

A study on pre-service mathematics teachers' criteria of proof evaluation

Tuğba Öztürk and Bülent Güven

Trabzon University, Turkey

Proof is foundational to mathematics, and constructing proofs and establishing their validity are both important mathematical studies. Determining the validity of a proof is a part of the process of proof evaluation. Proof evaluation contributes to students' ability to construct and revise their own proofs. The purpose of this study was to determine the criteria pre-service mathematics teachers take into account in evaluating a proof. The study was carried out with 50 first-year university students enrolled in an elementary mathematics teacher education program. The data were collected through activities relating to proving. The results of this study revealed that, when evaluating a proof, the participants regarded use of appropriate definitions, axioms, or theorems in the steps of the proof as the most important criterion with reference to justification, while in terms of mathematical language; they regarded appropriate use of symbolic language as the most important criterion. However, they tended to ignore situations where non-symbolic language was used. To address this issue, it is recommended that mathematics-learning environments include the use of non-symbolic language, as well as symbolic representations, in the definition of mathematical concepts.

Keywords: Proof, proof evaluation, justification, mathematical language

1 Introduction

Mathematical proof, being foundational to mathematics, is a concept that distinguishes mathematics from other disciplines. Scholars have for this reason often defined the concept of proof as the heart of mathematics (Anderson, 1996; Hanna, 2000; Herbst & Brach, 2006; Hoyles, 1997; Öztürk, 2020; Rav, 1999). Flores (2002) notes that recognition of the relationships between mathematical concepts, realization of mathematical thinking, and making sense of mathematical concepts all depend on proof. As such, the stages of proof writing have an important place in the formation of the specific language of mathematics (Forman, Larreamendy-Joer, Stein & Brown, 1998) and proving should be an essential part of learning mathematics for students of all ages (Stacey & Vincent, 2009). Furthermore, reasoning and proof are each one of significant activities characterized as working mathematically (Clarke, Goos & Morony, 2007). The scope of reasoning proficiency is directly associated with proof at all levels of mathematics education (Lesseig, Hine, Na & Boardman, 2019). This proficiency includes several actions as analyzing, proving, evaluating, explaining,

ARTICLE DETAILS

LUMAT General Issue
Vol 10 No 1 (2022), 210–241

Received 9 August 2021
Accepted 28 April 2022
Published 19 May 2022

Pages: 32
References: 77

Correspondence:
tugbaozturk@trabzon.edu.tr

[https://doi.org/10.31129/
10.1.1649](https://doi.org/10.31129/10.1.1649)



inferring, justifying, and generalizing (Clarke, Clarke & Sullivan, 2012). National Council of Teachers of Mathematics [NCTM] (2000) particularly recommended that importance be attached to the teaching of proof at all levels of mathematics education.

Proving is a comprehensive process that involves several actions as exploring, conjecturing, reasoning, and formulating arguments (Stylianides & Ball, 2008). Furthermore, this process occurs in stages (Hanna, Bruyn, Sidoli & Lomas 2004) representing these kinds of the actions. Many researchers also identified several stages for the proving process and these stages can be followed in the proof-related activities (e.g., Boero, 1999; Edwards, 1997; Perry, Molina, Camargo & Samper, 2011). Thompson, Senk and Johnson (2012) state that the proof-related activities include some stages as make a conjecture, investigate a conjecture, evaluate an argument or proof, find a counter-example to a mathematical claim. Öztürk (2016), expressing that proving is a multi-step process, defined an instructional model that can be used in teaching proof. This model consists of seven stages that reflects the nature of a mathematical proof such as *understanding the problem, working on the structure and conjecturing, postulation of the relationship, proving, investigation of the coherence of the proof*, [and] *formalization of the proof*. *Investigation of the coherence of the proof* as a stage of the model involves examining the validity and coherence of the proofs and expressing any deficiencies and errors, and so it represents the action of proof evaluation. Therefore, proof evaluation is one of the significant actions that construct the final version of the proof.

According to Pfeiffer (2011a), the most significant part of the process of accepting a proof is validation, as a determination of the correctness of an argument. During determining when a mathematical argument becomes a proof of a theorem, mathematicians employ several criteria to decide whether it can be a new mathematical knowledge (Dickerson & Doerr, 2014). Bass (2009) emphasizes the social aspect of proof, noting the act of proving and accepting new proofs constitutes certification of new knowledge within a community of practice. This notion that is upheld by Manin's (2010) assertion that acceptance of a proof occurs as a social process. Given the critical role of proof in the mathematical community, Powers, Craviotto and Grassl (2010) emphasize that proof validation skills are important for every individual who will be involved in the teaching of mathematics.

2 Proof validation and evaluation in mathematics practice

Examining a mathematical proof with the aim of determining whether it is valid is a significant activity for pre-service teachers and those who may be involved in instruction or training as a graduate student or supervisor (Powers et al., 2010; Selden & Selden, 2000); however, Inglis and Alcock (2012) indicate that validation is only one of the reasons for reading a proof. As Pfeiffer (2011b) points out, proof validation falls within the scope of proof evaluation, which Selden and Selden (2003a) describe as a complex process involving evaluating statements, posing and answering questions, constructing sub-proofs, and recalling definitions and theorems. Thus, evaluating a proof is useful in determining its appropriateness with reference to a range of features, such as its relation to the functions of proof (*exploring, conjecturing, generalizing, association, etc.*), as well as establishing the truth of the statement (*validation*) (Pfeiffer, 2011a). In other words, proof evaluation involves assessment of the significance and merits of a proposed proof.

In this respect, Tabach et al. (2010) emphasize that proof evaluation embraces not only the validity of a proof, but also an assessment of evaluators' knowledge as it comes to light during the process. Similarly, Pfeiffer (2011b) states that proof evaluation includes not only determining whether proof is correct but also establishing the truth of a statement (*validation*). Proof evaluation is more comprehensive than proof validation; while proof validation is restricted to determining whether a proof is successful in establishing the truth of a statement, proof evaluation examines proofs in terms of dimensions such as mathematical language, justifications, proof steps, and so on (Moore, 2016).

The activities of validating and evaluating proofs in an educational context enable students to acquire insights about their own and others' existing mathematical knowledge, as well as learning to follow the steps necessary to carry out a proof (Stylianides & Stylianides, 2009). The ability to validate proofs (i.e., to determine their correctness) has a positive effect on proof writing (Powers et al., 2010). The more adept the writers of proofs become at determining whether a proof written by a peer is correct, the better able they will be to apply this ability to their own proofs, as with Selden and Selden's (2003a) contention that the ability to validate proofs relates to the ability to construct them. Furthermore, as a more comprehensive activity than proof validation, proof evaluation significantly contributes to students' ability to construct and revise their own proofs. However, many researchers have emphasized that not only the construction but also the validation and evaluation of proofs present

a significant challenge for many students (Inglis & Alcock, 2012; Lin, Yang & Chen 2004; Pfeiffer, 2011a; Selden & Selden; 2003a; Selden & Selden, 1995; Varghese, 2011). Pfeiffer (2011a) noted that students generally focus on the presence or absence of examples or of algebraic formalism in their evaluations, while Selden and Selden (2003a) expressed that students mainly concentrate on surface features such as algebraic notation and computation in their evaluation of proofs.

Inglis and Alcock (2012) assert that there are several reasons for students' difficulties in proof validation, including insufficient understanding of what constitutes a proof and a tendency to focus only on certain proofs. Furthermore, according to Knuth (2002), even teachers have serious difficulties with the task of proof validation, and many are prepared to accept flawed arguments as valid mathematical proofs. As such, both mathematics students and pre-service mathematics teachers should be exposed to proof validation and evaluation applications in the course of their learning. In this manner, students at all educational levels may be supported in making sense of their mathematical knowledge, as well as understanding what constitutes a proof and how a proof is constructed. Pre-service teachers, in particular, should be exposed to these applications, so that they will be better equipped to develop their future students' ability to construct and evaluate proofs. In this respect, it is important to determine what factors pre-service teachers take into account in the evaluation of proofs. In addition, studies about proof evaluation (Knuth, 2002; Stylianides & Stylianides, 2009; Weber, 2009) were limited and the applications in these studies are generally performed as giving various proofs to a certain group of participants and asking them for evaluating the proofs. Their mathematical context was also mainly algebra, number theory and calculus. In this study, carrying out both proof construction and evaluation activities, their implementation through group work and the fact that its mathematical context was geometry were the originality of the study.

Proving is one of the general objectives of geometry education (Marrades & Gutiérrez, 2000; MEB, 2013; Nasibov & Kaçar, 2005). Besides, geometry provides a rich content for developing mathematical reasoning, which includes deductive and inductive reasoning, making assumptions and evaluating the validity of these assumptions (NCTM, 2000). Therefore, geometry is a course that enables to occur actions inherent in proof. In Imamoğlu and Yontar-Togrol's (2015) study, both proof construction and evaluation practices are existing. However, particularly for proof evaluation, they focused on investigating whether a mathematical argument can be

accepted as a valid proof. It can be stated that the evaluation of proofs is an examination in general in the study. Accordingly, it is necessary to investigate in depth which criteria are taken into account during proof evaluation. Considering the pre-service mathematics teachers are responsible for instilling the fundamentals of mathematics to students after they have completed education (Lesseig, 2016), it is worth to examine their criteria of proof evaluation in the proving process. In consideration of this issue, we proposed to determine the criteria that pre-service mathematics teachers take into account when evaluating their peers' proofs.

3 Method

The data of this study were acquired from twenty activities in a geometry course carried out in an elementary mathematics teacher education program in a state university in Turkey. These activities were designed in alignment with the weekly topics in this course and during the course, two or three of these activities were implemented in every week. The details of the geometry course and the activities are below for this study.

3.1 The structure of the geometry course

The geometry is a course conducted in the spring semester of the first grade of elementary mathematics teaching program. The research participants included 50 first-year university students enrolled in this program. Each of these students is a pre-service teacher and they will educate middle school students in the future. The pre-service mathematics teachers have recently completed their high school education and that they are not likely to have any structured ideas about mathematical proof. In the fall semester, they took the abstract mathematics. Thus, they had enough knowledge of basic concepts about proof, proof methods, proving. In the scope of the geometry course, the activities about angles, congruence and similarity of triangles, angle bisector, altitude, parallel lines, orthocenter, centroid, incenter, excenter, circumcenter, polygons previously designed were implemented. The pre-service teachers gained experiences about proof construction and evaluation through the activities.

