Is This a Proof? Students' Initial Criteria

Gülseren Karagöz Akar	Dru Horne	AnnaMarie Conner
Boğaziçi University	University of Georgia	University of Georgia
Aida Alibek	Ngutor Tembe	Uyiosa O. Ugiagbe
University of Georgia	University of Georgia	University of Georgia

Criteria for proof are important because future teachers will need to evaluate students' arguments as proofs or nonproofs. This preliminary report examines how a group of prospective and practicing teachers in a reasoning and proof course began to collectively construct criteria for proof. Given a set of hypothetical student arguments, their initial criteria differed, despite common mathematics coursework, including an introduction to proof course. Ongoing analysis uses documenting collective activity to understand how collective criteria and definition were constructed.

Keywords: proof definition, prospective teachers, practicing teachers, proof criteria, argumentation

Proofs are an essential component in doing, communicating, and recording mathematics (Schoenfeld, 1994). Hersh (1993) contrasted proofs from a mathematical point of view and in a classroom setting. Within the creation and development of mathematics, proofs are social constructs for convincing others such that "without complete, correct proof, there is no mathematics" (p. 396). From a humanistic view, a proof is an explanation. Considering the role of proof in a classroom setting, Stylianides and Stylianides (2009) described proof as "general, valid, and accessible to the members of the community" (p. 239). Notably, Stylianides (2007) viewed proofs as mathematical arguments with certain properties: 1) proofs contain statements which are familiar to all students in a classroom, 2) the reasoning behind the steps of the proof should be valid for and known to all students in the class; and 3) the representation of the proof should be accepted by the classroom community.

The National Council of Teachers of Mathematics (NCTM, 2000) asserted that reasoning and proof "should be a consistent part of students' mathematical experience in prekindergarten through grade 12" (p. 56). Along with the consideration of proof within a classroom setting, the Common Core State Standards for Mathematics (CCSSM, 2010) further emphasized that students need to build a logical progression of statements to explore the truth of their conjectures by justifying their conclusions, communicating and responding to the arguments of others, and recognizing and using counter examples. Notably, students are expected to develop a proof conception that involves "a person's ability to prove and analyze proofs and their beliefs about proof and justification" (Conner, 2007, p.11). Thus, teachers are asked to embed proof as a core practice in their daily mathematical instruction (NCTM, 2000).

However, previous research has shown that secondary school students, collegiate students, and pre-service teachers mostly focus on empirical arguments while constructing proofs (e.g., Knuth, 2002a, 2002b; Reyhani et al., 2012; Sears, 2019). Despite their beliefs about the limitations of empirical examples (e.g., Stylianou et al., 2015), students have difficulties in distinguishing proofs from non-proofs. They have a tendency towards considering generic arguments and pictorial representations as non-proofs (e.g., Lesseig et al., 2019), and they consider the role of proof as mostly a verification of the truth of a mathematical statement (e.g., Lesseig et al., 2019; Likando & Ngoepe, 2014; Varghese, 2009). Though research has shown that some practicing teachers considered proof as socially constructed (Knuth, 2002a), some prospective teachers did not consider proof as a product of collective engagement (e.g., Sears, 2019; Varghese, 2009). On the other hand, research has shown that given education, students' and prospective teachers' conception and comprehension of proofs can be developed (e.g., Belin & Karagöz Akar, 2020; Miller, 2020). These results align with Stylianides et al. (2017):

> "There is a great need for interventions to introduce students of different ages to an appropriate set of criteria for judging whether an argument meets the standard of proof…Unless the field of mathematics education designs effective ways to instill in students expandable as well as developmentally and mathematically appropriate views of what proof means, it is unrealistic to expect that students will make sustained progress in their learning of proof during their mathematical education" (p. 257-258).

Taking this as a challenge and building on the previous studies, in this study, we further investigated how prospective and practicing mathematics teachers come to collectively develop criteria for proof, by which we mean criteria to determine whether an argument can be considered a proof. Acknowledging that defining proof and identifying proofs and non-proofs are important components of knowledge for teaching proof (Steele & Rogers; 2012) and considering that teachers are key figures in the education of future generations (NCTM, 2000), we specifically investigated the following research question: *How do practicing and prospective teachers come to collectively construct and use a community definition of proof?*

In this paper, we show how teachers and prospective teachers who had a significant amount of mathematics coursework prior to the course, often disagreed as to whether an argument was a proof or not. We include an excerpt of our data that depict how at times their written and verbal work seemed to disagree. These data exemplify why it is important for a class to have explicit collective criteria for proof as they continue to evaluate arguments. We are interested in how different individual warrants and criteria for proof evolved into a collective understanding or definition. At the beginning of the course, we observed that when presented with the same student work, teachers can have very different understandings of what constitutes proof; ultimately in the course, they constructed a collective definition. We are in the process of analyzing how the teachers moved from the initial activity shared here to the collective criteria.

