Prospective teachers' evaluations of students' proofs by mathematical induction

Hyejin Park
University of Georgia

This study examines how prospective secondary teachers validate several proofs by mathematical induction (MI) from hypothetical students and how their work with proof validations relates to how they grade their students' proofs. When asked to give criteria for evaluating a student's argument, participants wished to see a correct base step, inductive step, and algebra. However, participants prioritized the base step and inductive step over assessing the correctness of the algebra when validating and grading students' arguments. All of the participants gave more points to an argument that presented only the inductive step than to an argument that presented only the base step. Two of the participants accepted the students' argument addressing only the inductive step as a valid proof. Further studies are needed to determine how prospective teachers evaluate their students' arguments by MI if many algebraic errors are present, especially in the inductive step.

Key words: Mathematical induction, Prospective secondary teachers, Proof validation, Proof grading

The proof method of mathematical induction (MI) is significant in the discipline of mathematics. In the *Principles and Standards for School Mathematics*, the National Council of Teachers of Mathematics (2000) asserts "students should learn that certain types of results are proved using the technique of mathematical induction" (p. 345). Secondary mathematics teachers are expected to teach MI (e.g., Australian Curriculum, Assessment, and Reporting Authority, 2012; California Department of Education, 2013; Korean Ministry of Education, Science, and Technology, 2012) and, therefore, are required to have a robust knowledge of MI as a prerequisite, including proficiency in reading and analyzing students' arguments that use MI. Most of the previous studies on the learning and teaching of MI have focused on examining either the students' or the teachers' knowledge of MI, showing their difficulties with MI, especially in their proof production or while exploring the pedagogy of MI for better supporting students' learning. Little research, however, has been devoted to how teachers read and reflect on students' arguments using MI. In this study, I examine the characteristics of five prospective secondary teachers when validating and grading student arguments using MI. These arguments were presented in an interview setting and situated in the context of teaching at the secondary level.

Relevant Literature

Proof validation is an important mathematical activity, especially for mathematics undergraduates, prospective and practicing teachers, and mathematicians (Selden & Selden, 2003). Weber (2008) stated, "Teachers need to be able to determine if the justifications and proofs that students submit are acceptable and to provide feedback when they are not" (p. 4). Some researchers have begun to examine how undergraduate students, practicing teachers, and mathematicians validate proofs, but there have been few studies focusing on MI. Knuth (2002) found that some practicing teachers accepted an argument by MI as a proof by relying on its form (appearance) rather than understanding its reasoning. Dickerson (2006) found the same result in his study with two prospective teachers. In the process of examining both prospective secondary and elementary teachers' knowledge of proof by MI, Stylianides, Stylianides, and Philippou (2007) asked participants to validate two arguments, which were invalid. Stylianides et al. reported that although both groups had similar difficulties with MI, the prospective secondary teachers validated arguments more accurately than the prospective elementary teachers. In their study, participants who provided correct answers recognized that the first argument that they were asked to validate omitted the base step and judged the argument invalid. However, some of them were not able to explain the necessity of the base step. Stylianides et al. concluded that their participants focused on the form of proof by MI during proof validation.

Grading students' proofs is also an important teaching practice, but there has been little attention to how teachers assess and respond to students' written work. The process of *proof grading* includes judgments about a proof's validity, clarity, and readability. Previous studies (e.g., Inglis, Mejía-Ramos, Weber, & Alcock, 2013; Weber, 2008) showed that mathematicians used different criteria when evaluating students' proofs and disagreed on what arguments are considered valid proofs. These studies lead us to expect that mathematicians might also use different criteria when grading students' proofs and the scores that they give students might vary. Moore (2015), in his preliminary study, reported that the scores that four mathematicians assigned to students' proofs varied drastically even though they agreed on their overall evaluations of the proofs.