3.1.1 The content of the activities

In this study, twenty activities were conducted over the course of nine weeks to determine which criteria are considered by pre-service mathematics teachers in the process of proof evaluation. The activities were developed with consideration for the opinions of experts in the field. The activities included questions related to geometric proof on topics such as angles, congruent and similar triangles, polygons, height, angle bisector, median, and circles. In the activities, the finished proofs were not directly presented to the pre-service teachers. They were asked to construct the proofs step-by-step. In this paper, the stages of activities were based on the instructional model designed by Öztürk (2016). Apart from the final stage of this instructional model, an activity was consisted of six stages as “understanding the problem, constructing a structure, working on the structure and conjecturing, postulation of the relationship, proving, [and] evaluating the proof.” Each activity is implemented through group work. The implementation of an activity is as follows:

A geometry problem, relevant to unfamiliar mathematical propositions, theorems or mathematical relations, is given to examine its content in the stage of *understanding the problem*. A geometric structure based on the statements given in the problem is constructed with dynamic geometry software in the stage of *constructing a structure*. In the stage of *working on the structure and conjecturing*, the measurements are performed on the structure, and then the relationships are searched. In the stage of *postulation of the relationship*, the hypothesis and conclusion are determined, and the proposition is stated. In the stage of *proving*, each of group members thinks how to prove individually and his/her opinions about proving are shared with the group. Then, the group designs the common proof plan, and writes the steps of the proof. It takes approximately 25 minutes. In the stage of *evaluating the proof*, the proof plans are switched between the groups, and then they express each other's errors and deficiencies in the proofs. The groups spend about 15 minutes in this stage. At the end of each activity, the designed proofs and their evaluations were criticized. During the whole-class discussion, the errors about their steps of proofs and proof evaluations were deliberated. The aim of these discussions was to enable the students to recognize other errors and deficiencies that were ignored during the evaluation of proofs. After each discussion, the researchers summarized errors and deficiencies in the proofs. Thus, it was provided to experience the pre-service teachers to construct proofs and to evaluate them.

3.2 Data analysis

Firstly, the pre-service teachers' proofs were evaluated by the researchers in terms of the participants' justifications and the proper use of mathematical language regardless of the students' evaluations of proofs. The researchers' evaluations enabled to determine certain dimensions and their subcategories for proof evaluation. Then, these evaluations were compared with the students' evaluations of the proofs in terms of determined dimensions and their subcategories. The final version of subcategories was constructed on the researchers' and pre-service teachers' evaluations. In other words, the data about these proof evaluations were categorized. Therefore, content analysis, a method for making replicable and valid inferences from texts to the contexts of their use (Krippendorff, 2004), was carried out.

Proving requires some competencies such as reasoning, giving justifications, using mathematical language etc. In this sense, proof has a structure including justifications for each step (Weber, 2008), proper use of mathematical symbols and expressions (Mercer, Dawes, Wegerif & Sams 2004), following the deductive steps and writing these steps systematically (de Villiers, 1999). Apart from these, focusing on only the results was identified as a dimension by the evaluation of the participants' proofs. The dimension as "Focusing on only the results" means that a proof is evaluated with an opinion as "If a proposition is expressed correctly in an activity, the proof of this proposition is constructed correctly." Along with main categories about evaluating the proofs, it is also determined the subcategories for the dimensions of justification and mathematical language.

At the result of analyzing their proof evaluations, the frequencies of these criteria were determined. The frequency of certain category represents the total number of different proof evaluations of pre-service teachers or researchers. In other words, it is accepted as a frequency that different evaluations about errors or deficiencies in a proof belong to the same category. As evaluations in the process of discussions had for purpose to enable pre-service teachers to recognize all of the errors or deficiencies in the proofs, frequencies about the categories of proof evaluation were determined on the basis of their evaluations about peers' proofs.

For coding reliability, a researcher being experienced in mathematics education evaluated 30% of the pre-service teachers' proofs. The coherence between the researchers' coding was determined as 87%.

4 Results

An examination of the proofs constructed by the pre-service teachers revealed that 58% considered justification as the criteria for proof evaluation while 35,9% considered the proper is of mathematical language as the important criteria. The criteria of proof evaluation about justification were divided into subcategories. The subcategories relating to justifications were identified as *Including explanations for mathematical statements (J1)*, *Expressing stages between the steps of proof (J2)*, *Qualifying as a proof (J3)* and *Using appropriate definitions, axioms or theorems in the steps of the proof (J4)*. J1 is related to giving justifications for mathematical statements used in each step of proof. In other words, it is about expressing how to write a step in a proof. J2 is about accounting for the relationship between the steps of a proof. Besides, this subcategory requires not occur any gaps from a step to another step in a proof. J3 illustrates that all the statements written for proving a mathematical relation is accepted as correct steps of a proof or a complete proof. J4 is about selecting appropriate mathematical definitions, axioms or theorems for proving a mathematical relation. With respect to mathematical language, several subcategories were identified, including *Using symbolic language appropriately (L1)*, *Using non-symbolic language appropriately (L2)*, and *Proof writing in a systematic way (L3)*. L1 is related to using appropriate mathematical symbols in the proof process. L2 is about using appropriate mathematical concepts and statements in the proof process. L3 is related to writing the steps of a proof regularly in an appropriate order. In addition, a separate category, *Focusing on only the results*, was defined. **Table 1** summarizes the distribution of frequencies of the pre-service mathematics teachers' criteria for proof evaluation and the researchers' proof evaluations.

Table 1. The frequencies of the criteria for proof evaluation

The criteria for proof evaluation	The pre-service teachers' evaluations	The researchers' evaluations
	f	f
Justifications		
J1. Including explanations for mathematical statements	81	220
J2. Expressing stages between the steps of proof	4	22
J3. Qualifying as a proof	25	38
J4. Using appropriate definitions, axioms or theorems in the steps of the proof	98	61
Mathematical language		
L1. Using symbolic language appropriately	96	159
L2. Using non-symbolic language appropriately	9	15
L3. Proof writing in a systematic way	24	9
Focusing on only the results	22	-

In [Table 1](#), the frequencies of the researchers' evaluations represent the number of evaluations that must be carried out in all the pre-service teachers' proofs. As [Table 1](#) demonstrates, J4, as a subcategory of justifications, had the highest frequency ($f = 98$); and in terms of frequencies, J1 ($f = 81$) followed J4. Furthermore, as a subcategory of mathematical language, L1 had the highest frequency ($f = 96$). On the other hand, J2 had the lowest frequency ($f = 4$) among the subcategories of the justifications, while L2 had the lowest frequency ($f = 9$) among the subcategories relating to mathematical language. When the frequencies of the pre-service teachers' evaluations are compared with the frequencies of the researchers' evaluations, they ignored some situations in the proofs. According to the comparison of frequencies, the pre-service teachers mostly ignored the situations about J1, as a subcategory of justifications, in the proofs. Situations about J2 is more ignored than situations about J3 criterion. However, the ignored situations about J2 and J3 have an approximate frequency. The pre-service teachers' evaluations relating to J4 have higher frequency than the researchers' evaluations relating to J4. The fact that the pre-service teachers mostly used general expressions as “*The steps of proof are reasonable. Because theorems were used appropriately.*” in their evaluations led to increase the frequency of J4 criterion. Similarly, stated as “*The sequence of the steps of proof is written correctly.*” by pre-service teachers contribute that the frequency of L3, as a subcategory of mathematical language, in their evaluations exceeds the frequency of researchers' evaluations about L3. Comparing with the researchers' evaluations in terms of mathematical language, there are also several situations that the pre-service teachers ignored in their

evaluations of proofs. These situations are related to L1 and L2. However, they ignored the situations about L1 more often than the situations about L2.

According to these frequencies, in the course of evaluating a mathematical proof, the pre-service teachers considered that a proof should include explanations for each mathematical statement, along with appropriate definitions, axioms or theorems. Furthermore, in terms of mathematical language, the participants emphasized the appropriate use of mathematical symbols. Furthermore, in addition to these criteria, some of the participants considered *Focusing on only the results* as a criterion for proof evaluation, wherein they evaluated a proof according to the correctness of a mathematical statement.

4.1 Pre-service mathematics teachers' criteria for proof evaluation relating to justifications

In the process of proof evaluation, the pre-service mathematics teachers primarily emphasized the criterion of *Using appropriate definitions, axioms, or theorems in the steps of the proof* (J4); this subcategory was preferred more than other criteria among the subcategories relating to justification. In addition, the criterion *Including explanations for mathematical statements* (J1) was given considerable importance. On the other hand, *Expressing stages between the steps of proof* (J2) and *Qualifying as a proof* (J3) were rarely preferred, with J2 receiving the least importance. This finding is not unexpected, as most of the pre-service mathematics teachers' proofs did not have many reasoning gaps between the steps. When the pre-service mathematics teachers' proof evaluations were compared with those of the researchers', it could be seen that not all of the evaluations were carried out appropriately. Namely, the participants did not fully evaluate the proofs with regard to justification; and even in cases where they took into account the suitability of definitions, axioms or theorems and the availability of explanations of mathematical statements, their evaluations were deficient with respect to these criteria. Moreover, their proof evaluations in relation to other criteria for justification were incomplete. One example from each of criteria identified by the pre-service teachers is presented below.

4.1.1 Results relating to the Criterion Including Explanations for Mathematical Statements (J1)

It is important that a proof includes adequate explanations for each mathematical operation or statement in order to clarify how a mathematical relationship has been determined. As such, J1 is one of the important criteria for proof evaluation. However, while the pre-service mathematics teachers considered the criterion of J1 in their evaluations, their results were not always in accord with the researchers' evaluations, in that they did not remark on certain issues relating to the criterion of J1. For instance, in the process of proof construction, certain questions should be answered, such as *What do the variables in a proof mean? How is the mathematical result reached? and Which axioms, definitions or theorems are used?* However, the participants did not always seek the answers to these questions while evaluating their peers' proofs. **Figure 1** presents an example of an evaluation that ignored the criterion of J1.

	<p>[AK and [BL are respectively exterior angle bisectors of the angles of A and B in ABC triangle. [CH] ⊥ [AK], [CT] ⊥ [BL]</p> <p>Find whether there is a relation between AB , AC , BC and HT by following the steps below.</p>
<p> $(AC) + (BC) + (AB) = 2(HT)$ $2b + 2a + 2c = 2(a+b+c)$ $[HT] \parallel [AB] \checkmark$ $\triangle HBA \sim \triangle CHA$ $\triangle TXB \sim \triangle CTB \checkmark$ $CB = CT = CL$ $CA = CH = CK$ $\triangle CBE$ ve $\triangle CAD$ ikizkenar üçgen </p> <p> $m(\angle TB) = m(\angle TBX)$ (z. Ters aç.) $m(\angle HA) = m(\angle HAY)$ (z. Ters aç.) $\triangle CCL \sim \triangle CAB \Rightarrow \frac{ CL }{ CA } = \frac{ CL }{ CB } = \frac{ CL }{ AB } = \frac{1}{2}$ </p>	<p>The researchers' evaluations</p> <ul style="list-style-type: none"> • It was expressed as $HYA \sim CHA, TXB \sim CTB, CKL \sim CAB$. However, the participant did not reference the similarity theorem on which the statement was based (J1). • It was noted as $[HT] \parallel [AB]$, but how this parallelism was determined was not stated (J1). • It was not stated why CBE and CAD triangles are isosceles (J1). <p>The pre-service teacher's evaluations</p> <ul style="list-style-type: none"> • They indicated the parallelism in an alternative way. The methods used are sufficient and correct (J4). • The result that they reached is correct (Focusing on only the results).
<p><i>farklı bir yöntemle paralelliği göstermektedir. Kullandıkları yöntemler ispat yeterli ve doğrudur. İspatla ulaştıkları sonuca abgöndür.</i></p>	

Figure 1. An example of an evaluation that ignored the criterion of J1.

