: At least one of the diagonals cuts the area of a quadrilateral in half (adapted from Ball, Hoyles, Jahnke, & Movshovitz-Hadar, 2002).

The tasks were in the conceptual reach of the PSMTs, yet they still provided a productive struggle since they required the PSMTs not only to solve the tasks but also justify their solutions. Additionally, the tasks could be solved by various solutions, which was taught to be essential to launch a forum for a rich discussion to engage the PSMTs in construction-evaluation-refining processes. All these consisted some of the reasons while choosing the tasks employed during the semester.

Data Collection

The participants were engaged in a course where they worked in groups of five or six. The PSMTs were engaged in class activities including integrated mathematical activities of generating (working on) conjectures, evaluating conjectures and developing arguments in support of these conjectures, dealing with counterexamples that contradict the conjecture or argument, and finally refining conjectures (or arguments). This process was iterative in nature so that the PSMTs had plenty of opportunities to engage in construction-evaluation-refining tasks in each class.

The PSMTs were informed about the study in the first class prior to the data collection process started and their consent was sought. All instructions were videotaped during three hours of the instruction time for 12 weeks in the semester. The video camera was located at the corner of the classroom where the board was captured. The conversations and the PSMTs' work on the board during whole class discussions were aimed to be captured by the video camera. Additionally, voice recorders were located at the tables to capture the groups' dialogs during instructions. The class videos served as the main data source for this study. In addition to the class videos, the PSMTs' written responses to the tasks and their class assignments were also collected to be analysed.

Data Analysis

A constant comparative method (Glaser & Strauss, 1967) was employed to construct a coding scheme: (1) the author independently reviewed all the videos of the instructions and created an initial coding scheme depending upon the related literature and the related proof definitions mentioned previously; (2) two research assistants, who were familiar with the literature related to reasoning-and-proving, and the author compared the descriptions of the codes in the preliminary coding scheme with the sample of responses to see whether the features of the responses captured by the codes or indicated any mismatches with the codes that could lead to the generation of new codes or adjustment of existing codes. After finalizing the coding scheme as displayed in Tables 2 and 3 by collaborating with the coders, coding of the data process started and occurred in two steps. In the first step of data analysis process, the parts of the instruction videos, where the PSMTs were engaged in constructing-evaluating-refining of mathematical conjectures and proofs were selected and transcribed. Later, the selected segments and the transcript of these segments were viewed again and the PSMTs' proof constructions were coded in one of the following categories: Type P_1: valid general deductive argument that justifies an assertion was true by standing of the underlying mathematical concepts, Type P_2: general argument that fall short of being acceptable proofs and Type P_3: unsuccessful attempt for a valid general argument (invalid, unfinished, or irrelevant responses (or potentially relevant response but the relevance was not made evident)). Categories Type P_1 through Type P_3 represent three different arguments constructed by the PSMTs in decreasing levels of sophistication (from a mathematical standpoint), with Type P_3 representing the least sophisticated argument. The constructions of conjectures were coded in one of the following categories: Type C_1: correct conjectures and Type C_2: incorrect conjectures.

After classifying types of proofs and conjectures constructed by the PSMTs, further attention was paid to Type P_2, P_3 arguments and Type C_2 conjectures given that the purpose of the study was also to understand the process of the PSMTs' ways of refining proofs and conjectures. Balacheff (1991) examined how students overcome contradiction when they encounter a counterexample and identified different approaches such as refuting conjectures, refining conjectures or considering counterexamples as exceptions. Larsen and Zandieh (2008) proposed a framework that describes three different student behaviours when dealing with counterexamples. Focusing on whether the students' attention was focused on the counterexample, definition, the conjecture and the proof, and whether the students' activity resulted in a modification to a definition or to the conjecture, Larsen and Zandieh (2008) categorized students' mathematical activities as monster-barring, exception-barring and proof analysis. The same approach was followed to make sense of the PSMTs' behaviour when they discovered a counterexample or missing parts of an argument, in other words, when they needed to refine a conjecture or a proof. The PSMTs behaviour to overcome contradiction when they encountered a counterexample or missing part of arguments were coded as: Rejecting the argument, Rejecting the conjecture, Refining the argument, and Refining the conjecture (Tables 2 and 3).

Table 2
Coding Scheme for Proof Constructing-Evaluating-Refining Activities

Conjecturing/working on a conjecture	Proving	Discovering missing part(s) of the argument/ a counterexample	Decision making
	Type P_1: Mathematical proofs		Acceptance of the argument
Type C_1: correct conjectures			Rejection of the argument
	Type P_2: General argument that falls short of being accepted as proofs	Discovering missing parts of the argument(s)	Refining the argument
Type C_2: incorrect conjectures	Type P_3: Unsuccessful attempt	Discovering missing parts of the argument(s) /Discovering a counterexample	Refining the conjecture

Table 2 demonstrates the PSMTs' behaviours for proof constructing-evaluating-refining activities. As seen in the table, the PSMTs constructed arguments when they verified the conjectures to be true. During evaluation of the constructed proofs, the PSMTs might choose accepting the constructed arguments as proofs (usually the case for Type P_1 arguments), rejecting the arguments or refining them (usually the case for Type P_2 and Type P_3 arguments). When the PSMTs faced a counterexample to the constructed arguments (usually the case for Type P_3

arguments), they either chose to reject the argument or to refine the conjecture in order to make the presented argument to work. Table 3 demonstrates the PSMTs' behaviours for refutation constructing-evaluating-refining process. The PSMTs constructed counterexamples when they verified the conjectures to be false. Their counterexamples were coded in one of the two categories as specific counterexample and general counterexample. During evaluation of the constructed counterexamples, the PSMTs' behaviours were coded either as rejecting the conjecture or refining the conjecture.

