



ISSN: 2148-9955

International Journal of Research in Education and Science (IJRES)

www.ijres.net

Exploring Prospective Teachers' Ability to Generate and Analyze Evidence-based Explanatory Arguments

Marta T. Magiera¹, Vecihi S. Zambak²

¹Marquette University, United States of America

²Monmouth University, United States of America

To cite this article:

Magiera, M.T. & Zambak, V.S. (2020). Exploring prospective teachers' ability to generate and analyze evidence-based explanatory arguments. *International Journal of Research in Education and Science (IJRES)*, 6(2), 327-346.

The International Journal of Research in Education and Science (IJRES) is a peer-reviewed scholarly online journal. This article may be used for research, teaching, and private study purposes. Authors alone are responsible for the contents of their articles. The journal owns the copyright of the articles. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material. All authors are requested to disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations regarding the submitted work.

Exploring Prospective Teachers' Ability to Generate and Analyze Evidence-based Explanatory Arguments

Marta T. Magiera, Vecihi S. Zambak

Article Info

Article History

Received:
24 August 2019

Accepted:
16 December 2019

Keywords

Mathematical explanation
Mathematical reasoning
Argumentation
K-8 pre-service teachers

Abstract

In this paper, using written responses of 37 PSTs preparing to teach grades 1-8 mathematics, we examined explanations they constructed to support their problem solutions and explanations they provided in support of their critiques of student-generated explanations. We also examined features of explanations on which PSTs drew in their critiques of mathematical explanations of students. Our results draw attention to the importance of helping PSTs develop competencies in constructing and critiquing mathematical explanations concurrently. Although explanations PSTs generated for their critiques of student explanations were weaker compared to the explanations PSTs formulated for their own problem solutions, PSTs proficient in generating mathematical explanations were also more proficient in analyzing and critiquing mathematical explanations. We identified seven criteria PSTs used while analyzing and critiquing student-explanations. These criteria reveal what PSTs might value, or pay attention to, as they critique student-explanations. We share implications for mathematics teacher educators to consider and suggest directions for further research.

Introduction

Although not unanimously defined in the mathematics education literature, explanations and justifications are frequently viewed as connected practices through which students should engage with mathematics content (Australian Curriculum and Assessment Reporting Authority [ACARA], 2015; National Council of Teachers of Mathematics, [NCTM], 2000; National Governors Association Center for Best Practices & Council of Chief State School Officers [NGA & CCSSO], 2010). Included in curricular documents are expectations that students “explain and justify their thinking and learn how to detect fallacies and critique others’ thinking” (NCTM, 2000, p. 188). These practices promote conceptual understanding because they serve as a vehicle for representing internal processes of reasoning and making sense of mathematics (Cross, 2009; Francisco, 2013).

Creating explanations that go beyond simple descriptions of steps and procedures allows students to clarify aspects of mathematical thinking that might not be readily apparent to others, providing them with an opportunity to articulate and revise their mental models (Chi, 2000). Prediger and Erath (2014) shared that formulating explanations facilitates “building and connecting knowledge in a systematic, structured way by linking an explanandum (i.e., the issue that needs to be explained) to an explanans (i.e., by which the issue is explained).” Others argued that engaging students from early grades on in producing and evaluating mathematical explanations helps them develop the kind of mathematical thinking that supports the ability to construct proofs (e.g., Bicknell, 1999) and reason about proofs (e.g., Hodds, Alcock, & Inglis, 2014).

Generating and critiquing mathematical explanations are complex practices that are not easy for students. They require that a student connects their evidence to the relevant information and effectively communicates their understanding using evidence to evaluate and revise claims. Focusing on mathematical explanations in the context of pre-service teachers’ (PSTs’) education is warranted. Both competencies (providing and critiquing explanations) help PSTs to advance the mathematical agenda within their classrooms. Teachers who develop these competencies are positioned to guide their students’ mathematical development, support and scaffold student learning, and promote the growth of students’ reasoning and mathematical understanding (Hoover, Mosvold, Ball, & Lai, 2016; McClain, 2009).