4.1.2 Results relating to the criterion expressing stages between the steps of proof (J2)

When proving a mathematical theorem, it is necessary to write all of the steps that show how to reach the proof. In this manner, a proof becomes more descriptive, and a connection is made between the steps. With this in mind, J2 is an essential criterion for proof evaluation. In the proofs performed by the pre-service mathematics teachers, instances where stages between the steps of a proof were not expressed occurred less often than other errors; therefore, references to problems concerning J2 were less frequent in the evaluations. However, when compared with the researchers' evaluations, it can be seen that less than half of the existing issues were noted in the participants' evaluations. For example, in constructing their proofs, some of the pre-service teachers performed mathematical operations (finding the measures of angles or lengths of sides, drawing the lines or line segments etc.) on geometric figures or described variables in the problems, and then wrote the existing mathematical relationships directly. During the evaluation of these proofs, their peers generally realized that there were gaps in the steps; accordingly, their evaluations reflected the J2 criterion, as with the example below (see Figure 2).

	<p>ABC is a triangle. D and E are respectively on the sides of [BC] and [AC]. [AF] is the bisector of \widehat{CAD}. [BF] is the bisector of \widehat{CBE}.</p> <p>Find whether there is a relation between $m(\widehat{AEB})$, $m(\widehat{ADB})$ and $m(\widehat{AFB})$ by following the steps below.</p>
	<p>The researchers' evaluations</p> <ul style="list-style-type: none"> • It was expressed as $m(\widehat{AEB}) = \alpha + 2x$, $m(\widehat{BDA}) = \alpha + 2\beta$, $m(\widehat{AGB}) = y + \alpha$, $m(\widehat{AFB}) = y + x + \alpha$. However, the participant did not state how to reach these statements (J1). • The mathematical relationship [$2m(\widehat{AFB}) = m(\widehat{AEB}) + m(\widehat{BDA})$] was written directly. Instead, the process for reaching this relationship should be given step by step (J2). <p>The pre-service teacher's evaluations</p> <ul style="list-style-type: none"> • The mathematical relationship should be shown by expressing the operations as mentioned above (J2).

Figure 2. An example of an evaluation that reflects the criterion of J2.

When mathematical operations were missing in the proofs, the participants had difficulties in recognizing the gaps between the steps; thus, their evaluations ignored some aspects of the criterion of J2. An example of this is as follows (see Figure 3):

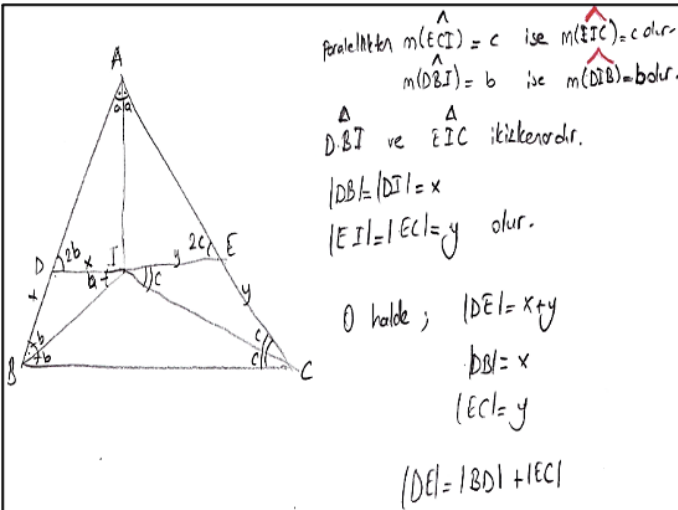
<p>ABC is a triangle. The point I is cut-off point of the internal angle bisector. A line being parallel to [BC] and passing through the point I is respectively cutting [AB] and [AC] at the points of D and E.</p> <p>Find whether there is a relation between DB , EC and DE by following the steps below.</p>	
 <p>paraleliten $m(\widehat{ECI}) = c$ ise $m(\widehat{ETC}) = c$ olur $m(\widehat{DBI}) = b$ ise $m(\widehat{DIB}) = b$ olur. ΔDBI ve ΔIEC ikizkenardır. $DB = DI = x$ $EI = EC = y$ olur. O halde ; $DE = x + y$ $DB = x$ $EC = y$ $DE = DB + EC$</p>	<p>The researchers' evaluations</p> <ul style="list-style-type: none"> It was expressed that because of parallelism if they are $m(\widehat{ECI}) = c$, $m(\widehat{DBI}) = b$; they are $m(\widehat{ETC}) = c$, $m(\widehat{DIB}) = b$. It can be written in a more descriptive way that because of parallelism, they are $m(\widehat{ICB}) = c$, $m(\widehat{ECI}) = c$ and so $m(\widehat{ETC}) = c$, $m(\widehat{ETC}) = c$. Also, because of parallelism they are $m(\widehat{DIB}) = m(\widehat{IBC}) = b$ and so $m(\widehat{DIB}) = m(\widehat{DBI}) = b$ (J2). Instead of $m(\widehat{ECI}) = c$, $m(\widehat{DIB}) = b$, it can be written as $m(\widehat{ETC}) = c$, $m(\widehat{DIB}) = b$ (L1).
<p>ispat doğru yapılmıştır. Matematiksel mutasyonlardan biri hariç, L1, m ve t teorim dahilindedir.</p>	
<p>The pre-service teacher's evaluations</p> <ul style="list-style-type: none"> The proof was correct. Apart from one of mathematical notation, definitions and theorems were correct (J4). It can be written as $m(\widehat{ETC}) = c$, $m(\widehat{DIB}) = b$ (L1). 	

Figure 3. An example of an evaluation that ignored the criterion of J2.

4.1.3 Results relating to the criterion qualifying as a proof (J3)

Efforts such as checking the validity of a mathematical theorem, examining special cases, or testing numerical values cannot be accepted as a proof; nor can stating the required steps for proving without including the related mathematical operations. Rather, the criterion of *Qualifying as a proof* depends on including the required steps for proving, along with their justifications. Moreover, it is important to examine whether a proof is valid for all cases, or only in special instances. As such, J3 is considered one of the important criteria for proof evaluation. Although certain issues relating to J3 were ignored in the participants' evaluations, the researchers' evaluations in this respect were sometimes in accord with those of the pre-service mathematics teachers. In some cases, they described the required variables for a proof; furthermore, they preferred writing a proof in the form of checking the validity of the mathematical theorem. In their evaluations of these proofs, they ignored some of the required evaluations about this more than other evaluations relating to J3. One

such example is given below (see Figure 4).

	<p>[AK and [BL are respectively exterior angle bisectors of the angles of A and B in ABC triangle. [CH] ⊥ [AK], [CT] ⊥ [BL]. Find whether there is a relation between AB , AC , BC and HT by following the steps below.</p>	<p>The researchers' evaluations</p>
<p> $\frac{HK}{ RA } = \frac{\frac{y}{2}}{y} = \frac{1}{2}$ oranını bulduk $\frac{LT}{ BS } = \frac{\frac{c}{2}}{c} = \frac{1}{2}$ olur. $\frac{y}{2} + k + \frac{c}{2} = \frac{y+c+2k}{2}$ $\frac{CA}{CK} = \frac{CB}{CK} = \frac{AB}{KL}$ (A.AA Benzer. Teoremi) </p>	<ul style="list-style-type: none"> • Instead of $\frac{HK}{ RA } = \frac{LT}{ BS } = \frac{CA}{CK} = \frac{CB}{CK} = \frac{AB}{KL}$, it should be written as $\frac{ HK }{ RA } = \frac{ LT }{ BS } = \frac{ CA }{ CK } = \frac{ CB }{ CK } = \frac{ AB }{ KL }$ (L1). • It was not explained why ACR and CBS triangles are isosceles (J1). At the end of proof, it was written as $\frac{y}{2} + k + \frac{c}{2} = \frac{y+c+2k}{2}$. Thus, the mathematical relationship that should be proven was not established (J3). 	
	<p>The pre-service teacher's evaluations</p>	
	<ul style="list-style-type: none"> • Instead of CBS triangle, it should be written as CBS (L1). • Instead of $\frac{HK}{ RA } = \frac{LT}{ BS } = \frac{CA}{CK} = \frac{CB}{CK} = \frac{AB}{KL}$, it should be written as $\frac{ HK }{ RA } = \frac{ LT }{ BS } = \frac{ CA }{ CK } = \frac{ CB }{ CK } = \frac{ AB }{ KL }$ (L1). 	

Figure 4. An example of an evaluation that ignored the criterion of J3.

Furthermore, most of the pre-service mathematics teachers stated that the proofs they evaluated examined only special cases or relied on trying numerical values; and thus, they did not qualify as a proof. In other words, they realized that such instances cannot be accepted as a proof, in line with the J3 criterion for proof evaluation. In addition, the proofs that contained the required steps of proof, but without expressing the mathematical operations, or those that stated only the mathematical relationships, were generally evaluated with reference to J3. Thus, the participants were able to explain that these instances did not qualify as a proof. One example of this is as follows (see Figure 5):

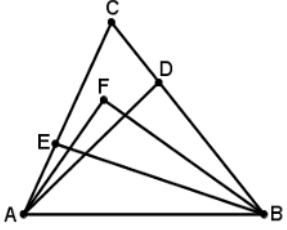
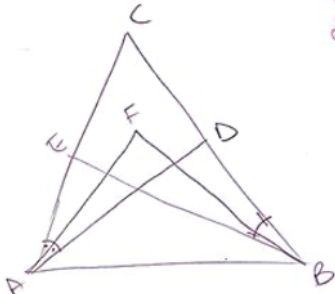
	<p>ABC is a triangle. D and E are respectively on the sides of [BC] and [AC]. [AF] is the bisector of \widehat{CAD}. [BF] is the bisector of \widehat{CBE}.</p> <p>Find whether there is a relation between $m(\widehat{AEB})$, $m(\widehat{ADB})$ and $m(\widehat{AFB})$ by following the steps below.</p>
<p>$m(\widehat{AEB}) + m(\widehat{ADB}) = m(\widehat{AFB})$</p> <p>2</p>  <p>Sadece iftihar yapılmıy Buna nasıl ulaşılmıy? on anlatılmıy</p> <p>* Bağntının geometri'ye yap. Üzerinde zenc-10 olması yeterlidir. Çünkü selul üzerinde bağntıları gösterilebiliriz. Bağntı iftihar için bu yeterlidir.</p>	<p>The researchers' evaluations</p> <ul style="list-style-type: none"> • Only examining special cases was accepted as a proof. This was not a proof (J3). <p>The pre-service teacher's evaluations</p> <ul style="list-style-type: none"> • Only mathematical relationship was provided. An explanation should be given on how to reach this relation (J3). <p>* It is sufficient to state that the mathematical relationship is valid on geometric construct, because we can show the relationship.</p>

Figure 5. An example of an evaluation that reflects the criterion of J3.

