Comparing STEM Majors, Practicing and Prospective Secondary Teachers' Feedback on Mathematical Arguments: Towards Validating MKT-Proof

Rebecca Butler
University of New Hampshire

Orly Buchbinder
University of New Hampshire

Sharon McCrone
University of New Hampshire

Mathematical Knowledge for Teaching Proof (MKT-P) has been recognized as an important component of fostering student engagement with mathematical reasoning and proof. This study is one component of a larger study aimed at exploring the nature of MKT-P. The present study examines qualitative differences in feedback given by STEM majors, in-service and pre-service secondary mathematics teachers on hypothetical students' arguments. The results explicate key distinctions in the feedback provided by these groups, indicating that this is a learnable skill. Feedback is cast as a component of MKT-P, making the results of this study significant empirical support for the construct of MKT-P as a type of knowledge that is unique to teachers.

Keywords: Secondary Teachers, Mathematical Knowledge for Teaching Proof, Feedback

Calls for students to experience deeper engagement with mathematical proof and reasoning across a variety of levels have been long standing (Nardi & Knuth, 2017; NGA & CCSSO, 2010; NCTM, 2009, 2014; Stylianides, Stylianides, & Weber, 2017). In order for this aim to be realized, mathematics teachers must be equipped with a substantial knowledge, both subject matter and pedagogical. The present study is part of a larger, NSF-funded research project, which studied how Mathematical Knowledge for Teaching Proof (MKT-P) of prospective secondary teachers (PSTs) evolves as a result of participating in a specially designed capstone course *Mathematical Reasoning and Proving for Secondary Teachers* (Buchbinder & McCrone, 2020). The course design and the assessment instruments draw on the MKT-P theoretical framework proposed by Buchbinder & McCrone (2020), which we describe below. At the same time, the project seeks to understand the nature of (MKT-P) empirically; specifically, whether this knowledge is unique to teachers and if it is learnable. To explore these questions, we compared the performance of STEM majors, in-service teachers, and prospective secondary teachers on an instrument designed to assess MKT-P (Buchbinder & McCrone, 2021). This examination of groups with distinct mathematical and pedagogical backgrounds aims to draw out the ways in which MKT-P is unique to teachers. Distinctions in performance of the practicing teachers with respect to individuals with little teaching experience or extensive theoretical knowledge of teaching, but strong knowledge of proof-specific subject matter, could reveal the components of MKT-P, which are unique to teachers (Krauss, Baumert, & Blum, 2008).

The study reported herein analyzes a particular aspect of the MKT-P - the ability to provide instructional feedback on student mathematical work. A set of items on the MKT-P instrument asked participants to provide hypothetical students with feedback on their mathematical arguments. This feedback was quantitively assessed in the larger study through a coding scheme concerned with correctness and mathematical richness of responses (Buchbinder, McCrone & Capozzoli, Butler, in preparation). As we coded these responses, we noticed significant qualitative differences between the feedback provided by the three groups, which could not be captured by the existing scheme. The current study provides in-depth exploration of the differences that caught our attention in the larger study. The research questions are:

1) How does the feedback to a hypothetical student on mathematical arguments compare across three groups of participants: practicing teachers, STEM majors, and PSTs.

2) How does the feedback of PSTs change before and after participation in the capstone course?

Background and Literature Review

In this study, providing feedback on students' arguments is taken as one practice specific to Mathematical Knowledge for Teaching Proof (MKT-P). Inspired by research on the notion of MKT-P (e.g., Corleis, Schwarz, Kaiser, & Leung, 2008; Lessig, 2016; Lin et al., 2011; Steele & Rogers 2012; Stylianides 2011), Buchbinder and McCrone (2020) proposed an MKT-P framework, which consist of three domains: knowledge of the logical aspects of proof (KLAP), knowledge of content and students (KCS-P), and knowledge of content and teaching (KCT-P). KLAP encompasses subject matter knowledge related to the mathematical content of proof, such as knowledge of definitions, proof types, and proof validity. KCS-P describes knowledge about student perceptions of various proof related concepts including knowledge of common difficulties or misconceptions. KCT-P is knowledge of pedagogical praxis specific to facilitating student engagement with reasoning and proof and encompasses such skills as design of classroom activities, the making of instructional moves, and the ability to provide students with productive feedback on their proof related activities.