Two research assistants coded a random sample of 20% of the PSMTs' responses during classroom instructions. The coders reached an agreement on 85% of these codes, and all disagreements were resolved through discussion. Three classroom episodes will be provided to exemplify these codes later in the manuscript.

Table 3
Coding Scheme for Refutation Constructing-Evaluating-Refining Activities

Conjecturing/working on a conjecture	Refuting	Decision making
Type C_2: incorrect conjectures	Specific counterexample	Rejection of the conjecture
	General counterexample	Refining the conjecture

Results

General Findings

Table 4 summarizes the distribution of proof-conjecture constructions during the semester. As it was evident in the table, the majority of the arguments constructed during the class were Type P_1 arguments, valid general argument that justifies an assertion was true or false by showing how the statement of a theorem coheres and connects with the key properties of the concepts involved in the proof. Of the remaining argument construction occurrences, nine of them were Type P_2 arguments, general argument that fall short of being acceptable proofs.

Type P_3 arguments, unsuccessful attempt for a valid general argument, were proposed only four times during the semester; however, it should be noted that the PSMTs were aware of the limitations of these arguments. Therefore, they were able to evaluate those arguments as not proofs or as not correct argument during the class discussions. Of these four unsuccessful attempts to prove the class tasks (Type P_3 arguments), two arguments were empirical arguments. The PSMTs who proposed these empirical arguments as well as the others in the class were able to state the fact that generalizing from specific cases was not a valid mode of argumentation.

The number of the cases where conjectures were constructed during the class happened significantly less than the number of cases where proofs were proposed. Correct conjectures were proposed ten times and incorrect conjectures were proposed seven times during the instructions. Additionally, the majority of the cases where the conjectures were constructed occurred as a response to a request made usually by the instructor as opposed to constructing as a natural extension of the instruction. Samples of different types of arguments and conjectures constructed by the PSMTs will be provided next.

Table 4
Distribution of Proof and Conjecture Constructions during the Class

	Proofs (N=48)			Conjectures (N=17)	
	Type P_1	Type P_2	Type P_3	Type C_1	Type C_2
PSMTs	35	9	4	10	7

Classroom Episodes That Represent Different Types of Proof-Conjecture Constructions

In this section three classroom episodes is shared to exemplify some of the codes used to codify the PSMTs' proof and conjecture construction-evaluation-revision activities.

Episode 1. Chicken tender task. In this episode, the PSMTs were engaged in working on chicken tender task in their groups (see Table 1 for details).

1. Orhan[1]: Our group has decided that the numbers that can be represented as 28k+27, where k is an integer, cannot be bought in the packets of 7 and/or 4.
2. Instructor: Ok. Where did 28 and 27 come from?
3. Orhan: 7 times 4 is 28, so 28 can be bought in packets of either 7 or 4 or its multiples.
4. Instructor: OK, if k=1 then how many nuggets do you think you cannot buy, umm, 55?
5. Orhan: Yes
6. Instructor: Can we have 55 nuggets in the packets of 7 and/or 4? Merve?
7. Merve: Yes. We can have 5 packets of 7s and 5 packets of 4s. So, we can get 55 nuggets in total.
8. Instructor: Selman?
9. Selman: You can represent all numbers by adding 4. For instance, if you add 4 to 11, you will get 15; if you add 4 to 12 you will get 16; if you add 4 to 14 you will get 18 and it will continue like this. These numbers cannot be bought in packets of 7 or 4 (highlighting the numbers underlined below in Figure 1). Umm, I needed to check 21 because it is to add 4 to 17 and I know 17 cannot be represented as addends of 4 (or multiples of 4) and 7 (or multiples of 7). However, I have found that 21 is 3 times 7, so it is okay too. Now you will continue this pattern for all numbers 21+4=25, 22+4=26, 23+4=27, 24+4=28...etc.

1	2	3	4	5	6	7	8	9	10
11	12	<u>13</u>	14	15	16	<u>17</u>	18	19	20
21	22	23	24	25	26	27	28	29	30

Figure 1. Selman's work on the board

10. Ayse: When you get modulo 7, the residue classes will be 1, 2, 3, 5 or 6. 4 cannot be counted here because we can get the packets of four). When you get modulo 4, the residue classes will be 1, 2 and 3. When you add all residue classes up, you will get the answer—17.
11. PSMT: I think she just got lucky!
12. Instructor: Will this work all the time? Let's try for different cases. What if the restaurant sells the chicken in packets of 6 and 5, what would be the largest number that you can not buy? Let's try both Selman's and Ayse's way.
13. PSMTs: (Trying the methods in their groups).

PSMTs: Ayse's way won't work.