4.1.4 Results relating to the criterion using appropriate definitions, axioms, or theorems in the steps of the proof (J4)

The process of proving is based on the use of appropriate definitions, axioms, or theorems. In order to construct a qualified proof, the mathematical operations for each step should be provided, and they should be supported by appropriate definitions, axioms or theorems. Therefore, in proof evaluation, it is necessary to examine the suitability of the definitions, axioms, or theorems according to the criterion of J4. When compared with the researchers' evaluations, the pre-service mathematics teachers' evaluations referenced the criterion of J4 more frequently. The reason for this was that they mainly evaluated the proofs with respect to the best mathematical statement for proving, or on the basis of similarity to their own proofs. Namely, they generally gave attention to whether the mathematical statements in their proofs were the same as those in the proofs developed by their classmates. On the other hand, they ignored other aspects of evaluation relating to J4. An example of these evaluations is below (see [Figure 6](#)).

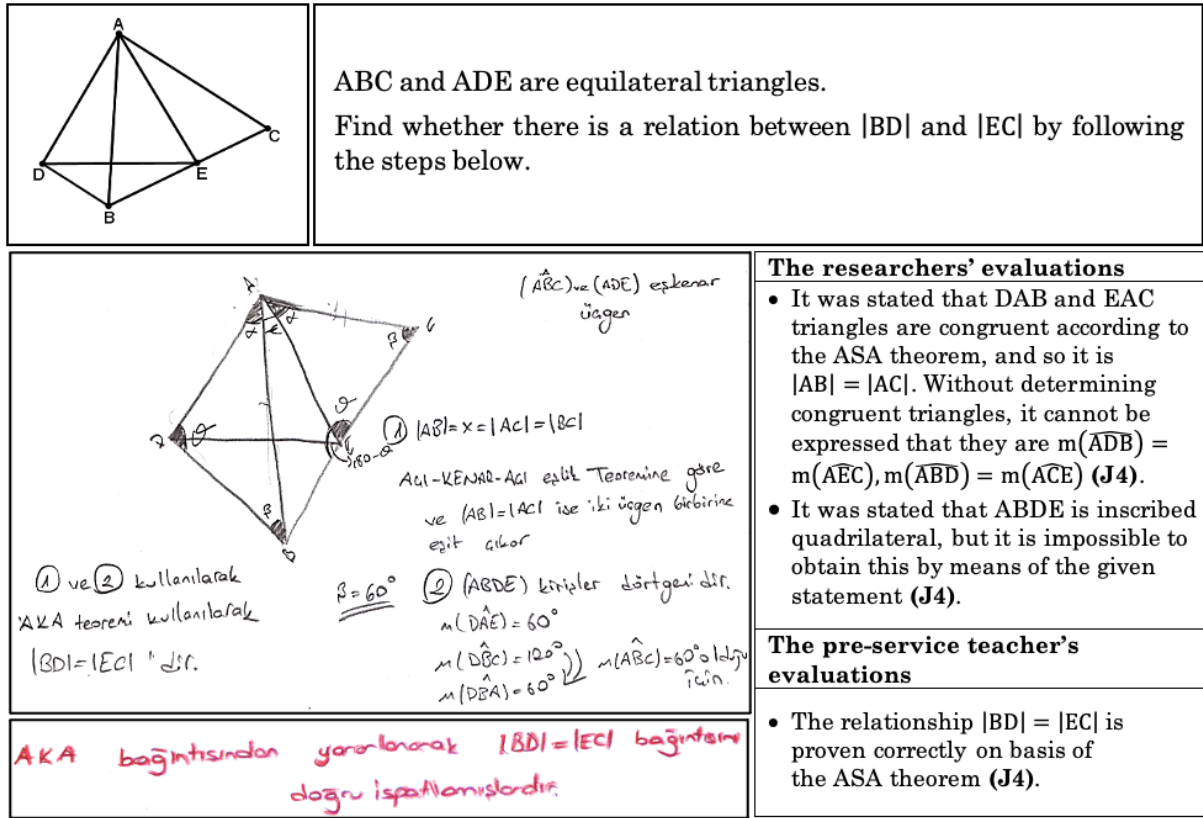


Figure 6. An example of an evaluation that related to the criterion of J4.

Another aspect of students' evaluations related to the use of inappropriate definitions, axioms, or theorems; however, these were limited to the evaluations wherein this was necessary. An example of the students' determination of the use of inappropriate definitions, axioms or theorems is provided below (see Figure 7).

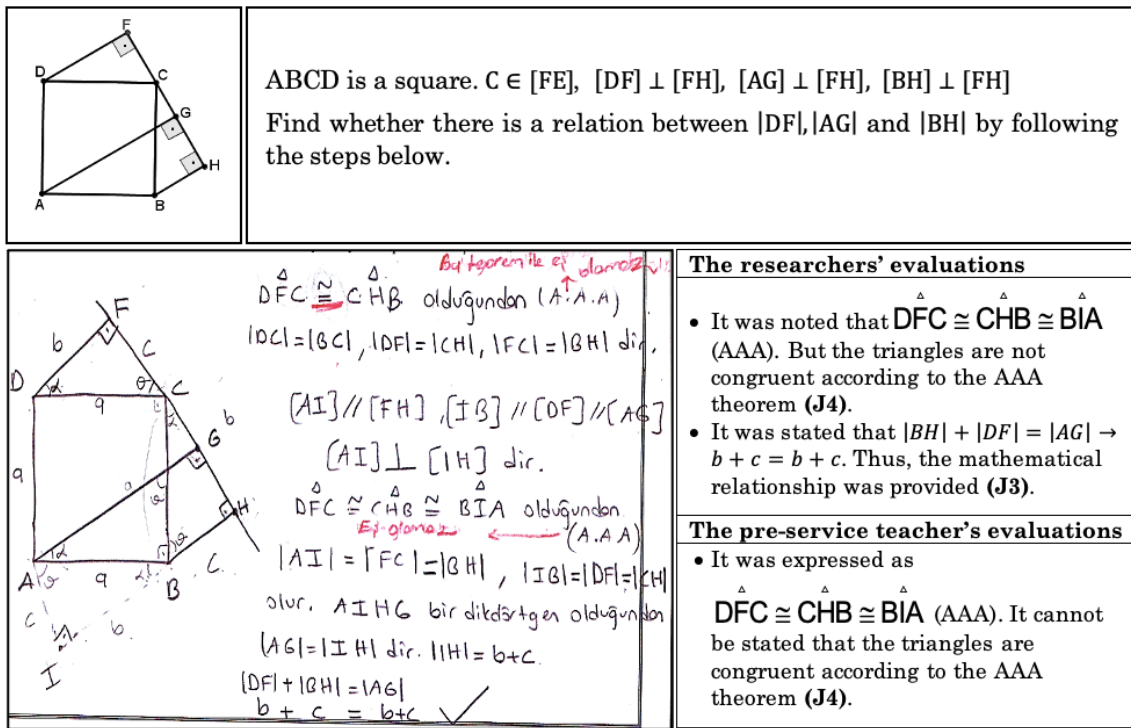


Figure 7. An example of an evaluation that reflects the criterion of J4.

In addition, the students sometimes regarded insignificant details about J4, such as the expression of *similar* versus *congruent* triangles. In other words, they did not realize that congruent triangles are also similar triangles, as demonstrated in the example below (see Figure 8):

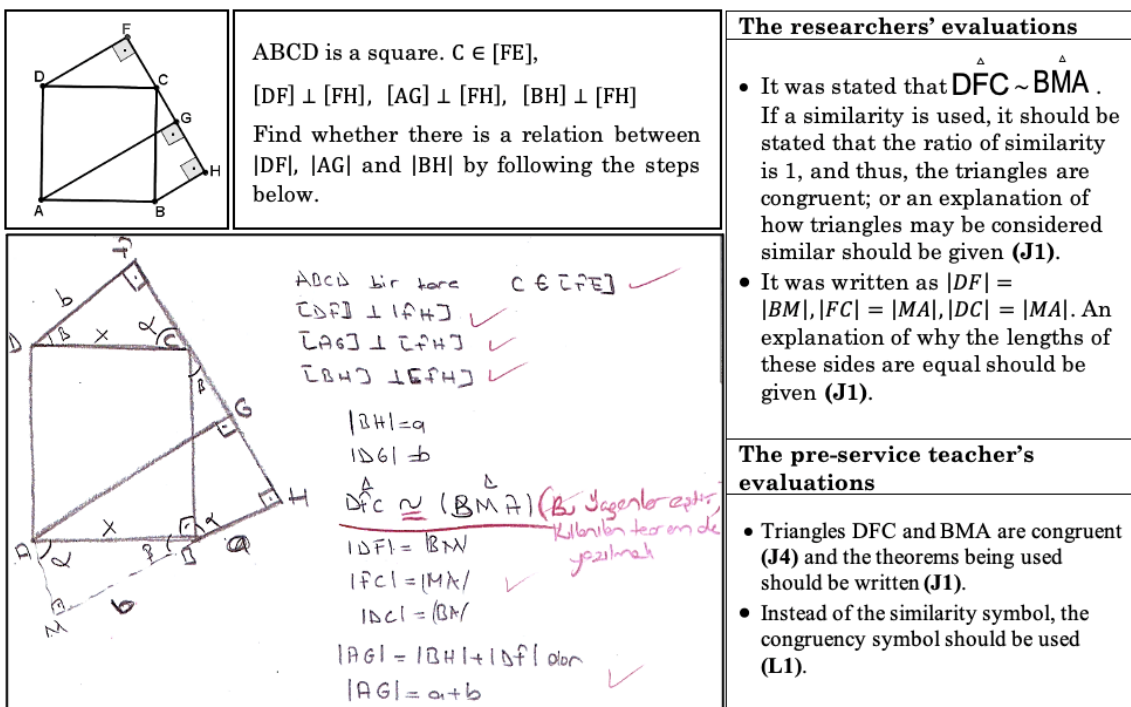


Figure 8. An example of an evaluation related to the criterion of J4.