Feedback is commonly described as "information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding" (Hattie & Timperley, 2007, p.81). Extensive research across populations and disciplines has explored the ways in which various characteristics of feedback are related to student achievement (e.g., length, timing, positive or negative nature, and format) (Shute, 2008). Recent work in mathematics education builds on these established relationships between feedback and student learning by focusing on feedback as a teaching practice.

Kastberg et al. (2016) found that mathematics PSTs tend to center their feedback on the performance of the task, processes involved in completing the task, and personal evaluations with more attention to correct than incorrect responses. This finding is echoed by Crespo (2002) who found that PSTs tended to focus their feedback on the process and correctness of student responses, with PSTs supplying students with correct solutions in the case of incorrect responses and issuing praise in the case of correct responses. Over the course of a mathematics methods course, Crespo (2002) found that PSTs shifted the focus of their praise toward the process students used to obtain their answers and shifted the focus of their critiques toward probing student processes via questions and offering guidance toward correct solutions. These shifts are also evidenced by Santos & Pinto (2010) who tracked the changes in written feedback given to secondary students by a single teacher over the first two years of her teaching. They found that the focus of her feedback shifted away from the task and toward the student, errors were corrected less often, symbolic feedback was given less often, the nature shifted from stating facts toward leading clues, and length of feedback increased. Bleiler et. al. (2014) analyzed feedback given by PSTs on sample student proof strategies, finding that PSTs tend to critique student work for use of specific cases over general mathematical representations. Overall, the literature indicates that feedback is a practice which teachers develop over time with this development centered on shifts from product to process and from telling to guiding. This supports our assumption that providing feedback on student mathematical arguments is a particular task of teaching specific to classroom situations involving reasoning and proof; in other words, an element of MKT-P. Comparing feedback across groups with comparable mathematical knowledge but different pedagogical background can provide empirical support for this assumption and illuminate the nature of MKT-P.

Methods

Participants and the Study Setting

Three groups of participants were involved in this study; 17 in-service secondary mathematics teachers, 22 university STEM students, and nine PSTs involved in the capstone course *Mathematical Reasoning and Proof for Secondary Teachers*. The in-service mathematics teachers (11 female and 6 male) had teaching experience ranging between 2 to 25 years ($\bar{x}$ = 12.18, *SD* = 8.26). The teachers were recruited via presentations given at departmental meetings and were compensated for their participation with a $35 ex gratia payment.

The STEM majors and PSTs came from the same 4-year university. The STEM students, mostly computer science and mathematics, second year majors, were enrolled in a mandatory mathematical proof course. The course instructors reviewed the MKT-P questionnaire and confirmed that the students should have all the relevant proof knowledge included in it. The students who volunteered to participate in the study received a small extra credit in the course.

All but one of the PSTs enrolled in the capstone course were in their final year of their secondary mathematics education program, so they had completed the majority of their mathematical coursework, including the same Mathematical Proof course as the STEM majors, as well as two courses specific to mathematics pedagogy. While these latter courses did require some classroom observations, the participants had no personal experience with teaching.

The capstone course (taught by the second author of this paper) intended to enhance PSTs' MKT-P by helping the PSTs to connect their university-level knowledge of mathematical reasoning and proof with secondary school teaching (Buchbinder, McCrone, 2018; 2020). The course was comprised of four modules, each devoted to a particular proof theme: (1) direct proof and argument evaluation, (2) conditional statements, (3) role of examples in proving and (4) indirect reasoning. In each module, the PSTs refreshed their subject matter knowledge of the proof theme and subsequently examined common student (mis)conceptions of the topic. Next, they applied this knowledge by planning and teaching a lesson on that proof theme at a local school. They recorded their lessons using 360 video cameras and wrote a reflection report on their teaching (Buchbinder, Brisard, Butler, McCrone, in press).