[1]All names used in the study are pseudonyms

In this episode, the conjecture proposed by Orhan was coded as Type C_2: incorrect conjecture. The instructor posed a specific counterexample (55 nuggets when *k*=1) as a necessary condition for the realization of falsity of the conjecture. The PSMTs were able to explain how the example 55 contradicted the conjecture and decided to reject Orhan's conjecture immediately. Selman's argument was coded as Type P_1 since it was built on the properties of numbers. It was a correct argument to justify that 17 was the highest number of chicken tenders that could not be bought in packets of 4 and/or 7. Ayse identified all residue classes of modulo 7 except 4 (since it could be a possible answer) and added them up to reach the answer of 17. However, her argument did not include a justification for the assertion that the residues would always be the highest number that could not be bought in the packets of 4 or 7. Indeed, the instructor suggested trying her method for different cases. The class then realized that when her argument was applied to different numbers such as packets of 6 and 5, it would not give the correct response. Therefore, her argument was coded as Type P_3 and got rejected by the PSMTs. Many researchers examined how students overcome contradiction when they encounter a counterexample and identified different approaches such as refuting the conjecture, refining the conjecture, considering the counterexample as an exception, refuting the argument or refining the argument (Balacheff, 1999; Larsen & Zandieh, 2008). The same approach was followed by the PSMTs in this episode. When the PSMTs were guided to discover a counterexample to Ayse's argument, they then chose to reject the argument instead of refining it.

Episode 2. Area perimeter task. In this episode, the PSMTs were engaged in working on geometry task—investigating the relationship between area and perimeter of rectangles. The instructor asked the PSMTs to construct conjectures about area and perimeter of rectangles. Cihat proposed the following conjecture: "With the same perimeter, the smaller the difference between the side lengths of a rectangle, the biggest the area". The instructor asked the PSMTs to evaluate the conjecture and prove whether it was correct.

1. Merve: (Drawing three rectangles with the side lengths of 12 by 6, 15 by 3, and 9 by 9). It is true. These rectangles have the same perimeter, 36. But, the area of the square is bigger than the other two rectangles.
2. Instructor: Do you think that Merve proved Cihat's conjecture?
3. PSMTs: No, she just demonstrated for those rectangles.
4. Instructor: What is missing in her argument?
5. PSMTs: It is not general!
6. Instructor: How many examples can I draw with a perimeter of 36?
7. PSMTs: 5? (Said as if they were asking if it was true). Infinitely many?
8. Cihat: Infinitely many, because in between whole numbers, there are infinitely many rational numbers
9. Instructor: So we can draw infinitely many rectangles with the perimeter of 36, will you be able to try all of these rectangles out like Merve attempted to do here?
10. PSMTs: No!
11. Yılmaz: Now we have the lengths of b,c (referring to the long and short sides of a rectangle in this order) and x (referring to a side length of a square). They should have the following relationships: $b>x>c$. Thus, $x^2>b.c$
12. Let's assume that $x=n$ and $c=n-1$ and $b=n+1$. Therefore, $x^2=n^2>b.c=n^2-1$

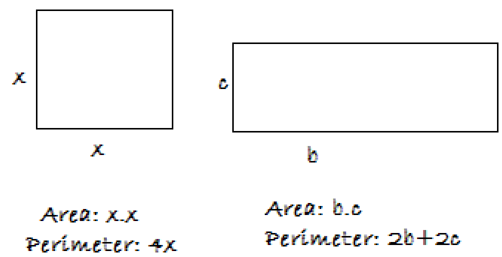

Figure 2. Yılmaz's work on the board

13. Instructor: Why does b have to be bigger than x and x has to be bigger than c?
14. PSMTs: If these two rectangles have the same perimeter, than this relationship should hold.
15. Instructor: Cihat?
16. Cihat: b and c should be different in lengths, because we consider the rectangles that are not squares, so $b \neq c$. Then, we know that $x=\frac{b+c}{2}$ since the perimeter of the two shapes should be equal. Thus x should be between b and c. We know that $x= \frac{b+c}{2}$ so, $x^2= \frac{b^2+c^2+2bc}{4}$. We know that $b-c>0$, so $(b-c)^2>0$. $b^2-2bc+c^2>0 \Rightarrow b^2+c^2>2bc$. If $b^2+c^2>2bc$, Then $\frac{b^2+c^2+2bc}{4}$ should be bigger than b.c [the area of the rectangle].

Cihat's conjecture was constructed as a response to the instructor's request and it was a correct conjecture. Therefore, it was coded as Type C_1. Merve provided three examples that demonstrated that Cihat's conjecture was true. Since Merve used an invalid mode of argumentation (inductive argument), her argument was coded as Type P_3. Stylianides (2007b) argues that the main difference between empirical arguments and proofs lies in the modes of argumentation: invalid versus valid modes of argumentation. Empirical arguments provide inconclusive evidence by verifying the statement's truth only for a proper subset of all covered by the generalization, whereas proofs provide conclusive evidence truth by treating appropriately all cases covered by the generalization. When asked to evaluate the argument, both Merve and the other PSMTs in the class were able to state this limitation of the argument. Stylianides and Stylianides (2009) argued that construction-evaluation tasks can better identify prospective teachers' who seem to possess the empirical justification scheme. Unlike Merve, Yilmaz attempted to construct a deductive argument. However, his argument did not provide justification for some of the assertions he used in his argument (i.e., $b>x>c$). Additionally, Yilmaz's argument was constructed based on a condition—the side lengths of the rectangle and the square should be consecutive. Yilmaz's argument was coded as Type P_2. Unlike Merve's argument, which was get rejected immediately by the PSMTs, Yilmaz's argument was modified. Cihat was able to provide the justifications for each step of his argument.