4.2 Pre-service mathematics teachers' criteria for proof evaluation in relation to mathematical language

In their evaluations, the pre-service mathematics teachers considered the criterion of *Using symbolic language appropriately* (L1) more frequently than other criteria in the subcategories of mathematical language. However, many of the instances in the proofs relating to L1 were not noted in the students' evaluations, as the comparison with the researchers' proof evaluations demonstrates. On the other hand, they considered L2, *Using non-symbolic language appropriately*, less frequently than the other criteria. This was due in part to the fact that this situation was encountered less often in their proofs; therefore, this result is expected. Among the other criteria in the subcategories of mathematical language, *Proof writing in a systematic way* (L3) occurred at an average frequency in the proof evaluations. All of the criteria for proof evaluation concerning mathematical language are presented in the following sections.

4.2.1 Results relating to the criterion using symbolic language appropriately (L1)

The appropriate use of mathematical symbols facilitates the comprehensibility of a mathematical proof, including the steps of the proof and the connections between them. As such, the criterion of L1 is considered as one of the subcategories of mathematical language for proof evaluation. In this study, the pre-service mathematics teachers took this criterion into account more often than other criteria among the subcategories of mathematical language. However, in comparison with the researchers' proof evaluations, most of the pre-service mathematics teachers ignored some of the errors or deficiencies in this respect, or they accepted some mathematical symbols as correct use of notations. An example of this is provided here (see [Figure 9](#)):

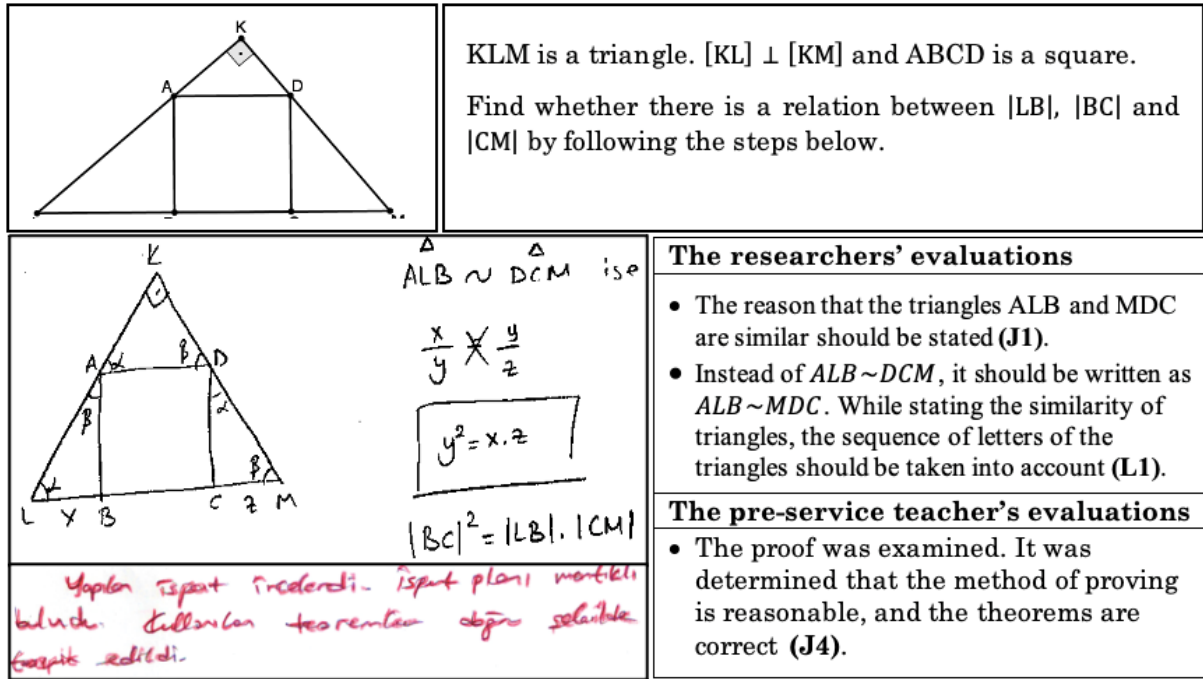


Figure 9. An example of an evaluation that ignored the criterion of L1.

The pre-service teachers also tended to realize errors concerning L1 for certain mathematical symbols, i.e., angle and length. However, they were seen to ignore some required evaluations in this respect. An example of this issue is given below (see Figure 10).

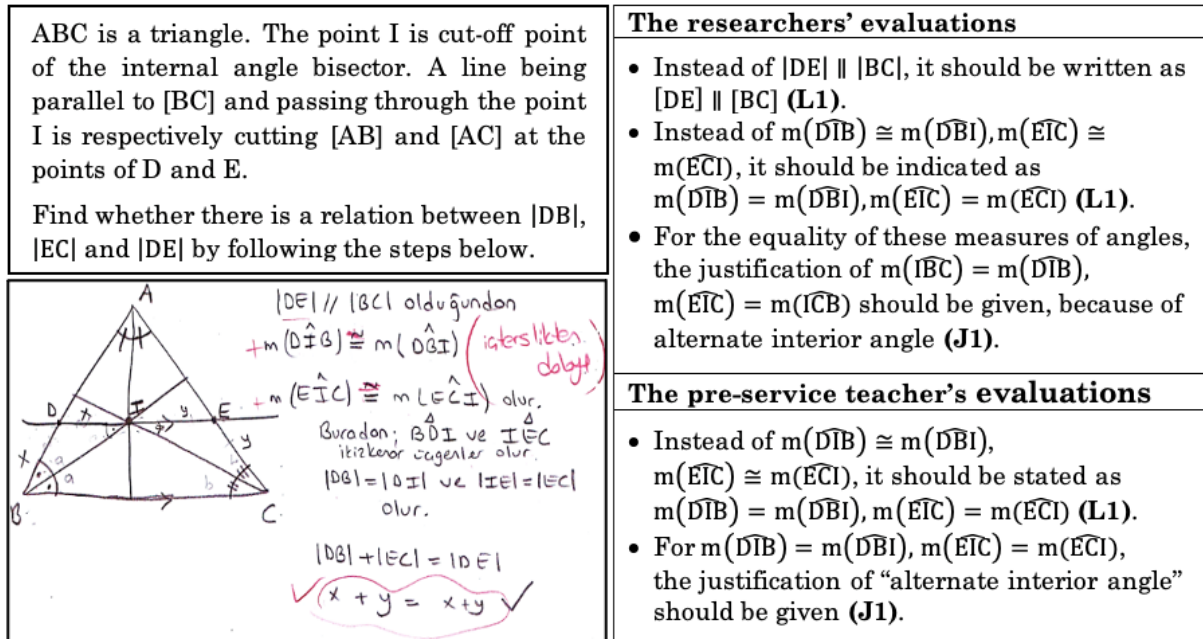


Figure 10. An example of an evaluation related to the criterion of L1.

4.2.2 Results relating to the criterion using non-symbolic language appropriately (L2)

When justifications are presented for the steps of a proof, mathematical concepts or explanations should be expressed using appropriate mathematical language; and in order to construct a quality proof, these concepts and explanations should be comprehensible. Accordingly, it is necessary that non-symbolic language be used appropriately throughout the steps of a proof; and therefore, L2 constitutes one of the criteria for proof evaluation. In the current study, while the pre-service mathematics teachers were evaluating their peers' proofs, they generally remarked on discrepancies between mathematical symbols and statements. For example, when the congruence symbol was used, but the similarity theorem was referenced in the proof, the students who evaluated the proof noted that the congruence theorem should have been applied. An example of this instance is below (see [Figure 11](#)).

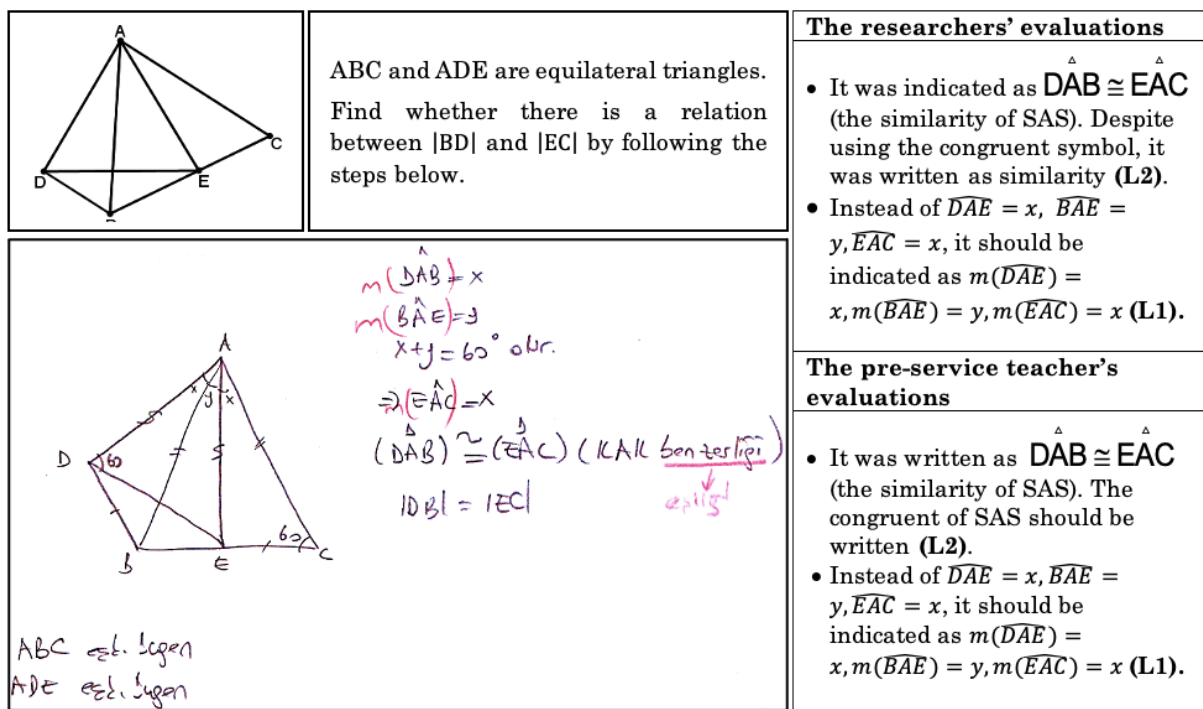


Figure 11. An example of an evaluation that reflects the criterion of L2.