Instruments

The MKT-P questionnaire contained 29 items designed to evoke participants' knowledge in each of the three domains of MKT-P. Ten questions were devoted to measuring KLAP, eleven to KCS-P, and eight to KCT-P. The content of the questions spanned the four proof-themes of the capstone course (direct proof, conditional statements, role of examples in proving, and indirect reasoning) framed across a range of topics in the secondary curriculum: number and operations, geometry, algebra and functions. The KCT-P items, which are the focus of this study, consisted of two parts: identifying errors in a sample student's argument and providing feedback to the student, with attention to both the strong and weak points of their work (Figure 1). This type of question requires participants to translate their proof-specific mathematical knowledge into the pedagogical practice of communicating with a student in a way that is productive for their learning. Hence these questions are appropriate for assessing KCT-P. By examining KCT-P items qualitatively, we hoped to detail the differences among the groups of participants, which we noticed in the larger study.

13. Mr. Briggs asked his students to prove the following statement: ***The sum of any two rational numbers is a rational number.***

a) **Paula's solution**:
Suppose r and s are rational numbers. Since r is a rational number, it can be written as a fraction. Similarly, s is a rational number, so it can be written as a fraction by definition of rational number. But the sum of two fractions is a fraction, and a rational number is a fraction, it follows that $r + s$ is a rational number. ■

i) Identify errors (if any) in the student's argument. If none, write "no errors".
ii) Provide feedback to the student, highlighting strengths and weaknesses of their argument.

Figure 1: Sample KCT-P item

Data Collection and Analytic Techniques

The same MKT-P instrument was administered to all three groups. The in-service teachers were given two weeks to complete the questionnaire in an on-line format. The STEM majors completed the paper-and-pencil questionnaire near the end of the course. The questionnaire took 1.5 – 2 hours to complete. The PSTs completed the MKT-P questionnaire *twice*; at the beginning of the course (as an ungraded assignment) and at the end of the course as a final exam.

The feedback provided by participants on KCT-P items was first analyzed for mathematical correctness and pedagogical depth, e.g., whether a student receiving this feedback would be able to improve their mathematical work. Through this analysis, detailed in Buchbinder et. al. (in preparation), we noticed stark differences in the three groups' feedback, which our coding scheme was not capturing. Consequently, we re-coded all data using a newly devised a coding scheme, along four dimensions. First, we counted the sheer *amount of words* in the feedback. Second, we analyzed the *perspective* - to whom the participants addressed their responses. Addressing a student directly (e.g., "you are on the right track") was coded as 1, using the collective "we" (e.g., "think what we are trying to show") was coded as 2, and referring to the student in the third person, was coded as 3. The third dimension was the presence of *questions* in the feedback. Lastly, since participants were asked to highlight both positive and negative aspects of the student's work, we analyzed the content of the feedback for complimentary and critical feedback. We used open coding and the constant comparative method (Strauss & Corbin, 1994) to identify recurring themes (Table 1). The unit of analysis was the appearance of a code in a given response, so occasionally multiple codes were assigned to a given response. Two researchers coded the data individually with initial Kappa agreement of 0.67 (moderate) for compliments and 0.53 (weak) for critiques. All disagreements were discussed and reconciled.

Table 1. Coding scheme for the content of the feedback

Code and Description	Examples
Empty Providing feedback which is devoid of substance beyond recognition of correctness	*Compliment*: Good proof *Critique*: Not proven enough
General Mathematical Comment A general statement about the mathematics of the student's work or describing work in a neutral manner	*Compliment*: Strength: showed multiple examples showing its true *Critique*: An example does not prove so a counterexample is needed.
Student Mathematical Work An explicit connection to the student's work is made in the feedback	*Compliment*: Molly had a great idea to show how the fractions add together to make a new fraction and used the closure property to argue details about the integers.