Mathematical Explanations as Arguments

In the mathematics education literature, the terms *mathematical explanation* and *argument* are not clearly defined. The relationship between explanations and arguments is also not clear. Some researchers interpret explanations and arguments as two separate entities delineated by their purpose (e.g., Yackel, 2004). They portray explaining as giving a *description* of an aspect of mathematical thinking, steps taken to solve a problem, or a mathematical statement. They portray constructing an argument as giving *reasons* that justify a claim, a mathematical procedure, result, or respond to challenges. Others distinguish between *descriptive* explanations, ones that answer the *how* questions, *interpretative* explanations, ones that clarify the statement to bring up its central meaning, and thus address the *what* questions, and *reason-giving* explanations which supply answers to the *why* questions (e.g., Hafner & Mancosu, 2005). Yet, others share the notion of explanations as arguments that provide meaning *and* reasons for that meaning (e.g., Balacheff, 2010; Evens & Houssart, 2004; Esmonde, 2009; Hanna, 2000; Krummheuer, 1995; Levenson, 2013; Levenson & Barkai, 2013; Morselli & Levenson, 2014; Prediger & Erath, 2014; Stacey & Vincent, 2009). They conceptualize explanations as a way of communicating meaning *and* providing grounds, evidence, and reasons for *why* a statement being explained is true, or not. In the process of explaining, one describes one's thinking, result, insight, *and* gives evidence and reasons to justify a plausibility of a process, result, or strategy.

Mathematical explanations are often conceptualized as arguments in the K-8 curricular materials, as exemplified by the use of the phrase "*explain and justify*." Within the K-8 curricular materials, this phrase is frequently used as a request for students to generate mathematical arguments (Bieda, Ji, Drwencke, & Pickard, 2014; Dolev & Even; 2015; Stacey & Vincent, 2009). This phrase choice might be deliberate to make explicit for the students that they need to move beyond providing merely procedural or interpretative explanations towards reason-giving explanations. It might also be used deliberately as a reference to a broad range of arguments and ways of reasoning, and to distinguish arguments that do not qualify as proofs from arguments that could be accepted as proofs. For example, in their survey of the 7th-grade Israeli mathematics textbooks, Dolev and Even (2015) identified that curricular tasks that intend to engage students in mathematical argumentation asked students to explain and justify their work or a mathematical claim stated in the textbook. Bieda et al. (2014) made a similar observation analyzing a collection of upper-elementary mathematics textbooks used in the U.S. Bieda and colleagues identified that curricular materials that foster student engagement in generating mathematical arguments (both proof-like and not) prompted for explaining and justifying why a claim, conjecture, or statement is true.

The notion of mathematical explanations as arguments is also present in the literature that addresses teaching strategies and teacher actions that contribute to engaging students in learning mathematics content with a focus on mathematical argumentation (e.g., Choppin, 2007; Esmonde, 2009). Choppin, for example, discussed the complexity of teacher actions that support students' mathematical argumentation in class discussions. He shared that while teachers make in-the-moment decisions to purposefully support students' learning, they need to decide which "students need to elaborate on their explanations in order to determine warrants in their claims," consider the "quality of the evidence presented for each [student-generated] explanation," or consider "the mathematical ideas embedded in each individual explanation" (p. 309).

In our work with PSTs, we conceptualize mathematical explanations as arguments that communicate meaning drawing on evidence and reasons (as opposed to providing a description of a procedure). In this paper, we discuss mathematical explanations in the context of problem-solving and use the term *explanation* to represent an evidence-based argument for a solution to a problem; one that provides meaning *and* justifies the solution. We also use the term explanation to represent evidence-based argument, which provides meaning and justifies one's critique of problem solution in the context of analyzing a student-generated explanation. Our interpretation of mathematical explanation in the context of problem-solving is consistent with descriptions of Chi, Bassok, Lewis, Reimann, and Glaser, (1989) who saw an explanation of a problem solution as a statement which either overtly or covertly articulates meaning expounding *why* a strategy or result makes sense.