Methods

We collected data from a reasoning and proof class that the third author taught in the fall semester of 2024. The class was comprised of 12 students who were prospective or practicing secondary mathematics teachers. The participants had previously completed an introduction-to-proofs course as well as several other mathematics courses. The class met weekly alternating between 165 minute in-person classes and asynchronous classes. The asynchronous classes were weeks 1, 3, 5 and 6, and the in-person classes were weeks 2, 4, and 7. We only collected data from the three in-person classes. The data consisted of videos of three groups of four participants each (except for week 2, in which one group had three), videos of whole class discussions, pictures of participants' written work, and pictures of whole class board work. Because we are focused in this paper on the first part of our research question regarding the construction of a collective definition of proof, we only describe the activities and tasks that preceded the agreed upon class criteria for proof.

In the first class, the participants completed three tasks. For the first task, we gave them 16 solutions to the question "How many 3 x 3 squares are there in an n x n square?" and asked them to discuss in their groups whether they were convinced by each solution and what modifications they would make to the solution. For the second task, we gave the students 23 arguments for "Prove that when you add any 2 odd numbers, your answer is always even" and asked them to sort the arguments according to their underlying ideas. For each of their sorted groups, we asked them to determine if the underlying idea had merit for proving the statement and to select the best argument in the group. The participants had written their own solutions to these two problems in preparation for class, so they were familiar with the problems. The third task we gave included 10 arguments for the statement "the sum of two odd numbers is even" and asked students to decide if the argument is a proof and write a rationale for their decision. An example of the conversation around one of these arguments is presented below.

At the end of the first class, the participants were asked to write down, as a group, criteria for determining if a given argument is a proof. These criteria were shared with the whole class and compiled by the instructor into a list of 14 criteria. The participants were then asked to refine the class list to include fewer criteria, which they were unable to finish due to time constraints. We began the second in-person class (week 4) by asking them to revisit the class list and to consolidate the list into a set of three criteria. As a result of small group and whole class discussion, the participants agreed on the following criteria for proof: 1) makes logical connections using mathematical and community-appropriate evidence, 2) accounts for all cases, and 3) comes to the conclusion of the initial conjecture.

Our analysis examined the participants' arguments as they discussed each presented argument. We used Rasmussen and Stephen's (2008) Documenting Collective Activity (see also Cole et al., 2012; Rasmussen et al., 2015) to examine how the class collectively constructed their understanding of criteria for proof. In this method, the data, claims, and warrants are identified and recorded; three criteria are then used to examine when ideas begin to function as if shared in the classroom community. To ascertain when an idea or way of reasoning is functioning as if shared in a community, we used three criteria: 1) when the warrant was initially present for a claim but is no longer necessary; 2) when a part of an argument serves a different function in a later argument (i.e., a claim originally needed justification but is later used as data or warrant for another claim); 3) when an idea is used as data or warrant for different claims in multiple arguments (Rasmussen & Stephen, 2008; Cole et al., 2012). Our research team members individually examined participant small group discourse, making tentative codes for data, claim, and warrant. We then came together and constructed argumentation logs of each small group and whole class discussion, discussing and reconciling differences. Note that our analysis was conducted from the perspective of a participant, so we assigned ideas the function that we believed most closely aligned with the participant's intent. Thus, warrants sometimes preceded claims, and statements were not assigned a function based on how we would construct an argument but by what the participants seemed to intend with their utterances. At times, the participants included qualifiers and rebuttals in their arguments, which we captured in our argumentation log. We also included participants' written arguments in our argumentation log. These were usually easier to identify because we specifically asked for claims and reasoning.

Results

Participants were given 10 arguments purporting to prove that the sum of two odd numbers is even, each from a hypothetical student. The following excerpt is from one small group's discussion of Student E's argument (see Figure 1), which was the fifth argument in the list.

Student E

Any odd number can be written as 2x + 1. So let's add two odd numbers.

2x + 1 + 2x + 1 = 4x + 2
4x + 2 is even since 4 and 2 are both even.
Or 2(2x + 1) shows that 4x + 2 is even.

Figure 1. Student E's work shared with participants for their evaluation of whether the argument is a proof or not.