Theoretical Framework

Teacher learning occurs in multiple contexts such as "university mathematics and teacher-preparation courses, preparatory field experiences, and schools of employment" (Peressini, Borko, Romagnano, Knuth, & Willis, 2004, p. 69). According to the situative perspective, only relying on an individual's acquisition of knowledge without consideration of his or her participation in social contexts leads to difficulties in understanding his or her practices. A situative perspective is relevant for understanding how a teacher's knowledge can be recontextualized across situations. Borko et al. (2000) showed that the situative perspective assisted in understanding how a teacher, Ms. Savant, transferred her conceptions of proof as she participated in the multiple contexts of teacher education and in her actual teaching. Because I was interested in participants' conceptions of proof by MI in different situations, I used this situative lens (following Peressini et al., 2004) and situated my interview questions and proof tasks in participants' roles as teachers and students. The situative perspective was useful in making sense of the participants' responses. Because they had encountered MI as students and could imagine themselves encountering MI as teachers, the participants often referenced the settings of university and middle or high school mathematics classes when evaluating students' arguments and giving answers about their conceptions of proof by MI in school mathematics.

The activity of proof validation requires judging the correctness of arguments. Validating arguments is an important part of a teacher's work in assessing student work. A validator's judgment of whether an argument is a valid proof or not occurs mentally in his or her work on proof validation and, therefore, might not be observable. For analysis of the participants' proof validations, I referred to Selden and Selden's (2003) description of proof validation, which demonstrates it as a complex process by which someone reads and reflects on an argument in order to determine its validity. They suggested that the activity of proof validation includes such things as "asking and answering questions, assenting to claims, constructing sub-proofs, remembering or finding and interpreting other theorems and definitions, complying with instructions, and conscious feelings of rightness or wrongness" (p. 5). In this study, I examined the participants' behaviors in their proof validations and how they judged whether arguments by MI were valid.

Methodology

Five prospective secondary teachers, who were concurrently enrolled at the University of Georgia in either the undergraduate secondary mathematics teacher education program or the master's degree program leading to certification as a secondary school mathematics teacher, participated in this study from July to the middle of November 2014. Three of the participants were pursuing dual degrees in mathematics and mathematics education. All had taken Introduction to Higher Mathematics offered by the Mathematics Department in which they learn mathematical reasoning and proof writing, including proof by MI. Pseudonyms were used for identifying the participants –Emily, Jason, David, Brad, and Blain – to protect their anonymity. For this study, I conducted semi-structured interviews of about 80 minutes in length (one interview per participant). In the interviews, participants were asked to communicate their thoughts about the teaching and learning of MI, to prove two mathematical statements (an equation problem and an inequality problem), and to evaluate students' arguments purported to be proofs by MI that respond to the same statements (three arguments per statement). I created the student arguments used in this study by referring to literature (e.g., Baker, 1996; Harel, 2002) showing students' common mistakes in proving by MI (see Table 1 for a summary of the proof tasks). When validating arguments, the participants were asked what they thought about each argument and whether each argument was a valid proof and was convincing. Also, they were asked to assign a grade (out of 10 possible points) for each argument. The following are some of the questions I asked during the interview: *Is this argument a valid proof? Why? How many points would you assign each argument? What factors would go into your grading?* I conducted, video-recorded and transcribed all the interviews, and the transcriptions were checked by another person to verify their accuracy.

Table 1

A Summary of the Proof Tasks Presented to the Participants

Problem	Argument	Argument summary
Prove that for any positive integer n, $1 + 2 + \cdots + \mathrm{n} = \frac{\mathrm{n}(\mathrm{n}+1)}{2}$	Rebecca's	No base step
	Shane's	No inductive step This argument concludes that the statement is true from three cases
	Polly's	One minor algebraic error in the inductive step
Prove that for any positive integer $\mathrm{n} \geq 4$, $2^{\mathrm{n}} < n!$	Kelly's	Incorrect base step This argument shows that the statement is not true for $n = 1$
	Garrison's	One equality/inequality error in the inductive step
	Laura's	In the inductive step, this argument addresses the induction hypothesis but does not show how $P(k+1)$ is derived from $P(k)$

For the data analysis, I used an open coding system (Strauss & Corbin, 1990). I first identified and coded parts of the data where participants talked about MI in general and then separated them from the parts in which the participants were working with the six arguments. After that, I summarized how each participant validated the students' arguments, including grades/scores they would give the arguments, and then compared their work across the participants.