Relevant Research on PSTs' Argumentation Skills

Prior research in the area of mathematical argumentation includes studies conducted primarily within the proof-related research paradigm. This body of research largely explored PSTs' conceptions of mathematical argumentation and their argumentation skills from the perspective of proof (e.g., Boyle, Bleiler, Yee, & Ko, 2015; Felton, 2007; Martin & Harel, 1989; Morris, 2007; Stylianides G. & Stylianides A, 2009). Some researchers studied PSTs' proof-related knowledge with a focus on specific proof methods (e.g., Stylianides,

Stylianides, & Phillippou, 2007). Others focused on pedagogical content knowledge exploring teacher actions and strategies that facilitate student engagement with proof-related tasks (e.g., Bostic, 2016) or PSTs' ability to evaluate student-generated arguments with a focus on proofs and proving (e.g., Morris, 2007). Collectively this body of research generates important insights into our understanding of mathematical argumentation, specifically as it relates to proofs in the context of PSTs' preparation, but it also raises questions.

First, building a comprehensive understanding of PSTs' argumentation skills requires expanding the research focus to include a broader range of arguments, such as earlier discussed evidence-based explanations. Stylianides A. and Stylianides G. (2009) argued for research that attends to different tasks and situations in the context of which PSTs' ability to write and critique a broad range of mathematical arguments is explored. Expanding the research focus beyond the proof-related research paradigm allows for generating a more ubiquitous understanding of mathematical argumentation in the context of PSTs' education.

Second, past research that addressed PSTs' competencies in writing and evaluating proofs brings attention to and raises questions about the relationship between PSTs' abilities to generate and evaluate mathematical arguments. By their design, past studies that explored PSTs' conceptions of proof did so by asking PSTs to either write or evaluate arguments from the perspective of proof. For example, having PSTs analyze researcher-provided arguments, Martin and Harel (1998) documented that PSTs' conceptions of proofs were influenced by an argument form or the mode of argumentation. The participants in their study frequently identified proofs based on their perceived argument sophistication or conventions they viewed as associated with proofs. Stylianides A. and Stylianides G. (2009) explored how elementary PSTs make sense of proofs as they first construct them on their own, and then evaluate their proofs. Their research revealed that most of their participants were able to provide an accurate assessment of the arguments they generated. Even if they failed to generate a proof, they were able to recognize and correctly assess the limitations of their arguments. Working with PSTs preparing to teach elementary, middle, or high school mathematics, Felton (2007) asked them to define a proof and then use their definition to analyze and identify proofs in a sample of student-generated arguments. His study revealed vast differences between PSTs' conceptions of proof (as identified in their definitions), and how they operationalized their conceptions (as identified in their analyses of student-generated arguments to identify proofs). While these studies generate an understanding of PSTs' conceptions of arguments that prove and their ability to analyze and assess proofs, researchers did not explore a possible relationship between PSTs' ability to generate and their ability to assess proofs (or broader mathematical arguments that might not necessarily be accepted as proofs). PSTs' broader argumentation skills need then to be explored with concurrent attention to their argument-construction *and* argument-evaluation skills and a possible relationship between the two.

Finally, building a comprehensive understanding of PSTs' argumentation skills requires also insights into the characteristics of student arguments on which PSTs draw and use as evidence in their assessments. To date, research that examined criteria PSTs use in their assessment of student-generated arguments is scarce. Nardi, Biza, and Zacheriades (2012) studied how *practicing* teachers evaluate students' written responses in a broader context of problem-solving and examined arguments that teachers made about students' solutions. They reported that teachers did not evaluate student arguments on strictly mathematical grounds. Teachers' judgments and evaluations were frequently influenced by personal, pedagogical, professional, or curricular views that teachers used to rationalize their assessments. While Nardi and colleagues' results describe practicing secondary school teachers, Morris (2007) shared similar results reporting on a variety of criteria elementary school PSTs use while asked to assess students' arguments. Unlike Nardi and colleagues who engaged teachers in assessing students' written arguments, Morris engaged PSTs in analyzing and evaluating arguments elementary school students generated in class discussions. PSTs used discussion transcripts, which captured how students collectively proved pattern generalizations. Morris identified a wide range of, often not-mathematically based, criteria PSTs used to judge students' arguments in classroom discussions; they frequently judged student arguments based on their perception of the level of student's understanding or engagement (e.g., whether a student *experimented*, *defended* the presented solution). Little is known about criteria PSTs might use while asked to critique students' written arguments.