Episode 3. Area of a kite. In this episode, the PSMTs were engaged in working on a geometry task—justifying the area formula of a kite. The instructor asked the PSMTs to justify that the area formula of a kite is $\frac{d_1 d_2}{2}$. Cihat justified the formula by drawing a rectangle around the kite and showing that the area of the kite is half of the area of the rectangle where the diagonals of the kite form the sides of the rectangle. Cihat's justification resulted in Selman's conjecture of

"The area of the quadrilateral formed by joining the four points on the consecutive sides of a rectangle is always half of the area of the rectangle".

1. Instructor: Ok, what was the area formula of a kite?
2. PSMTs: Multiply the diagonals
3. Instructor: Let's call the diagonals as d_1 and d_2, so A=$\frac{d_1 d_2}{2}$
4. Instructor: Why does this formula works?
5. Cihat: The area of the kite will be half of the area of the rectangle.

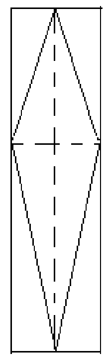

Figure 3. Cihat's work on the board

6. Instructor: What is the side of this rectangle?
7. PSMTs: d_1 and d_2 (referring to the diagonals of the kite)
8. Instructor: $d_1.d_2$ is the area of the rectangle and there are congruent triangles in this rectangle (pointing to the congruent triangles). You can see that the area of the kite is exactly half of the area of the rectangle around.
9. Selman: It does not have to be a kite, the quadrilateral inside the rectangle. If the corners of a quadrilateral are on the consecutive sides of a rectangle, then the area of the quadrilateral inside is always half of the area of the rectangle.

Another lesson….

10. Instructor: We have talked about kite last week. How did we calculate the area of a kite?
11. PSTs: $\frac{d_1 d_2}{2}$
12. Instructor: We have proved it last time. Selman said that the shape inside could be any quadrilateral, but the area would still be half of the rectangle. Do you guys remember?
13. PSMTs: Yes
14. Instructor: We did not spend much time on this last time. We just said yes it could be true. Now, I would like you to spend more time on Selman's conjecture and decide whether or not it would always be true and then prove your decision.
15. Omer: We have found that it was not true.
16. Instructor: Not true?
17. Let's talk all together. How many of you think that the conjecture is true? 3? I think you have not decided yet. Let's hear from someone who thought it was true. Ahmet, can you come up to the board to explain?
18. Ahmet: (Drawing a rectangle on the board).

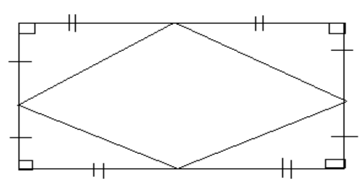

Figure 4. Ahmet's work on the board

19. Omer: He used the midpoints. Will it be true if the corners of the inner quadrilateral were not the midpoints?
20. Instructor: Let's look at why it would work for this case. Are these vertices opposite to each other? (Connecting a pair of opposite vertices). So what can we say about the areas of these triangles (referring to the triangles formed by constructing a diagonal of the inner quadrilateral)?

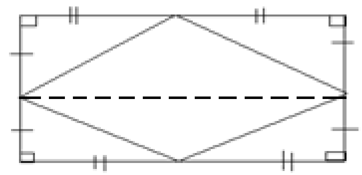

Figure 5. Instructor's drawing on Ahmet's work

21. PSTMs: Same.
22. Instructor: Why?
23. PSMTs: All sides are equal in length.
24. PSMTs: Congruent triangles.
25. Omer: Half of the area of the rectangle around them [referring to the two rectangles formed by the line that connects the two opposite midpoints].
26. Ahmet: [Trying to calculate if the vertices were not the midpoints].

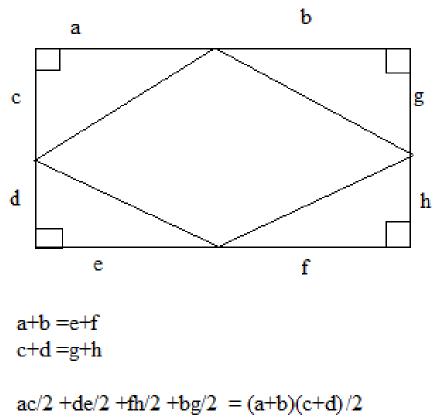

Figure 6. Ahmet's attempt to refine his argument

27. Instructor: While you are working on this, is there anybody else who would like to share?
28. Durmus: (Drawing his example on the board).

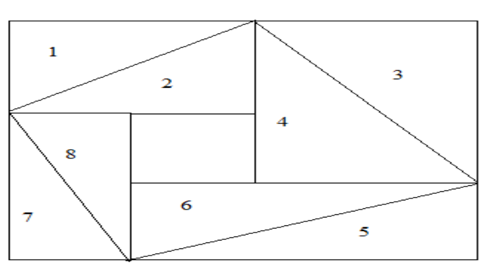

Figure 7. Durmus' counterexample

29. Instructor: Durmus showed us that the area of the inner quadrilateral is sometimes bigger than half of the area of the rectangle. Let's look at how we could use Geometer's Sketchpad (GSP) to investigate a conjecture like this. (Projecting a rectangle constructed by GSP on the board). I am going to show you how I constructed this rectangle in a second, but let's find the midpoints of each side and connect them. What shape would be constructed?