Results

In this section, I report characteristics of the participants' work when validating and grading students' arguments that use MI and what relationships existed between their work with proof validation and their grading work on the students' arguments. Participants used similar criteria when evaluating students' arguments such as a correct base step, induction hypothesis, inductive step, and algebra. Some of the participants considered whether the arguments addressed the concluding statements and used the $P(n)$ notation, but they did not focus as much on these aspects in their evaluations.

The Equation Problem (Prove that for any positive integer *n,* $1 + 2 + \cdots + \mathrm{n} = \frac{\mathrm{n(n+1)}}{2}$)

When analyzing Rebecca's argument, all participants recognized that there was no base step, and three of them concluded that this argument was not a valid proof. Two of participants, David and Brad, accepted this argument as a valid proof based on either their past learning experiences with MI or their perceptions of proof by MI. For example, when asked whether Rebecca's argument was a valid proof, Brad said, "It's a valid proof. Like I said, the only problem is, basically, there is no base case and there was no checking that the statement is true with that." When grading Rebecca's argument, participants took two or three points off (out of ten possible points) on average (See Table 2). For Shane's argument, all participants pointed out that there was no inductive step and concluded that this argument was not a valid proof. When grading Shane's argument, they gave him lower grades than they had given Rebecca's argument by observing that either the inductive step was an important part of proof by induction or that the inductive step was harder for students to understand than the base step. David, for instance, stated, "I think any students could be able to prove the base case, because that's not that hard, and any students could be able to look at the inductive step and then to say if the statement is true or not. But, actually defining a statement from a given problem and then taking out the inductive step takes a lot more careful effort and more cognitive demand. And, so that's why they put more emphasis on those parts of the questions." For Polly's argument, the participants checked each step of the argument, including whether the base step and inductive step were using algebra correctly. However, none of the participants recognized one minor algebraic error in the inductive step, even though four of them had correctly proven this statement before examining the students' arguments (Jason was not able to complete the inductive step). They determined that this argument was a valid proof in that everything – the base step, inductive step, and algebra – was correct and gave it full credit (see Table 2).

The Inequality Problem (Prove that for any positive integer $\mathrm{n} \geq 4$, $2^{\mathrm{n}} < n!$)

When validating Kelly's argument, all of the participants recognized that she used an incorrect base case, determined that this argument was not a valid proof and gave her a small amount of credit (less than 3 out of the 10 points; see Table 2). In evaluating Garrison's argument, two of the participants found one algebraic error in the inductive step, but they did not put as much emphasis on this minor error in their validation and even in their grading. Brad said, "I'm less concerned with so much of the algebra. I'm looking at the logic and the use of induction," while completing his validation work. The other three participants did not recognize the error. However, all of participants concluded that Garrison's argument was a valid proof, and all except David gave him full credit. David took one point off Garrison's argument by pointing out that "he did not define $P(n)$." David was the only participant who discussed the proper use of notation in his evaluation of Garrison's argument. As for Laura's argument, all of participants recognized that Laura addressed only the inductive hypothesis and did not show the inductive step. So, they concluded that this argument was not a valid proof and gave her 5.4 points (out of 10) on average (see Table 2).

Table 2

Scores Assigned to the Students' Arguments by the Participants

Participant	Argument (out of 10 possible points)					
	Rebecca's	Shane's	Polly's	Kelly's	Garrison's	Laura's
Emily	8	5	10	2	10	6
Jason	7	5	10	—*	10	7
David	8	2	10	0	9	4
Brad	8 or 9	3 or 4	10	2 or 3	10	6
Blain	6 or 7	2	10	1	10	4
Average score	7.6	3.4	10	1.25	9.8	5.4

* Jason did not assign a score for Kelly's argument, but instead, he stated, "I won't give her zero. I would say…just some kind of credit" when asked to assign a grade for Kelly's argument.