Research Questions

As suggested by prior research, building a comprehensive understanding of PSTs' argumentation skills requires concurrent attention to PSTs' argument-construction *and* argument-critique skills and understanding of a possible relationship between these skills. It also requires an understanding of the features of student arguments on which PSTs draw (and which they use as evidence) as they critique student arguments. Our work extends

current research on PSTs' argumentation skills by focusing on their explanations in the context of problem-solving. We selected problem-solving as a context for our research because problem-solving constitutes an important aspect of the elementary school mathematics curriculum. Francisco and Maher (2005) emphasized that problem-solving can provide a context for engaging K-8 students in explaining and justifying their reasoning. In the context of problem-solving, students can learn the practice of explaining and convincing others of the validity of their ideas and develop skills of critically examining and evaluating arguments and ideas shared by their peers. Our decision of situating this work in the context of problem-solving was also motivated by opportunities we saw in bringing PSTs' attention to problem-solving as a context for engaging in mathematical argumentation, specifically in the elementary grades. To this end, situating our research in a semester-long problem-solving course, we explored (1) mathematical explanations PSTs provide to support their problem solutions (i.e., arguments they generate in support of their solution to a problem); (2) explanations they provide in support of their critique of student-generated explanations (i.e., arguments they generate in support of their assessment of explanations they analyze and critique); and (3) the specific features of student-generated explanations PSTs emphasize as they critique student explanations. The specific research questions were:

1. How does the quality of explanations PSTs formulate to support their problem solutions compare and relate to the quality of explanations they provide to support their critique of student-generated explanations?
2. What specific criteria do PSTs use as they critique student-generated written explanations?

Assessing the Quality of Mathematical Explanations

Mathematics education researchers frequently draw on Toulmin's (1958/2003) model to study a broad range of mathematical arguments in different contexts: group interactions (e.g., Krummheuer, 1995; Wagner, Smith, Conner, Singletary, & Francisco, 2014), individuals' written (e.g., Evens & Houssart, 2004) or verbal arguments (e.g., Knipping, 2008), and development of an argument from informal towards a valid proof (e.g., Stylianides G. & Stylianides A., 2009; Weber & Alcock, 2005). Toulmin described six interrelated components of an argument: a claim, data, warrant, backing, modal qualifier, and rebuttal. The claim (conclusion) is the assertion made about an issue. Data embrace the relevant to the claim evidence that provides the foundation for the claim. The warrant and the backing justify the connection between the data and the claim by providing a rationale for why the data support the claim; the backing serves as additional support for the warrant. The modal qualifier expresses the degree of confidence. And finally, the rebuttal includes a rejection of a claim providing support for a counter-argument. Given that not all six components are consistently present in every argument, Krummheuer (1995) argued that the reduced model: claim, data, and warrant together with backing (if present) could be effectively used to analyze student-generated arguments.

The reduced Toulmin's model has been frequently employed in mathematics education research. However, many researchers recognize the model's limitations. One of the challenges is that the model does not account for one's conceptual understanding of a situation, a phenomenon of interest that provides a context for an argument (Pedemonte & Balacheff, 2016; Yopp, 2015). To generate or to evaluate mathematical arguments, one must understand the context. For example, Pedemonte and Balacheff (2016) argued that students' conceptions (which they defined as situated mental constructs, an understanding of a situation or problem context) strongly impact the activity of argumentation. Yopp (2015) asserted that one's conceptual insights might impact how one derives at a conclusion, generates the data, and uses the data as a support for the claim. Pedemonte and Balacheff, and Yopp proposed to enhance Toulmin's model to account for the knowledge system at the basis of argumentation.