1 *Lindsey:* Nooo! (shaking her head and mouthing in response to reading Student E's response)
2 *Chloe:* I feel like it is almost there.
3 *Debbie:* It is almost there.
4 *Lindsey:* Yeah, but it is adding two of the same number.
5 *Chloe:* Mm-hmm (affirmative).
6 *Lindsey:* They also don't tell you what x is.
7 *Chloe:* I feel like if they use x and y (inaudible).
8 *Lindsey:* Yeah, yeah.
9 *Chloe:* If you used two different variables. And they don't really say why $4x + 2$ is even.
10 Well, I guess they say because four and two are even. Having that four times x is
11 like they don't really say what that represents.

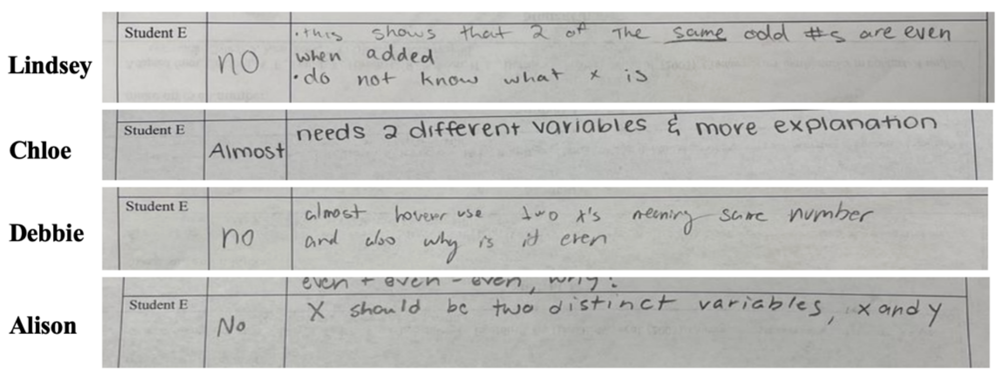

Figure 2. Participants' written responses to Student E's work.

We inferred that Chloe and Lindsey had two different claims. Notably, Chloe claimed that Student E's work was "almost there" (line 2), and Debbie supported her claim (line 3); Lindsey was very assertive, saying "noooo" (line 1), to claim that Student E's work was not a proof. In lines 4 and 6, Lindsey provides warrants for her claim that this was not a proof; the other members of the group seem to agree and accept this as a warrant for their claim that it is almost a proof. That is, each one of them seemed to be interpreting the same warrant as a warrant for their personal claim (potentially without realizing that the other person was using the warrant in their own way). Analysis of the participants' written work (Figure 2) further suggested that they did not come to consensus on whether Student E's work was a proof or not. Interestingly, although Debbie agreed with Chloe during their discussion, in her written response, she made the claim that Student E's work was not a proof. She made a definite claim, but she also explained in her warrant "almost however use two x's meaning same number" (see Figure 2).

Discussion and Conclusion

Our preliminary conclusions about the participants' initial criteria include that participants came into the course with their own somewhat idiosyncratic criteria for whether a student argument was a proof. At times, it seemed that participants translated their own criteria for proof into an assessment of a student argument by suggesting that they could tell what the student "meant" by something that was not necessarily correct. They also wondered if the grade level of a student impacted whether an argument could be considered a proof, suggesting that even if the argument would not be a proof for themselves, it could be for a student. This is consistent with the perspectives of other prospective teachers who have examined student arguments (Suominen et al., 2018).

In our analysis, we have begun to see groups of participants use previous claims as they assessed additional students' arguments. Although we have not completed our analysis of how ideas began to function as if shared, we have preliminary evidence of such ideas even within small group discussions. We believe it is important to document that the participants began their discussions with differing conceptions and criteria for proof. Otherwise, coming to consensus and shared criteria would not be necessary. In our presentation, we will include additional evidence of initial difference as well as ideas that began to function as if shared, culminating in our whole class discussion of criteria for proof. We will engage attendees in examining initial criteria for proof from our data and ask attendees to discuss the following questions:

- What criteria for proof would you expect undergraduate students to use at the beginning of a course?
- How are criteria from advanced mathematics classes translated to expectations for students' proofs?
- Have you attended to students' initial and developing criteria for proof?
- On the surface, students can seem to have a consensus, but in our case, students had different claims (even with a similar warrant). How do you engage with these kinds of situations in your own practice?

In this study, we are investigating how prospective and practicing teachers come to collectively construct and use a community definition of and criteria for proof to evaluate different student arguments. We shared some data from the beginning of the study to show that the participants did not necessarily begin with a shared set of criteria for proof. In fact, our analysis of the data showed that using the same explanations (warrants), the participants made different claims about the same argument. We suggest that this lack of consensus provides a rationale for investigating how a collective definition and set of criteria can be constructed. The Committee on the Undergraduate Program in Mathematics (CUPM, 2015), appointed by the Mathematical Association of America, emphasized that mathematics undergraduates, including prospective teachers, "should learn to read, understand, analyze, and produce proofs at increasing depth as they progress through a major" (p. 11). This further provides motivation for mathematics teacher educators to consider how a learning environment may influence practicing teachers' and prospective teachers' conceptualizations of proof for different audiences. This is especially important because teachers have difficulties in distinguishing proofs from non-proofs (e.g., Lesseig et al., 2019), and some do not consider proofs as a product of collective engagement (e.g., Sears, 2019; Varghese, 2009).