	Critique: Nicole only uses one example to argue that the statement is true.
Student Understanding Feedback gauges what the student knows or does not know	*Compliment*: You show your strong understanding of what a rational number is and how to use variables to generalize a situation. *Critique*: Calvin lacks an understanding in conjecture which is good to know for the teacher.
Mathematical Value Centered on some mathematical value (brevity/clarity), mathematical writing, or argument structure	*Compliment*: I like that the student attempted to use algebra in a general case to prove it *Critique*: Try to write it more in the 3rd person
Directing Solution The participant tells the student how to correct their work, either through suggestion or direct explanation	*Compliment*: N/A *Critique*: Nicole needs to use variables to show that the statement holds true for all rational numbers

Results

Although the *number of words* in the feedback does not give insight to its substance, it gives a broad sense of the participant's willingness to take seriously the hypothetical student and the mathematics of the problem. The groups showed notable differences in the mean number of words, with teachers averaging 37.43 words per response and STEM majors averaging 27.18 words per response. The PSTs'-Pre[1] feedback was initially closer to the length of the STEM majors', averaging 29.33 words per response; however, PSTs-Post surpassed the teachers averaging 48.69 words per response. This gain could be attributed to the high-stakes nature of the post-questionnaire, but it is nevertheless a valuable indication of engagement.

The groups also differed with respect to *perspective* in which they framed their responses. The teachers had an average perspective rating of 1.87 while the STEM majors had an average rating of 2.29. This means that teachers more often framed their feedback as an address to the student while STEM majors tended to talk about the student or to the collective. The PSTs-Pre aligned with the STEM majors, with an average rating of 2.54, which slightly lowered upon their second attempt to an average of 2.17 (for PSTs-Post). This shift indicates that PSTs began to address students directly in their feedback more often, but still less frequently as the teachers.

Participant's responses were also analyzed for the presence of questions, although these only appeared in about 25% of the data: across all groups, only 74 responses contained some form of question. Of these, 13 (17.6%) were generated by STEM majors and 37 (50%) by teachers, meaning that teachers tended to pose questions to the hypothetical students more frequently than the STEM majors did. The PSTs-Pre posed 9 questions as a group, and 15 questions on the Post, representing 12.2% and 20.2% of all questions in the data set respectively. Yet again, the PSTs-Pre performance on the questionnaire resembles that of the STEM majors, while the PSTs-Post performance shifted toward the teachers' performance.

Whenever possible, the content of the feedback was analyzed for the presence of compliments and /or critiques on student arguments. Complimentary feedback appeared in 62% of STEM majors' responses, 64% of teachers' responses, 73% of the PSTs'-Pre, and 91% of the PSTs'-Post responses. Figure 2 shows the frequency of each *compliment* type (detailed in Table 1) as a proportion of the complimentary feedback from each group to facilitate cross-group comparison.

[1] For ease of communication, we refer to the pre-course data as "PSTs-Pre" and the post-course data as "PSTs-Post"