On the other hand, unlike the researchers' evaluations, they ignored some issues relating to the criterion of L2. Although fewer errors relating to L2 were encountered in their proofs, there were some situations where the pre-service teachers did not recognize an error, or they sometimes made errors in their evaluations. For example, one of their evaluations noted that a statement should have been written as "ABC

isosceles triangle” instead of \hat{ABC} isosceles. During the construction of a proof, the students sometimes did not state mathematical definitions, axioms, or theorems appropriately; and in some of these situations, the evaluations did not reflect the L2 criterion, as demonstrated below (see Figure 12).

	<p>ABC is a triangle. D and E are respectively on the sides of [BC] and [AC]. [AF] is the bisector of \widehat{CAD}. [BF] is the bisector of \widehat{CBE}.</p> <p>Find whether there is a relation between $m(\widehat{AEB})$, $m(\widehat{ADB})$ and $m(\widehat{AFB})$ by following the steps below.</p>
	<p>The researchers' evaluations</p> <ul style="list-style-type: none"> • The same measure of the angle has two different variables: x and y. The measure of angle AFB is y (L1). • Instead of $AEB = y, AEB = x, ADB = z$, it should be indicated as $m(\widehat{AEB}) = y, m(\widehat{AFB}) = x, m(\widehat{ADB}) = z$ (L1). • The justification given was “because two interior angles are equal to an exterior angle. Instead, it should be stated that the measure of an exterior angle of a triangle is equal to the sum of the measure of its two non-adjacent interior angles (L2). <p>The pre-service teacher's evaluations</p> <ul style="list-style-type: none"> • It was expressed as $AEB = y$. However, the measure of angle AFB should have been expressed (L1). • Instead of $AEB = y, AEB = x, ADB = z$, it should be indicated as $m(\widehat{AEB}) = y, m(\widehat{AFB}) = x, m(\widehat{ADB}) = z$ (L1).

Figure 12. An example of an evaluation that ignored the criterion of L2.

4.2.3 Results relating to the criterion proof writing in a systematic way (L3)

By writing the successive stages of a proof, it is possible to understand which mathematical operations have been performed and which steps have been followed. In this manner, the connection between the steps of a proof can be easily seen, and those steps that are necessary or unnecessary can be distinguished. Through this process, a proof becomes comprehensible. With this in mind, it is important to write a proof in a systematic way; therefore, L3 can be considered as a criterion for proof evaluation. Accordingly, during their evaluations of their peers' proofs, the pre-service

mathematics teachers generally considered whether the appropriate steps had been followed. When the steps of their peers' proofs were similar to their own, they indicated in their evaluations that their peers had followed the appropriate steps; furthermore, most of their notations relating to L3 concerned the suitability of the steps of the proofs. However, in comparison with the researchers' evaluations, they ignored situations in which the mathematical operations were vague. Apart from that, there were certain instances in which they did not remark on unnecessary steps. An example of this is given below (see Figure 13).

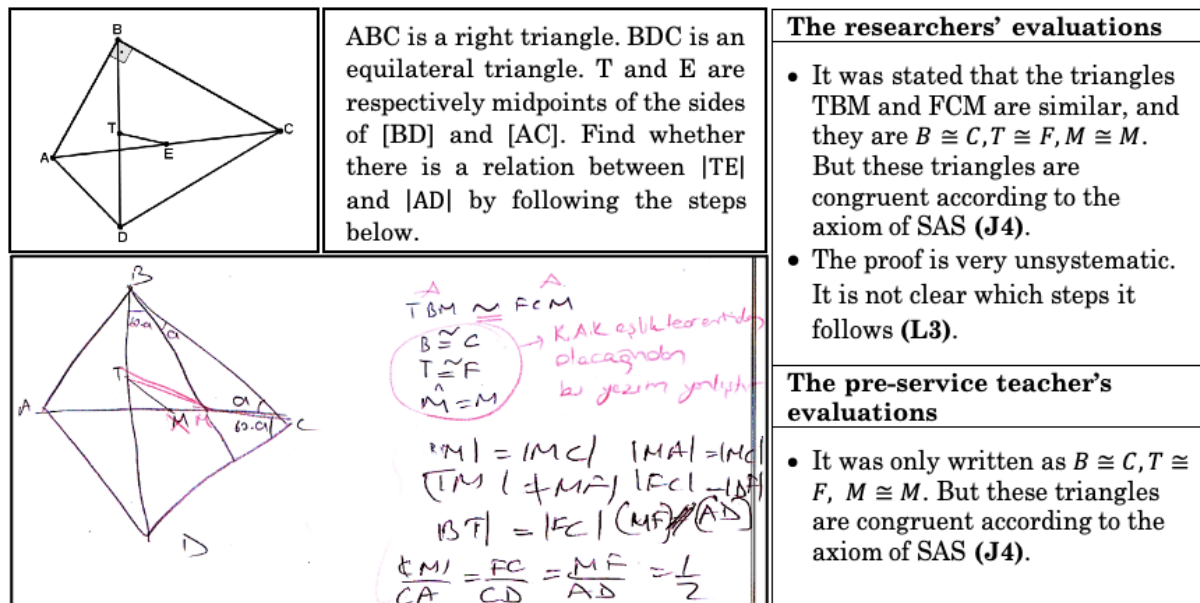


Figure 13. An example of an evaluation related to the criterion of L3.

4.3 Pre-service mathematics teachers' criteria for proof evaluation in relation to focusing on only the results

Aside from justifications and mathematical language, the pre-service mathematics teachers also regarded a criterion identified as *Focusing on only the results*. Namely, while they were evaluating the proofs, they sometimes considered them only in terms of focusing on the results; and in cases where the results of a peer's proof were the same as their own; they determined it to be valid. One such instance of this is illustrated below (see Figure 14).

	<p>ABCD is a parallelogram. [AN], [BT], [CL] and [DK] are respectively bisector of the angles of the parallelogram. Find whether there is a relation between TE and AD by following the steps below.</p>	<p>The researchers' evaluations</p>
	<p>[AB] // [DC] olduğundan $m(\hat{C}) + m(\hat{A}) = 180^\circ$ (T.A.S.) \rightarrow birbirine Düşünelim $EFGH \rightarrow$ dikdörtgen olur. $m(\hat{D}EK) = b^\circ$ $AC = x$ $TM = y$ $m(\hat{D}NA) = b^\circ$ $NC = x - y$</p> <p>[DC] // [MS] old. $m(\hat{E}HB) = b^\circ$ $n(\hat{E}GH) = a^\circ$ olur. Yani [ENCG] paralelkenar oldu $NC = x - y$ ise $\Rightarrow EG = x - y$ dir. $EG = HF = x - y$</p> <p><i>in celendi. Detaylı yazılmış. ve sonuç doğru.</i></p>	<ul style="list-style-type: none"> • Instead of $m(\hat{EHG}) = b$, it should be written as $m(\hat{HEG}) = b$. The name of angles was stated incorrectly (L1). • It was indicated as "because of [ENCG] parallelogram," but a parallelogram cannot be expressed in this way (L1).
		<p>The pre-service teacher's evaluations</p>
		<ul style="list-style-type: none"> • It was examined, it was written in detail (L3), and the result is correct (Focusing on only the results).

Figure 14. An example of an evaluation related to the criterion of focusing on only the results.

5 Discussion and Conclusion

At the end of the series of activities relating to proof construction and evaluation in this study, there were cases that the pre-service mathematics teachers determine or ignore errors or deficiencies in the proofs. However, it can be stated that they determined the errors and deficiencies of them better in the process of proof evaluation through observation in class discussions. Some criteria about proof evaluation had more high frequency at the end of all the activities. For example, as subcategories of the justification, these criteria were J1, J4; as a subcategory of mathematical language, it was L1. In other words, throughout the process of both constructing proofs and evaluating them, the frequencies of stating some criteria for proof evaluation changed, as they became better at elaborating their criteria for evaluation. A similar situation was noted by Powers et al. (2010), who argued that using specific activities for proof validation may contribute to the development of proof writing and validation abilities. Considering that proof evaluation is more comprehensive than proof validation, their contention supports the results of the current study.

In this study, three main categories for proof evaluation emerged as justification, mathematical language, and focusing on only the results. Furthermore, the fact that the subcategories related to justification and mathematical language were determined presents a framework for proof evaluation. Although focusing on only the results is one of main categories for proof evaluation, it is only a criterion that pre-service

teachers take into account in evaluating a proof. In the studies on proving (e.g., Lee, 2011; Pulley, 2010; Senk, 1983), proofs were generally evaluated in terms of reasoning, mathematical language, justification that were not identified as a separate category. The fact that justification and mathematical language were among the categories for proof evaluation in current study coincides with considering these dimensions for proof evaluation in these studies. However, this study differs from the others by identifying certain dimensions and subcategories for proof evaluation.

Moreover, the criteria for proof evaluation in this study took form in line of reasoning errors during the proof process. For example, J1 and J2 as subcategories of justification is related to holes, which occurs as result of claiming that a statement follows immediately from previously established results when in reality a considerable argument is required, defined by Selden and Selden (2003b). In addition, Demir (2017) classified reasoning errors into three categories. Reasoning gaps is one of the categories. A part of subcategories of reasoning gaps is about not giving justifications and omitting some steps of proof. Therefore, J1 and J2 are directly associated with these subcategories.

In the process of proof evaluation, the pre-service teachers generally expected that the steps of a proof should be related to one another, and that a proof should include appropriate mathematical statements (axioms, theorems etc.). These expectations are related to J1 and J4 as subcategories of justification. A deductive approach is originated from justification (de Villiers, 1990; Hanna, 1990). Moreover, deductive logic is well-exemplified by proof (Remillard, 2009). As a result, some criteria relating to justification (J1, J4) are likely to be considered more by pre-service teachers while evaluating proofs. On the other hand, proving can be characterized as problem solving task (Weber, 2005). Completing a problem solving task requires to perform appropriate mathematical actions and operations. Similarly, proving is a mathematical task including actions as using some initial information (e.g., assumptions, axioms, definitions) and applying rules of inferences (Anderson, 2000). In both problem solving and proving, these actions play an explanatory role. Thus, it is probable that during proof evaluations, the pre-service teachers paid more attention to the explanations for mathematical statements.