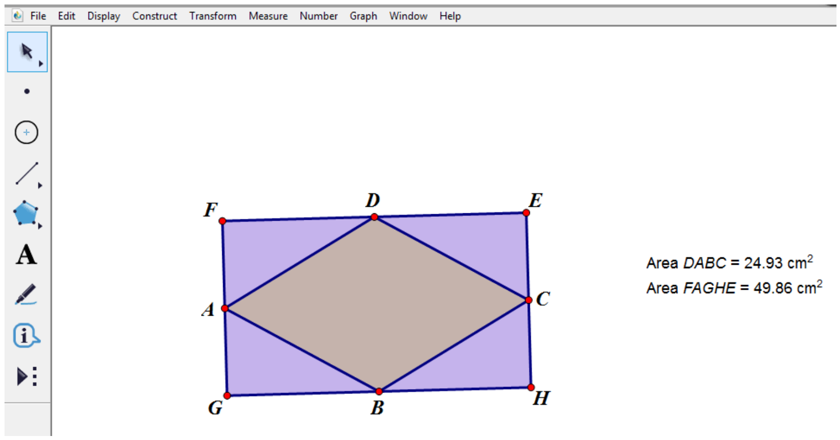

Figure 8. GSP demonstration 1

30. PSMTs: Rhombus
31. Instructor: Rhombus. Would each side of the shape be the same?
32. PSMTs: Yes.
33. Instructor: When I calculate the area of the rhombus, it will be 24,93 and the area of the rectangle is 49,86. It is exactly twice as much. Let's investigate what would happen if we did not choose the mid points. Let's construct a rectangle. What is the definition of a rectangle?
34. Merve: It has opposite sides parallel.
35. Instructor: Would it be enough to say that it is a quadrilateral with opposite sides parallel?
36. PSMTs: No.
37. Instructor: Parallelogram has also this property but it is not always a rectangle.
38. PSMTs: The sides are perpendicular.
39. Instructor: It is a quadrilateral with all 90-degree angles. Do you need to include opposite sides being parallel after including 90-degree angles?
40. PSMTs: No.
41. Instructor: Having right angles would ensure opposite sides being parallel. So, yes we do not need to add that property as well. In order to construct a rectangle, you need to have 4 right angles. (Constructing a rectangle on GSP). Let's select four random points on this rectangle. When I calculated the area of the inner quadrilateral it is 23cm^2 and the rectangle is 49,86 cm^2. It would not always hold true (Dragging the points to show that the area would not always be half of the area of the rectangle). However, when two of the vertices of the inner quadrilateral became exact opposite, the conjecture would hold true. Why?