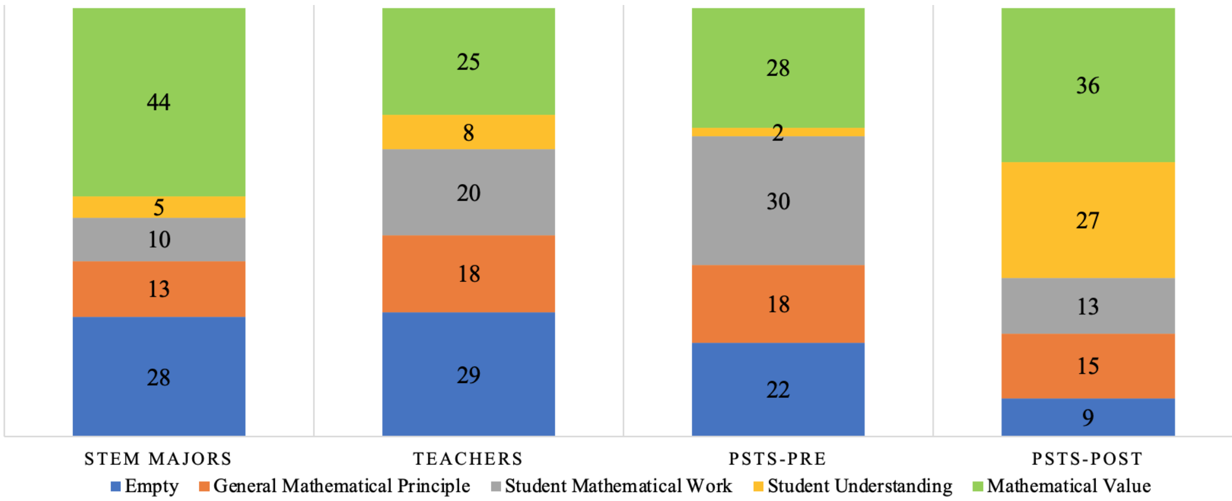

Figure 2: Distribution of compliment types by group (in percent of total complimentary feedback in each group)

Figure 2 shows that the only category for which all groups displayed similar proportions (within 5% of one another) is that of general mathematical principles. Among the notable differences between the groups is the drop in empty compliments of PSTs-Pre to PSTs-Post. While all other groups had similar proportions of empty praises (within 7%), PSTs-Post gave such feedback less frequently, with only 9% of their compliments considered empty. PSTs-Post also showed a large drop (17%) in complimenting student mathematical work, while dramatically increasing their complements of student understanding (2% to 27%). Both teachers and STEM majors showed low rates of such feedback (5% and 8% respectively) making this category a distinguishing feature of the PSTs'-Post complimentary feedback. PSTs also slightly increased their compliments of mathematical values from 28% to 36%. While PSTs-Pre were closer to the teachers on this category (25%), the PSTs-Post fell between the teachers and the STEM majors (44%).

Feedback of a *critical nature* was more frequent than complimentary feedback for both STEM majors and teachers, appearing in 86% of STEM majors' responses, 84% of teacher responses. 68% of PSTs-Pre and 89% of PSTs-Post feedback contained some sort of critique.