References

Belin, M., & Karagöz Akar, G. (2020). The effect of quantitative reasoning on prospective mathematics teachers' proof comprehension: The case of real numbers. *The Journal of Mathematical Behavior*, *57*, 1-21, 100757. https://doi.org/10.1016/j.jmathb.2020.100757

Cole, R., Becker, N., Towns, M., Sweeney, G., Wawro, M., & Rasmussen, C. (2012). Adapting a methodology from mathematics education research to chemistry education research: Documenting collective activity. International Journal of Science and Mathematics Education, 10(1), 193–211. https:// doi. org/ 10. 1007/ s10763- 011- 9284-1

Committee on the Undergraduate Program in Mathematics (2015). *CUPM Curriculum Guide to Majors in the Mathematical Sciences.* The Mathematical Association of America.

Conner, A., (2007). *Student teachers' conceptions of proof and facilitation of argumentation in secondary mathematics classrooms*, Ph.D. Thesis, The Pennsylvania State University.

Common Core State Standards Initiative. (2010). *Common core state standards for mathematics (CCSSM)*. National Governors Association Center for Best Practices and the Council of Chief State School Officers. Retrieved September 6, 2023, https://learning.ccsso.org/wp-content/uploads/2022/11/ADA-Compliant-Math-Standards.pdf

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

Knuth, E. J. (2002a). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education*, *33*(5), 379-405.

Knuth, E. J. (2002b). Teachers' conceptions of proof in the context of secondary school mathematics. *Journal of Mathematics Teacher Education*, *5*(1), 61-88.

Lesseig, K., Hine, G., Na, S. G., & 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.

Likando, K. M. & Ngoepe, M. G. (2014). Investigating mathematics trainee teachers' conceptions of proof writing in algebra: A case of one college of education in Zambia. *Mediterranean Journal of Social Sciences, 5*(14), 331-338.

Miller, D. & CadwalladerOlsker, T. (2020). Investigating undergraduate students' view of and consistency in choosing empirical and deductive arguments, *Research in Mathematics Education, 22* (3), 249-264, https:// 10.1080/14794802.2019.1677489.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. National Council of Teachers of Mathematics

Rasmussen, C., & Stephan, M. (2008). A methodology for documenting collective activity. In A. Kelly, R. Lesh, & J. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics teaching and learning* (pp. 195–215). Routledge.

Rasmussen, C., Wawro, M. & Zandieh, M. (2015) Examining individual and collective level mathematical progress. Educational Studies in Mathematics, 88, 259–281. https://doi.org/10.1007/s10649-014-9583-x

Reyhani, E., Hamidi, F., & Kolahdouz, F. (2012). A study on algebraic proof conception of high school second graders. *Procedia-Social and Behavioral Sciences*, *31*, pp. 236–241.

Sears, R. (2019). Proof schemes of pre-service middle and secondary mathematics teachers. *Investigations in Mathematics Learning*, *11*(4), 258-274.

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

Steele, M. D. & Rogers, K. C. (2012). Relationships between mathematical knowledge for teaching and teaching practice: The case of proof. *Journal of Mathematics Teacher Education*, *15*(2), 159–180.
Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education*, *38*(3), 289–321.
Stylianides, A. J. & Stylianides, G. J. (2009). Proof constructions and evaluations. *Educational Studies in Mathematics, 72*(2), 237–253.
Stylianides, G. J., Stylianides, A. J., & Weber, K. (2017). Research on the teaching and learning of proof: Taking stock and moving forward. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 237–266). Reston, VA: National Council of Teachers of Mathematics.
Stylianou, D. A., Blanton, M. L. & Rotou, O. (2015). Undergraduate students' understanding of proof: Relationships between proof conceptions, b beliefs, and classroom experiences with learning proof. International Journal of Research in Undergraduate Mathematics Education, 1, 91–134. https://doi.org/10.1007/s40753-015-0003-0
Suominen, A., Conner, A., & Park, H. (2018). Prospective mathematics teachers' expectations for middle grades students' arguments. *School Science and Mathematics*, *118*(6), 218–231. doi:10.1111/ssm.12292
Varghese, T. (2009). Secondary-level student teachers' conceptions of mathematical proof. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, *1*.