Conclusion

Overall, when grading the arguments, the participants gave more points to an argument that presented only the inductive step, rather than an argument that presented only the base step. Participants gave an argument full credit when they concluded that it included the correct base step, inductive step, and algebra. Even when they noted a minor algebraic error, most participants gave the student full credit, as was the case with Garrison's argument. Such criteria were also used when validating whether the arguments were valid proofs or not. When asked what criteria they used for proof validation, they wished to see the correct base step, inductive step, and algebra. All participants accepted the student arguments, recognizing three components as determinants of the proofs being valid or invalid. However, when given the student argument that addressed only the inductive step, two participants accepted that as a valid proof. Most participants compared the students' arguments to their own work when checking the correctness of the algebraic manipulations in the inductive step. However, some of the participants had difficulties understanding the students' algebraic manipulations and completely disregarded the algebraic manipulations in the proofs or presumed that all of the algebra in the inductive step was correct. Participants who recognized algebraic errors in the inductive step also did not put as much emphasis on the correctness of the algebra when validating students' proofs. Rather, both groups focused on the form of the arguments, whether they included the base case, inductive hypothesis, and inductive step, while validating and grading the proofs, without considering algebraic details. This finding raises questions about how the participants would evaluate student arguments if more algebraic errors were present. Future research should examine whether similar results can be found with other cohorts and how participants respond to student arguments by MI that include more errors in the algebraic manipulations.

References

Australian Curriculum, Assessment and Reporting Authority. (2015). *The senior secondary Australian curriculum: Mathematics V7.5*. Sydney, Australia: Author.

Baker, J. (1996, April). *Students' difficulties with proof by mathematical induction*. Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY.

Borko, H., Peressini, D., Romagnano, L., Knuth, E., Yorker, C., Wooley, C., Hovermill, J., & Masarik, K. (2000). Teacher education does matter: A situative view of learning to teach secondary mathematics. *Educational Psychologist, 35*, 193–206.

California Department of Education. (2013). *California common core state standards: Mathematics*. Retrieved from http://www.cde.ca.gov/be/st/ss/documents/ccssmathstandardaug2013.pdf

Dickerson, D. S. (2006). Aspects of preservice teachers' understandings of the purposes of mathematical proof. *Proceedings of the 28th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Mérida, México: Universidad Pedagógica Nacional.

Harel, G. (2002). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. Campbell & R. Zaskis (Eds.), *Learning and teaching number theory* (pp. 185–212). Westport, CT: Ablex.

Inglis, M., Mejía-Ramos, J. P., Weber, K., & Alcock, L. (2013). On mathematicians' different standards when evaluating elementary proofs. *Topics in Cognitive Science, 5*, 270–282. doi: 10.1111/tops.12019

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

Korean Ministry of Education, Science, and Technology. (2012). *The detailed exposition of mathematics curriculum of high school*. Seoul, South Korea: Author.

Moore, R. C. (2015, February). *When mathematicians grade students' proofs, why don't the scores agree?* Poster presented at the 18th Annual Conference on Research in Undergraduate Mathematics Education, Pittsburgh, PA.

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

Peressini, D., Borko, H., Romagnano, L., Knuth, E., & Willis, C. (2004). A conceptual framework for learning to teach secondary mathematics: A situative perspective. *Educational Studies in Mathematics, 56*, 67–96.

Selden, A., & Selden, J. (2003). Validations of proofs written as texts: Can undergraduate tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 36*, 4–36.

Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* London: Sage.

Stylianides, G., Stylianides, A., & Philippou, G. (2007). Preservice teachers' knowledge of proof by mathematical induction. *Journal of Mathematics Teacher Education*, *10*(3), 145–166.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39,* 431–459.

Weber, K., & Mejia-Ramos, J. P. (2011). Why and how mathematicians read proofs: An exploratory study. *Educational Studies in Mathematics, 76*, 329–344.