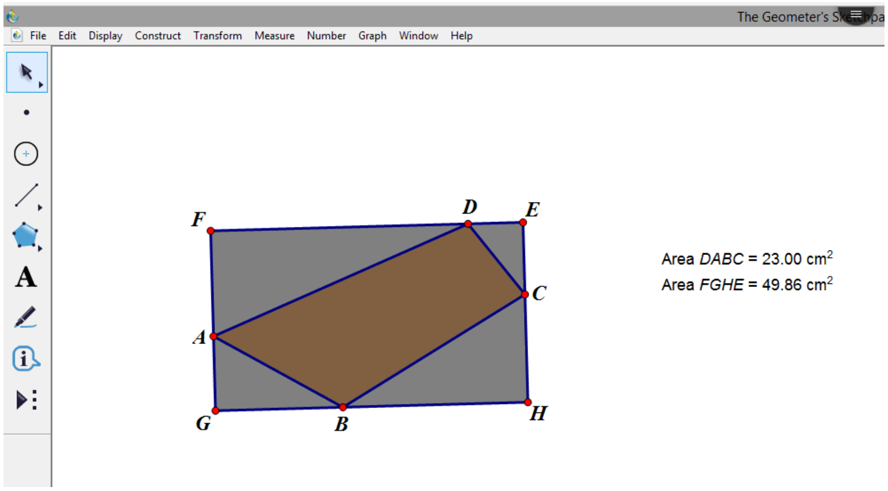

Figure 9. GSP demonstration 2

42. PSMTs: (Silence).
43. Instructor: Let me give you a hint (Constructing a line that connects opposite vertices).

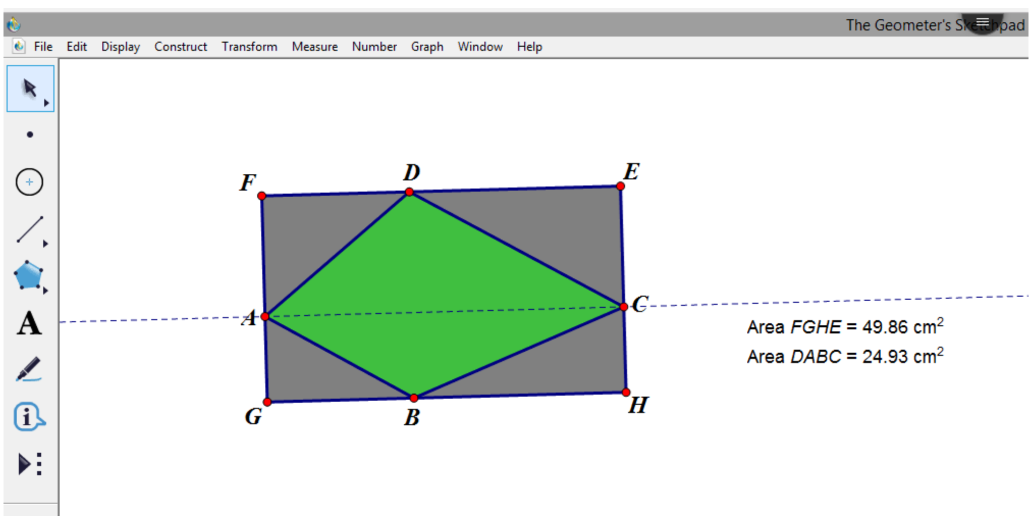

Figure 10. GSP demonstration 3

44. Orhan: If we take the two rectangles formed by the line you just drew through A and C, the areas of the triangles inside would always be half of the areas of the rectangles around them, no matter where point D and B would be moved. Because, the bases of the triangles will be the same and so do the heights of them.
45. Instructor: Did everyone understand what Orhan just explained?
46. PSMTs: Yes.

In this episode, Selman posed a conjecture spontaneously as a natural extension of Cihat's proof to the kite task, which was coded as Type C_2 since it was false. Although the PSMTs did not recognize that it was a false conjecture at first, they did in the next class when the instructor asked them to evaluate the conjecture. Ahmet constructed a proof based on a specific condition—choosing midpoints of a rectangle. His proof was coded as Type P_3 since it was only a general argument to prove not the original conjecture but the restricted conjecture—based on choosing the midpoints as opposed to choosing arbitrary points on the consecutive sides of a rectangle. Although he attempted to refine his proof when the shortage of his proof was mentioned by another PSMT, he could not complete modification of his initial proof (see Figure 6). Therefore, the class rejected his proof. Marrades and Gutierez (2000) argued that failed justifications are necessary to conceptualize proofs since justification and proof skills cannot be associated only to correct ones. When the instructor prompted the PSMTs to analyse Ahmet's proof (Type P_3), it provided a forum to launch justification and better conceptualize mathematical relations, which resulted the PSMTs to modify Selman's conjecture as: "The area of the quadrilateral formed by joining the four points on the consecutive sides of a rectangle is half of the area of the rectangle if at least one of the diagonals of the inner quadrilateral is parallel to the respective sides of the rectangle". All these results are discussed under the light of current research findings next.

Discussion and Conclusion

Incorporating proofs as an essential part of mathematics classrooms has recently been emphasized by both mathematics educators (Ball & Bass, 2000; Ball, Hoyles, Jahnke, & Movshovitz-Hadar, 2002; Knuth, 2002; Stylianides, 2007b; Yackel & Hanna, 2003) and by policy makers (e.g., Ministry of National Education [MEB], 2018; NGA & CCSSO, 2010). Current reform efforts are calling for substantial changes in the nature and role of proof in school mathematics (e.g., MEB, 2018; NCTM, 2000).Yet, it is widely documented that students at all levels struggle with proofs (Balacheff, 1988; Bell, 1976; Chazan, 1993; Coe & Ruthven, 1994; Harel & Sowder, 1998; Healy & Hoyles, 2000; Knuth, Choppin, & Bieda, 2009). The majority of the arguments constructed by the PSMTs was Type P_1 arguments—valid deductive arguments that justifies an assertion was true or false by showing how the statement of a theorem coheres and connects with the key properties of the concepts involved in the proof. Although the number of times in which the PSMTs constructed valid mathematical arguments—Type P_1— were higher, the PSMTs also proposed arguments that fell short of being accepted as mathematical proofs which were referred as Type P_2 and invalid arguments including empirical arguments—Type P_3 —during the semester. Many researchers demonstrated that the empirical reasoning is pervasive among school students including advanced or high-attaining secondary students (e.g., Coe & Ruthven, 1994; Healy & Hoyles, 2000), university students including mathematics majors (e.g., Goetting, 1995) as well as prospective and in-service teachers (e.g., Morris, 2002; Simon & Blume, 1996). However, these studies focus mainly on individuals' proof constructions or their evaluation of researcher-generated arguments. Stylianides and Stylianides (2009) reported that some prospective teachers in their study provided empirical argument while being aware that their constructions were invalid. Other researchers reported similar results (e.g., Weber, 2010). The results of this study align well with those studies. The PSMTs who proposed Type P_2 and Type P_3 arguments were not convinced entirely that their arguments could be counted as proofs. Especially the constructed arguments, which were coded as Type P_3 during the semester, were evaluated as invalid by the PSMTs and got rejected immediately (i.e., Line #10-14 in Episode 1 and Line# 1-5 in Episode 2). The constructed arguments which were coded as Type P_2, on the other hand, were more often chosen to be refined by the PSMTs as opposed to get rejected (i.e., Line #11-18 in Episode 2). Although these findings could be interpreted as that the PSMTs were aware of the limitations of Type P_2 and Type P_3 arguments and recognized that valid deductive argument was needed, yet they still had difficulty in constructing sufficient arguments. Thus, the question remains is that whether the arguments that fall short of being accepted as valid proofs and invalid arguments such as Type P_2 and Type P_3 arguments in the study can develop into valid proofs—Type P_1.