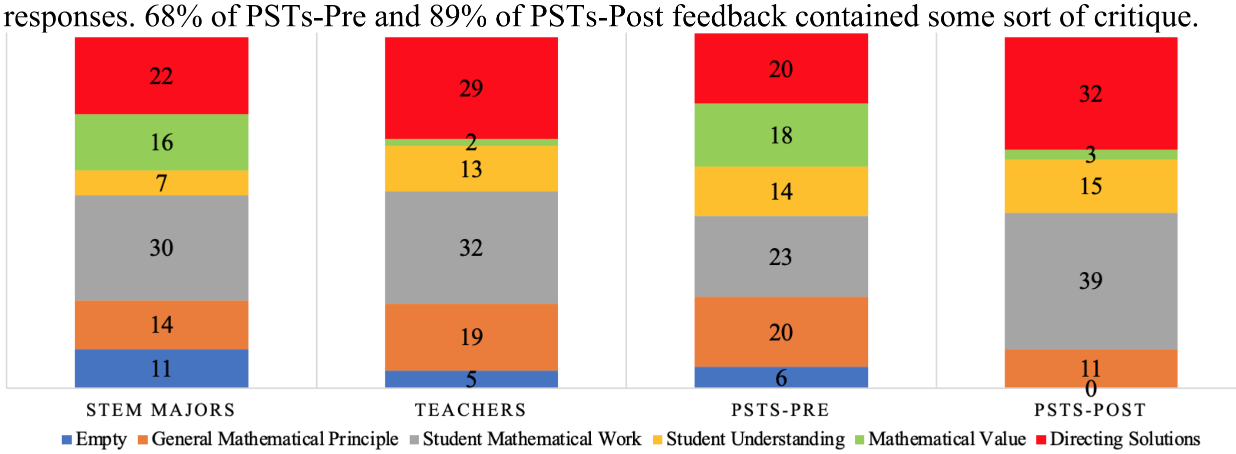

Figure 3: Distribution of critique types by group (in percent out of total critical feedback in each group)

Figure 3 shows the distribution of critical feedback categories for each group in percent of total critical feedback in each group. Compared to Figure 2, each group provided empty critiques less often than empty compliments, with such critiques being entirely absent for PSTs-Post. This indicates that the critiques provided by all groups are substantive and contain more information regarding the nuances of student's arguments. The types of critiques vary by group. Notably, PSTs increased their critiques of student mathematical work and directing solutions. PSTs-Post decreased their focus on general mathematical principles and mathematical values compared to PSTs-Pre. This indicates shifts toward specific aspects of existing student work and revisionary work rather than making general comments. The increase in directing solutions and decrease in mathematical values mark a shift toward the critique profile of the teachers while the decrease in general mathematical principles marks a shift toward the critique profile of the STEM majors.

Discussion

The two research questions of this study asked about qualitative differences in the feedback given by the groups and changes in PSTs feedback before and after the capstone course. Significant differences were found through these comparisons. Teachers tended to provide longer feedback, address the student directly, and frequently posed questions. STEM majors tended to write shorter responses in the third person. Content wise, teachers' feedback tended to focus on the particularities of student work while STEM majors focused on the mathematical conventions and qualities of mathematical writing. Considered holistically, the feedback provided by the PSTs-Pre is qualitatively close to that of the STEM majors, while PSTs-Post feedback is qualitatively close to that of the teachers. This is primarily evident in the shift toward longer feedback addressed to the student and inclusion of questions. It is also evident in the shift of critical feedback toward *directing solutions* and away from critiquing *mathematical values*.

These shifts in the PSTs feedback are consistent with the existing literature exploring PST feedback. PSTs in this study tended to provide complimentary feedback frequently, concurring with Kastberg et al. (2016) and Crespo (2002) that PSTs tend to focus on the correctness of responses. This study also echoes the trend in which PSTs write longer feedback focused on processes as they gain experience teaching (Santos & Pinto, 2010). The literature is furthered by this study as it compares the work of PSTs to both experienced teachers and STEM majors, revealing aspects of feedback specific to teachers and providing insight to their development.

While limited by small sample sized and the high-stakes nature of the PSTs post assessment, the results of this study elucidate MKT-P as a construct; namely, these qualitative distinctions contribute to the empirical validation of this theoretical construct. By drawing out the differences in feedback provided by groups with varying mathematical and pedagogical backgrounds, this study illustrates that there is an underlying difference in the MKT-P of these groups, thereby lending credibility to the construct as something which is both unique to teachers and learnable.

This connection of feedback to MKT-P also carries implications for teacher education. While the course that the PSTs of this study were enrolled in was not specifically aimed at cultivating specific feedback practices, changes in PST feedback were evidenced after taking this course. These changes along with the substantial differences in teachers' and STEM majors' feedback indicate that giving feedback is a learnable practice. By consciously attending to the development of MKT-P in PSTs, teacher preparation programs can better equip their participants to evaluate student work regarding proof and communicate their thoughts about this work back to the student in meaningful ways.

Acknowledgments

This research was supported by the National Science Foundation, Award No. 1711163. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Bleiler, S. K., Thompson, D. R., & Krajčevski, M. (2014). Providing written feedback on students' mathematical arguments: Proof validations of prospective secondary mathematics teachers. *Journal of Mathematics Teacher Education*, 17(2), 105-127.