Furthermore, in order to accept successive mathematical steps as a proof, it is necessary that these steps include explanations about how to reach a mathematical relationship; and in this process, appropriate mathematical symbols and concepts should be applied. Therefore, when evaluating a proof, justifications and

mathematical language are important considerations. In this respect, the pre-service teachers did consider some of the required criteria, but they ignored others. With the instances where issues relating to justification and mathematical language were ignored, there were generally no obvious errors or deficiencies. Therefore, it may be concluded that the pre-service teachers were more conscious of obvious errors. Imamoglu and Yontar-Togrol (2012) likewise indicated that pre-service teachers have difficulties in evaluating a proof when there are no obvious errors in an argument. In contrast with the researchers' evaluations, this situation was valid for all of the criteria for proof evaluation relating to justification and mathematical language. Namely, although there were fewer differences between the evaluations of the researchers and the pre-service teachers relating to J3 than to other criteria, the pre-service teachers tended on the whole to ignore the gaps between the steps of the proofs that they did not recognize directly.

Moreover, the researchers' evaluations in relation to the criteria of J1 and L1 were more frequent than those of the pre-service teachers; thus, it was concluded that they mainly ignored these criteria. On the other hand, the criteria of J4 and L3 were featured more prominently in the pre-service teachers' evaluations than in the researchers' evaluations. In this sense, their evaluations dealt with J4 in terms of whether a peer's proof resembled their own in terms of method, theorems, and so on; and with respect to L3, they mainly considered whether the steps of a proof were neatly written. As such, it can be inferred that their evaluations were sometimes superficial, and when their peers' proofs resembled their own, they assumed them to be correct. This circumstance relates to the work of Selden and Selden (2003a), who also found that undergraduate students generally focused on the surface features of arguments, rather than their underlying logical structure.

In terms of mathematical language, using symbolic language appropriately was a criterion that the pre-service teachers primarily considered when evaluating proofs. As such, it can be asserted that they paid more attention mathematical symbols than non-symbolic statements during proof evaluations. Hill, Ball and Schilling (2008) stated that teachers can be unaware of their students' verbal justifications. Moreover, Stylianides, Stylianides and Philippou (2004) pointed out that it is plausible to assume that symbolic reasoning is a focal point in the collegiate mathematics curriculum. The more emphasis on symbolic mathematical statements during their undergraduate education can lead them to disregard verbal justifications. Therefore,

the pre-service teachers may also tend to evaluate the proofs in terms of symbolic language.

On the other hand, when evaluating their peers' proofs, the participants generally did not accept specific examples as proof, as with Imamoglu and Yontar-Togrul (2012), who pointed out that most pre-service teachers were good at indicating the arguments that did not support the truth of the statement for all cases and understood specific examples cannot be accepted as proof. Goetting (1995) similarly determined that pre-service teachers were aware of the distinction between proofs and empirical arguments, unlike Weber (2010), who expressed that some university students accepted empirical arguments as valid proofs. In Weber's study, aside from specific cases or tried numerical values, students were generally unsuccessful in determine whether a proof is valid or not, in line with Selden and Selden's (2003a) finding that undergraduate students had difficulty in distinguishing invalid from valid proofs.

With respect to the present study, when there were discrepancies between mathematics symbols and statements in a proof, the pre-service teachers generally managed to note the use of inappropriate statements in their evaluations. However, in cases where only inappropriate statements were given, they sometimes had difficulty in recognizing these. Moreover, they sometimes regarded insignificant details in their evaluations. For example, some of the pre-service teachers remarked in their evaluations that a statement about similar triangles was incorrect in a proof involving congruent triangles.

Finally, the pre-service teachers demonstrated that they accepted focusing on only the results as a criterion for proof evaluation, which represents a deficiency in their skills in this regard. In order to overcome this issue, provisions should be made for activities involving proof construction and evaluation in the learning process. Furthermore, as Pfeiffer (2011a) notes, interaction and discussion can contribute to skills in validating, evaluating and constructing proofs; thus, opportunities should be provided for students to discuss their proofs.

The main contribution of this study can be summarized in four items.

1. The activities of this study require both proof construction and proof evaluation. Activities of other studies about proof validation and evaluation were generally carried out as determining whether the written steps of proof are valid or sufficient for any proof (Inglis, Mejia-Ramos, Weber & Alcock 2013; Panse, Alcock & Inglis, 2018; Pfeiffer, 2010, 2011a, 2011b; Powers et al., 2010; Segal, 1999). In other words, the participants of these studies evaluated the proofs as

valid or invalid, enough, or not enough to be a proof etc. Unlike these studies, Inglis et al. (2013) allowed the participants to explain the reasons of their proof evaluations. Accordingly, the proofs were not constructed by the participants and the finished proof is generally evaluated without going into details. Moreover, the suggestions of revising the steps of proofs were not presented to the participants in other studies. The present study enables the participants to assert the errors or deficiencies of the proofs.

2. This study has a diversity of categories for proof evaluation. The proofs were particularly not evaluated in terms of different categories (*mathematical language, concepts etc.*) in other studies. Mejia-Ramos et al. (2012) presented a multidimensional model for assessing proof comprehension in undergraduate mathematics in their study. This model had a purpose of assessing students' understanding through the proofs. Criteria about assessing students' understanding were related with seven different aspects of a proof. Therefore, conceptual aspect of the study predominates between other studies and it takes the context of the proofs to forefront. In brief, the evaluations of the proofs were carried out as determining the validity of mathematical arguments, statements and conjectures asserted for proving in the studies about validating and evaluating proofs (Goetting, 1995; Inglis et al., 2013; Pfeiffer, 2010, 2011a; Segal, 1999; Stylianides & Stylianides, 2009; Weber, 2010). However, in this study the evaluations were carried out in terms of actions taking part in a proof (*giving justifications, use of mathematical language, writing the steps deductively etc.*) thoroughly. Besides, the study enabled the participants to be aware of errors or deficiencies in their own proofs and to make up for them in other evaluations. Therefore, the study has a nature of providing the participants to develop their proof construction and evaluation.
3. The categories for proof evaluation in this study can serve as an inspiration for how to write a proof. In other words, this study can have a role in revealing the necessary qualities for acceptance a proof. Therefore, the present study can contribute to notice the necessities of proof construction, and the categorization of proof evaluation criteria.
4. This study can contribute to overcome the scarcity of studies on proof evaluation. The studies on proof validation and evaluation have a less ratio among research on proof. Mejia-Ramos and Inglis (2009) and Moore (2016)

support this by stating that studies on proof construction dominates among published research.

The results from comparing the proof evaluations of pre-service teachers and researchers show that pre-service teachers are not completely capable of evaluating proofs. Proof activities in undergraduate education should be implemented in an interactive manner. In other words, instructors should encourage their students to think about proof and they should allow them to discuss the proofs thoroughly in the class.

Proof writing requires the use of justification and mathematical language. In the current study, they are among the dimensions for proof evaluation. It is critical that students evaluate their own or peers' proofs in terms of several dimensions in the learning environment. It can be suggested that the instructors lead their students to provide justifications and to analyze proofs. Thus, the students may improve the skills in both constructing and evaluating proofs.

In this study, the activities of proof construction and evaluation were implemented through group work. This implementation represents both a strength and a weakness of the instructional strategy of the study. The strength of the instructional strategy was to provide the students to construct and evaluate the proofs interactively. Thus, the students not only learn something from each other but also overcome their deficiencies. However, working in groups may sometimes bring only one person into forefront. Accordingly, the weakness of the instructional strategy was to prevent the students to think and reason about proof construction and evaluation.

Acknowledgements

Avoiding identifying any of the authors prior to peer review.

The authors would like to express their appreciation to the anonymous reviewers and the editor Päivi Portaankorva-Koivisto for making useful suggestions regarding the presentation of this paper.

References

- Anderson J. (1996). The place of proof in school mathematics. *Mathematics Teaching*, 155, 33–39.
- Anderson, J. R. (2000). *Cognitive psychology and its implications* (5th ed.). Worth.
- Bass, H. (2009). *How do you know that you know? Making believe in mathematics*.
<https://deepblue.lib.umich.edu/handle/2027.42/64280>
- 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. Retrieved February 13, 2021 from
<http://www.lettredelapreuve.org/OldPreuve/Newsletter/990708Theme/990708ThemeUK.html>
- Clarke, D. M., Clarke, D. J., & Sullivan, P. (2012). Reasoning in the Australian Curriculum: Understanding its meaning and using the relevant language. *Australian Primary Mathematics Classroom*, 17(3), 28–32.
- Clarke, D., Goos, M., & Morony, W. (2007). Problem solving and working mathematically: An Australian perspective. *ZDM Mathematics Education*, 39(5-6), 475–490.
<https://doi.org/10.1007/s11858-007-0045-0>
- de Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras*, 24, 17–24.
- de Villiers, M. (1999). *Rethinking proof with the Geometer's Sketchpad*. Key Curriculum.
- Demir, E. (2017). *Ortaöğretim matematik öğretmen adaylarının muhakeme hatalarının ispatlama bağlamında incelenmesi* [Investigation of pre-service secondary mathematics teachers' reasoning errors reasoning deficiencies and reasoning gaps within the context of proof] [Unpublished doctoral dissertation]. Karadeniz Technical University.
- Dickerson, D. S., & Doerr, H. M. (2014). High school mathematics teachers' perspectives on the purposes of mathematical proof in school mathematics. *Mathematics Education Research Journal*, 26(4), 711–733. <https://doi.org/10.1007/s13394-013-0091-6>
- Edwards, L. D. (1997). Exploring the territory before proof: Students' generalizations in a computer microworld for transformation geometry. *International Journal of Computers for Mathematical Learning*, 2, 187–215. <https://doi.org/10.1023/A:1009711521492>
- Flores, A. (2002). How do children know that what they learn in mathematics is true? *Teaching Children Mathematics*, 8(5), 269–274. <https://doi.org/10.5951/TCM.8.5.0269>
- Forman, E. A., Larreamendy-Joer, J., Stein, M. K., & Brown, C. A. (1998). You're going to want to find out which and prove it: Collective argumentation in a mathematics classroom. *Learning and Instruction*, 8(6), 527–548. [https://doi.org/10.1016/S0959-4752\(98\)00033-4](https://doi.org/10.1016/S0959-4752(98)00033-4)
- Goetting, M. (1995). *The college students' understanding of mathematical proof* (Publication No. 9539653) [Doctoral dissertation, University of Maryland]. ProQuest Dissertations and Theses Global.
- Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational Studies in Mathematics*, 44, 5–23. <https://doi.org/10.1023/A:1012737223465>
- Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange*, 21, 6–13.
- Hanna, G., Bruyn, Y., Sidoli, N., & Lomas, D. (2004). Teaching proof in the context of physics. *ZDM Mathematics Education*, 36(3), 82–90. <https://doi.org/10.1007/BF02652775>
- Herbst P., & Brach C. (2006). Proving and doing proofs in high school geometry classes: What is it that is going on for students? *Cognition and Instruction*, 24(1), 73–122.
https://doi.org/10.1207/s1532690xci2401_2
- Hill, H., Ball, L. D., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 39(4), 372–400.