Stylianides and Stylianides (2009) argued that to comprehend individuals' understanding of proof and to help them recognize the need for mathematical proofs, the combined "construction-evaluation" activities could be beneficial. This study demonstrated that in addition to constructing arguments and evaluating whether those arguments could be counted as mathematical proofs, refining the ones that fall short of being acceptable proofs can better illuminate pre-service teachers' understanding of what constitute a mathematical proof and help them better conceptualize what mathematical proofs consist of. Knuth (2002) argues that one way to incorporate mathematical proofs meaningfully into mathematics classrooms is to have student present their arguments. This study demonstrated that classroom presentations of different arguments even the ones that fall short of being acceptable as proofs and refining them could furnish a forum for discussing with students the question of what constitutes a proof. Marrades and Gutierez (2000) argued that failed justifications are necessary because the assessment of students' justification and proof skills cannot be associated only to correct ones. Thus, refining proofs that fall short of being acceptable proofs in classrooms could be an essential instructional tool to help pre-service teachers comprehend what constitutes proof.

This study also demonstrated that in addition to refining proofs, conjecturing could also be considered to not only launch justification but also to develop what it means to make mathematics. Stylianides (2008) used a hyphenated term—reasoning-and-proving—to encompass four related activities in which one of them was making conjectures. In recent years, researchers draw increasing attention to conjecturing as an essential step to launch justification (Flores, 2002; Keith, 2006). Conjecturing leads students to think about mathematical ideas that constitute a basis for mathematical argumentation and undoubtedly, plays a significant role in an inquiry-based learning process (Cañadas et al., 2007). The results of the study demonstrated that the number of instances where the PSMTs constructed conjectures, which is referred as one of the essential parts of the process of making sense of and establishing mathematical knowledge (Stylianides, 2008), were limited and usually occurred when asked explicitly. The PSMTs constructed ten conjectures that were correct, which were coded as Type C_1, seven conjectures that were incorrect, which were coded as Type C_2. Cañadas and her colleagues (2007) argue that not all problems lead to conjecturing and different problems lead to different kinds of conjecturing. Align with this result; this study demonstrated that not all tasks encouraged the PSMTs to construct conjectures during the classes. Among these seventeen conjecturing cases, the majority of the constructions were constructed when the PSMTs were engaged in conjecturing tasks such as a geometry task that asked PSMTs explicitly to investigate about relationships between areas and perimeters of rectangles and to propose conjectures. However, very few justification tasks resulted into PSMTs' conjecture constructions (i.e., Selman's conjecture in Episode 3 in Line #9). It should be mentioned that Cañadas and her colleagues (2007) suggested ways of turning problems to prove (justification) tasks into problems to find (conjecturing) tasks. However, since the purpose of this study was not solely on conjecturing but also proving, both types of tasks were aimed to be included during instructions.

Harel and Sowder (1998) stated that "a conjecture is an observation made by a person who has doubts about its truth" (p. 241). Fischbein (1982), on the other hand, considered conjectures as expressions of intuitions. Although Cihat's proof to the kite task in Episode 3 could be considered as an observation about a specific case, Selman based his conjecture more on his intuition than this observation. Cañadas and her colleagues (2007) laid out several types of conjectures in which one of them was called as perceptually based conjecturing. They stated "the basis of this type of conjecturing is a careful representation of the content of the problem either concretely or as a mental image. There is not an immediate attention to the relationships existing between the problem's elements because instead the initial focus is on creating a new representation of the problem (p.61)". In Selman's conjecturing case, Cihat's proof for the area formula of a kite and his representation of his justification into a graphical form encouraged Selman to formalize his conjecture and then generalize. It was evident in classroom episodes that both Selman's conjecture in Episode 3 and Cihat's conjecture in Episode 2 provided a forum to launch justification and better conceptualize mathematical relations. Lampert (1992) argued that conjecturing about relationships is at the heart of mathematical practice. Therefore, conjecturing and refining conjectures could be considered as an essential approach to not only launch justification but also to better comprehend mathematical ideas and relations.

References

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

Balacheff, N. (1991). Treatment of refutations: Aspects of the complexity of a constructivist approach to mathematics learning. In E. von Glasersfeld (Ed.), *Radical constructivism in mathematics education* (pp. 89–110). Dordrecht: Kluwer Academic Publishers.

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on mathematics of teaching and learning* (pp. 83-104). Westport, Conn.: Ablex Publishing.

Ball, D. L., Hoyles, C., Jahnke, H. N., & Movshovitz-Hadar, N. (2002). The Teaching of Proof. In L.I. Tatsien (Ed.). *Proceedings of the International Congress of Mathematicians* (Vol. III, pp. 907–920). Beijing: Higher Education Press.

Bell, A. W. (1976). A study of pupils' proof –explanations in mathematical situations. *Educational Studies in Mathematics, 7* (1/2), 23-40.

Canadas, M.C., Deulofeu, J., Figueiras, L., Reid, D., & Yevdokimov, O. (2007). The conjecturing process: Perspectives in theory and implications in practice. *Journal of Teaching and Learning, 5*(1), 55-72.

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

Coe, R., & Ruthven, K. (1994). Proof practices and constructs of advanced mathematics students. *British Educational Research Journal, 20*(1), 41-53.

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

de Villiers, M. D. (1999). The role and function of proof. In M. de Villiers (Ed.), *Rethinking proof with the Geometer's Sketchpad* (pp. 3–10). Key: Curriculum Press.

de Villiers, M. D. (2012). An illustration of the explanatory and discovery functions of proof. *Pythagoras, 33*(3), 193-201.