Buchbinder, O., Brisard, S., Butler, R., McCrone, S. (2021) Preservice Secondary Mathematics Teachers' Reflective Noticing from 360-degree Video Recordings of Their Own Teaching. Manuscript submitted for publication.

Buchbinder, O., & McCrone, S. (2021). Mathematical Reasoning and Proving for Secondary Teachers.

Buchbinder O., McCrone S., Capozzoli M., & Butler, R. (2021) Mathematical Knowledge for Teaching Proof: Comparing Secondary Teachers, Preservice Secondary Teachers, and Undergraduate STEM Majors. Manuscript in preparation.

Buchbinder, O., & McCrone, S. (2020). Preservice teachers learning to teach proof through classroom implementation: Successes and challenges. *The Journal of Mathematical Behavior*, *58*, 100779.

Buchbinder, O., & McCrone, S. (2018). Mathematical Reasoning and Proving for Prospective Secondary Teachers. In *Proceedings of the 21st Annual Conference of the Research in Undergraduate Mathematics Education, Special Interest Group of the Mathematical Association of America*.

Corleis, A., Schwarz, B., Kaiser, G., & Leung, I. K. (2008). Content and pedagogical content knowledge in argumentation and proof of future teachers: a comparative case study in Germany and Hong Kong. *ZDM Mathematics Education*, *40*(5), 813-832. doi.org/10.1007/s11858-008-0149-1.

Crespo, S. (2002). Praising and correcting: Prospective teachers investigate their teacherly talk. *Teaching and teacher education*, 18, 739-758.

Kastberg, S. E., Lischka, A. E., & Hillman, S. L. (2016). Exploring Prospective Teachers' Written Feedback on Mathematics Tasks. *North American Chapter of the International Group for the Psychology of Mathematics Education*.

Krauss, S., Baumert, J., & Blum, W. (2008). Secondary mathematics teachers' pedagogical content knowledge and content knowledge: Validation of the COACTIV constructs. *ZDM – Mathematics Education, 40*(5), 873-892.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112.

Lesseig, K. (2016). Investigating mathematical knowledge for teaching proof in professional development. *International Journal of Research in Education and Science, 2*(2), 253-270.

Lin, F. L., Yang, K. L., Lo, J. J., Tsamir, P., Tirosh, D., & Stylianides, G. (2011). Teachers' professional learning of teaching proof and proving. In *Proof and proving in mathematics education* (pp. 327 – 346). Springer, Dordrecht.

Nardi, E., & Knuth, E. (2017). Changing classroom culture, curricula, and instruction for proof and proving: How amenable to scaling up, practicable for curricular integration, and capable of producing long-lasting effects are current interventions? *Educational Studies in Mathematics*, *96*(2), 267-274. DOI: 10.1007/s10649-017-9785-0

National Council of Teachers of Mathematics (2009). *Focus in High School Mathematics: Reasoning and Sense Making.* Reston, VA: NCTM.

National Council of Teachers of Mathematics (2014). *Principles to Actions: Ensuring Mathematical Success for All*. Reston, VA: NCTM.

National Governors' Association Center for Best Practices, Council of Chief State School Officers (2010). *Common Core State Standards: Mathematics.* Washington DC: Council of Chief State School Officers.

Santos, L., & Pinto, J. (2010). The evolution of feedback practice of a mathematics teacher. In *34th Conference of the International Group for the Psychology of Mathematics Education* (pp. 142-152). PME.

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, 78(1), 153-189.

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. DOI: 10.1007/s10857-012-9204-5

Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 273 – 285). Thousand Oaks: Sage Publications.

Stylianides, A. J. (2011). Towards a comprehensive knowledge package for teaching proof: A focus on the misconception that empirical arguments are proofs. *Pythagoras, 32*(1), 10 pages.

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.