- Hoyles, C. (1997). The curricular shaping of students' approaches to proof. *For the Learning of Mathematics*, 17(1), 7–16.
- Imamoglu, Y., & Yontar-Togrol, A. (2012, July 8-15). *An investigation of senior mathematics and teaching mathematics students' proof evaluation practices* [Conference presentation]. 12th International Congress on Mathematical Education, Seoul, Korea.
<http://www.lettredelapreuve.org/pdf/ICMI-12/TSG14-4-Imamoglu.pdf>
- Imamoglu, Y., & Yontar-Togrol, A. (2015). Proof construction and evaluation practices of prospective mathematics educators. *European Journal of Science and Mathematics Education*, 3(2), 130–144.
- Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, 43(4), 358–390.
<https://doi.org/10.5951/jresmetheduc.43.4.0358>
- Inglis, M., Mejia-Ramos, J. P., Weber, K., & Alcock, L. (2013). On mathematicians' different standards when evaluating elementary proofs. *Topics in Cognitive Science*, 5(2), 270–282.
<https://doi.org/10.1111/tops.12019>
- Knuth, E. (2002). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education*, 33, 379–405. <https://doi.org/10.2307/4149959>
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Sage Publications.
- Lee, K. (2011). *Students' logical reasoning and mathematical proving of implications* (Publication No. 3454803) [Doctoral dissertation, Michigan State University]. ProQuest Dissertations and Theses Global.
- Lesseig, K. (2016). Investigating mathematical knowledge for teaching proof in professional development. *International Journal of Research in Education and Science (IJRES)*, 2(2), 253–270.
- Lesseig, K., Hine, G., Na, G. S., & Boardman, K. (2019). Perceptions on proof and the teaching of proof: A comparison across preservice secondary teachers in Australia, USA and Korea. *Mathematics Education Research Journal*, 31, 393–418.
<https://doi.org/10.1007/s13394-019-00260-7>
- Lin, F. L., Yang, K. L., & Chen, C. Y. (2004). The features and relationships of reasoning, proving, and understanding proof in number patterns. *International Journal of Science and Mathematics Education*, 2, 227–256. <https://doi.org/10.1007/s10763-004-3413-z>
- Manin, Yu. I. (2010). *A course in mathematical logic for mathematicians* (2nd ed.). Springer-Verlag.
- Marrades, R., & Gutiérrez, Á. (2000). Proofs produced by secondary school students learning geometry in a dynamic computer environment. *Educational Studies in Mathematics*, 44(1), 87–125.
- Mejia-Ramos, J. P., & Inglis, M. (2009). Argumentative and proving activities in mathematics education research. In F.-L. Lin, F.-J. Hsieh, G. Hanna & M. de Villiers (Eds.), *Proceedings of the ICMI study 19 conference: Proof and proving in mathematics education* (Vol. 2, pp. 88–93), Taipei, Taiwan.
- Mejia-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, 79, 3-18. <https://doi.org/10.1007/s10649-011-9349-7>
- Mercer, N., Dawes, L., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: Ways of helping children to use language to learn science. *British Educational Research Journal*, 30(3), 359–377. <https://doi.org/10.1080/01411920410001689689>
- Millî Eğitim Bakanlığı [MEB]. (2013). *Ortaöğretim matematik dersi* (9, 10, 11 ve 12. sınıflar) *öğretim programı* [Curriculum for secondary school mathematics (9th, 10th, 11th and 12th grades)]. MEB Yayınları.

- Moore, R. C. (2016). Mathematics professors' evaluation of students' proofs: A complex teaching practice. *International Journal of Research in Undergraduate Mathematics Education*, 2(2), 246–278. <https://doi.org/10.1007/s40753-016-0029-y>
- Nasibov, F., & Kaçar, A. (2005). Matematik ve matematik eğitimi hakkında [On the mathematics and mathematics education]. *Kastamonu Eğitim Dergisi* [Kastamonu Education Journal], 13(2). 339–346.
- National Council of Teachers of Mathematics [NCTM]. (2000). *Principles and standards for school mathematics*. NCTM.
- Öztürk, T. (2016). *Matematik öğretmeni adaylarının ispatlama becerilerini geliştirmeye yönelik tasarlanan öğrenme ortamının değerlendirilmesi* [The evaluation of the learning environment designed for improving pre-service mathematics teachers' proving skills] (Doctoral dissertation). Karadeniz Technical University, Trabzon, Turkey.
- Öztürk, T. (2020, Mart). Matematğin kalbi: İspat [The heart of mathematics: Proof]. *APSİSTEK*. <https://www.apsistek.com/index.php/mart-2020/178-matematigin-kalbi-ispata>
- Panase, A., Alcock, L., & Inglis, M. (2018). Reading proofs for validation and comprehension: An expert-novice eye-movement study. *International Journal of Research in Undergraduate Mathematics Education*, 4, 357–375. <https://doi.org/10.1007/s40753-018-0077-6>
- Perry, P., Molina, Ó., Camargo, L., & Samper, C. (2011, February 9-13). *Analyzing the proving activity of a group of three students* [Conference presentation]. The Seventh Congress of the European Society for Research in Mathematics Education (CERME 7), Rzeszów, Poland. http://www.cerme7.univ.rzeszow.pl/WG/1/Cerme7_WG1_Perry.pdf
- Pfeiffer K. (2011b, February 9-13). *A schema to analyse students' proof evaluations* [Conference presentation]. The Seventh Congress of the European Society for Research in Mathematics Education (CERME 7), Rzeszów, Poland. http://www.cerme7.univ.rzeszow.pl/WG/1/Cerme7_WG1_Pfeiffer.pdf
- Pfeiffer, K. (2010). How do students evaluate and compare mathematical proofs? *Research in Mathematics Education*, 12(2), 161–162. <https://doi.org/10.1080/14794802.2010.496986>
- Pfeiffer, K. (2011a). *Features and purposes of mathematical proofs in the view of novice students: observations from proof validation and evaluation performances* (Publication No. U576792) [Doctoral dissertation, National University of Ireland]. ProQuest Dissertations and Theses Global.
- Powers, R. A., Craviotto, C., & Grassl, R. M. (2010). Impact of proof validation on proof writing in abstract algebra. *International Journal of Mathematical Education in Science and Technology*, 41(4), 501–514. <https://doi.org/10.1080/00207390903564603>
- Pulley, C. A. (2010). *Using instruction to investigate the effects of assessing reasoning tasks on students' understanding of proof* (Publication No. 3438561) [Doctoral dissertation, Illinois State University]. ProQuest Dissertations and Theses Global.
- Rav, Y. (1999). Why do we prove theorems? *Philosophia Mathematica*, 7(1), 5–41. <https://doi.org/10.1093/philmat/7.1.5>
- Remillard, K. S. (2009). *The mathematical discourse of undergraduate mathematics majors: The relation to learning proof and establishing a learning community* (Publication No. 3369995) [Doctoral dissertation, Indiana University of Pennsylvania]. ProQuest Dissertations and Theses Global.
- Segal, J. (1999). Learning about mathematical proof: Conviction and validity. *Journal of Mathematical Behaviour*, 18(2), 191–210. [https://doi.org/10.1016/S0732-3123\(99\)00028-0](https://doi.org/10.1016/S0732-3123(99)00028-0)
- Selden, A., & Selden, J. (2000). *Proof, validation, and trains of thought*. Retrieved February 10, 2021, from <https://www.maa.org/question-2-proof-validation-and-trains-of-thought>

- Selden, A., & Selden, J. (2003a). Validations of proofs written as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, 34(1), 4–36. <https://doi.org/10.2307/30034698>
- Selden, A., & Selden, J. (2003b). *Errors and misconceptions in college level theorem proving* (Report No. 2003-3). Cookeville, TN: Tennessee Technological University.
- Selden, J., & Selden A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, 29(2), 123–151. <https://doi.org/10.1007/BF01274210>
- Senk, S. L. (1983). *Proof-writing achievement and Van Hiele levels among secondary school geometry students* (Publication No. T-28618) [Doctoral dissertation, The University of Chicago]. ProQuest Dissertations and Theses Global.
- Stacey, K., & Vincent, J. (2009). Modes of reasoning in explanations in Australian eighth-grade mathematics textbooks. *Educational Studies in Mathematics*, 72(3), 271–288. <https://doi.org/10.1007/s10649-009-9193-1>
- Stylianides, A. J., & Ball, D. L. (2008). Understanding and describing mathematical knowledge for teaching: Knowledge about proof for engaging students in the activity of proving. *Journal of Mathematics Teacher Education*, 11, 307–332. <https://doi.org/10.1007/s10857-008-9077-9>
- Stylianides, A. J., & Stylianides, G. J. (2009). Proof constructions and evaluations. *Educational Studies in Mathematics*, 72, 237–253. <https://doi.org/10.1007/s10649-009-9191-3>
- Stylianides, A. J., Stylianides, G. J., & Philippou, G. N. (2004). Undergraduate students' understanding of the contraposition equivalence rule in symbolic and verbal contexts. *Educational Studies in Mathematics*, 55(1), 133–162. <https://doi.org/10.1023/B:EDUC.0000017671.47700.0b>
- Tabach, M., Levenson, E., Barkai, R., Tirosh, D., Tsamir, P., & Dreyfus, T. (2010). Secondary school teachers' awareness of numerical examples as proof. *Research in Mathematics Education*, 12(2), 117–131. <https://doi.org/10.1080/14794802.2010.496973>
- Thompson, D. R., Senk, S. L., & Johnson, G. J. (2012). Opportunities to learn reasoning and proof in high school mathematics textbooks. *Journal for Research in Mathematics Education*, 43(3), 253–295. <https://doi.org/10.5951/jresmetheduc.43.3.0253>
- Varghese, T. (2011). Considerations concerning Balacheff's 1988 taxonomy of mathematical proofs. *Eurasia Journal of Mathematics, Science & Technology Education*, 7(3), 181–192. <https://doi.org/10.12973/ejmste/75192>
- Weber K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education*, 39(4), 431–459. <https://doi.org/10.5951/jresmetheduc.39.4.0431>
- Weber, K. (2005). Problem-solving, proving, and learning: The relationship between problem solving processes and learning opportunities in the activity of proof construction. *Journal of Mathematical Behavior*, 24, 351–360. <https://doi.org/10.1016/j.jmathb.2005.09.005>
- Weber, K. (2009, February). *Mathematics majors' evaluation of mathematical arguments and their conception of proof* [Conference presentation]. Proceedings of the 12th Conference on Research in Undergraduate Mathematics Education, Raleigh, North Carolina. http://sigmaa.maa.org/rume/crume2009/Weber1_LONG.pdf
- Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning*, 12, 306–336. <https://doi.org/10.1080/10986065.2010.495468>