Duval, R. (2002). Proof understanding in mathematics: What ways for students? *Proceedings of 2002 international conference on mathematics: Understanding proving and proving to understand* (pp. 61–77).

Fischbein, E. (1982). Intuition and proof. *For the Learning of Mathematics, 3*(2), 9-18

Flores, A. (2002). How do children know what they learn in mathematics true? *Teaching Children Mathematics, 8*(5), 269-274.

Glaser, B, & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research.* New York: Adline de Gruyter.

Goetting, M. (1995). *The college students' understanding of mathematical proof* (Unpublished doctoral dissertation). University of Maryland: College Park.

Griffiths, P. A. (2000). Mathematics at the turn of the millennium. *American Mathematical Monthly, 107*, 1-14.

Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange, 21*(1), 6-13.

Hanna, G., & Barbeau, E. (2008). Proofs as bearers of mathematical knowledge. *ZDM –The International Journal on Mathematics Education, 4*0, 345-353.

Hanna, G., & Jahnke, H. N. (1996). Proof and proving. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp.877-908). Dordrecht: Kluwer Academic Publishers.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput, & E. Dubiensky (Eds*.), Research in collegiate mathematics education III* (pp. 234-283). Providence, R.I.: American Mathematical Society.

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 805–842). Greenwich, CT: Information Age Publishing.

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

Hersh, R. (1993). Proof is convincing and explaining. *Educational Studies in Mathematics*, *24*(4), 389-399.

Keith, A. (2006). Mathematical argument in a second grand class: generating and justifying generalized statements about odd and even numbers. In S. Z. Smith & M. E. Smith (Eds.), *Teachers engaged in research: Inquiry into mathematics classrooms, grades preK-2* (pp. 35-68).

Knuth, E. J. (2002). Proof as a tool for learning mathematics. *Mathematics Teacher*, *95*(7), 486-490.

Knuth, E. J., Choppin, J., & Bieda, K. (2009). Middle school students' productions of mathematical justification. In M. Blanton, D. Stylianou, & E. Knuth (Eds.) *Teaching and Learning Proof Across the Grades: A K-16 Perspective* (pp. 153-170). Routledge: New York.

Lampert, M. (1992). Practices and problems in teaching authentic mathematics. In F. K. Oser, A. Dick, & J. Patry (Eds.), *Effective and responsible teaching: The new synthesis* (pp. 295–314). San Francisco: Jossey-Bass Publishers.

Lannin, J., Ellis, A., Elliott, R., & Zbiek, R. (2011). *Developing essential understanding of mathematical reasoning for teaching mathematics in grades pre-K-8*. Reston: NCTM.

Larsen, S., & Zandieh, M. (2008). Proofs and refutations in the undergraduate mathematics classroom. *Educational Studies in Mathematics*, *67*(3), 205-216.

Mariotti, M. A. (2000a). Introduction to proof: The mediation of a dynamic software environment. *Educational Studies in Mathematics*, *44*, 25-53. 66

Mariotti, M. A. (2006). Proof and proving in mathematics education. In A. Gutiérrez & P. Boero (Eds.), *Handbook of research on the PME: Past, present and future* (pp. 173-204). Rotterdam, The Netherlands: Sense Publishers.

Marrades, R., & Gutiérrez, Á. (2000). Proofs produced by secondary school students learning geometry in a dynamic computer environment. *Educational Studies in Mathematics*, *44*, 87-125.

MEB (2018). *Matematik dersi öğretim programı (ilkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. sınıflar)*. Retrieved from https://ttkb.meb.gov.tr.

Morris, A. K. (2002). Mathematical reasoning: adults' ability to make the inductive-deductive distinction. *Cognition and Instruction*, *20*(1), 79–118

National Council of Teachers of Mathematics (NCTM) (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.

National Governors Association Center for Best Practices, & Council of Chief State School Officers (NGA & CCSSO) (2010). *Common core state standards mathematics*. Washington, DC: Authors.

Simon, M. A., & Blume, G. W. (1996). Justification in the mathematics classroom: A study of prospective elementary teachers. *Journal of Mathematical Behavior*, *15*, 3-31.

Stylianides, A. J. (2007a). Introducing young children to the role of assumptions in proving. *Mathematical Thinking and Learning*, *9*, 361-385.

Stylianides, A. J. (2007b). Proof and proving in school mathematics. *Journal for Research in Mathematics Education*, *38*, 289–321.

Stylianides, G. (2008). An analytic framework of reasoning-and-proving. *For the Learning of Mathematics*, *28*(1), 9-16.

Stylianides, A.J., & Stylianides, G. J. (2009). Proof constructions and evaluations. *Educational Studies in Mathematics*, *72*, 237-253.

Weber, K. (2002). Beyond proving and explaining: Proofs that justify the use of definitions and axiomatic structures and proofs that illustrate technique. *For the Learning of Mathematics*, *22*(3), 14-17

Weber, K. (2003). Students' difficulties with proof. *MAA Research Sampler 8*. Retrieved from http://www.maa.org/t.

Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning*, *12*, 306–336.

Wilburne, J. M. (2014). What is the largest number you cannot make? *Teaching Children Mathematics.* Retrieved from http://www.nctm.org/Publications/Teaching-Children-Mathematics/Blog/What-is-the-Largest-Number-You-Cannot-Make_/

Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 22-44). Reston, VA: National Council of Teachers of Mathematics.