As previously discussed, we interpret explanations in the problem-solving context as evidence-based arguments that give meaning and justify problem solution (or a critique of a problem solution, in case of explanations PSTs generated for their assessment of student explanations). We also use the reduced Toulmin's model as a guide for our analysis of PSTs' explanations. Like Pedemonte and Balacheff (2016) and Yopp (2015), we augment the model to account for PSTs' understanding (of the problem situation, articulated evidence, and solution strategy) from which the written explanation stems, as revealed in the problem explanation or explanation critique. Accordingly, we define the quality of mathematical explanation in both contexts (i.e., generating an explanation for a problem solution and generating explanations for one's critique of explanation given by a student) in terms of the quality of four interrelated components that collectively contribute to its strength: a *conclusion*, *supporting evidence* (data), *articulated reasoning* (warrants and backing), and *conceptual reference* that provides the basis for explanation. We further examined the existing mathematical education literature (e.g., Baner, López, Skubal, & Perfecto, 2017; Forman, McCormick, & Donato, 1997; Kline & Ishii, 2008; Lepak,

2014; Reid & Zack, 2009; Wall, Selmer, & Bingham Brown, 2016) for descriptions of the quality of each of the four components. Our operational definitions of the four components contributing to the overall strength of mathematical explanation, derived from the surveyed literature, are summarized in Table 1.

Table 1. Components Contributing to the Strength of Mathematical Explanation

Explanation Component	Generating Explanation for a Problem Solution		Generating Explanation in Support of an Assessment of Analyzed Explanation	
	Description	Operational definition of component quality	Description	Operational definition of component quality
Conclusion (C)	Final and intermediate results	Explanation for a problem solution includes clearly articulated and correct final and intermediate results	Statement that provides an assessment of analyzed explanation	Explanation critique includes well-articulated and valid statements about analyzed explanation
Supporting Evidence (SE)	Facts (stated or derived) that provide the foundation for the conclusion	Explanation for a problem solution includes comprehensive, correct, and relevant evidence in support for the conclusion. The evidence supports all cases within a given problem situation, and it is transparently articulated	Observations, statements identified within analyzed explanation used in support of ones' assessment of analyzed explanation	Explanation critique is supported by comprehensive, relevant, and valid evidence. The evidence is transparently articulated
Reasoning (R)	Links articulated within a problem explanation which show <i>why</i> the evidence supports the claim	Explanation for a problem solution provides logical and plausible connections which validly justify why the evidence supports the conclusion	Links articulated within an explanation analysis that show why the evidence supports the explanation assessment	Explanation critique shows logical and specific connections between discussed evidence and the assessment of analyzed explanation
Conceptual Reference (CR)	Demonstrated understanding that provides the basis for problem explanation	Explanation for a problem solution documents one's understanding of the problem and embedded concepts and relationships	Demonstrated understanding of reasoning on which the analyzed explanation rests	Explanation critique documents one's understanding of strategies and reasoning that provide the basis for analyzed explanation

Method

Participants and Study Context

This research was conducted at a large, private university in the Midwestern United States. In this paper, we report on data collected in the mathematics course for elementary and middle grades education majors. Participants ($n = 37$) were grades 1-8 teaching license candidates enrolled in two sections of that course. Both sections of the course were taught by the same instructor.

The one-semester course—Problem Solving and Reasoning for Teachers—was designed to support PSTs' understanding of mathematical argumentation, reasoning, and proof in the context of school mathematics. The

75 minutes long class sessions were scheduled twice a week for 14 weeks. As a part of the larger project, we collected information about class instruction (e.g., video-records of class sessions, field notes, instructional materials, PSTs' work samples). Below, drawing on our semester-long observations, we describe the course and the major instructional activities in which PSTs were engaged.

Throughout the semester, in the context of solving a large variety of problems, PSTs were engaged in generating and critiquing mathematical arguments. To provide them with language for class discussions and to help them understand the class expectations, at the beginning of the semester PSTs were introduced to the Toulmin's (1958/2003). In the context of the initial class activities, they analyzed several samples of written problem explanations and explicitly discussed (embedded in those explanations) claims, evidence, and reasons. They also discussed the quality and sufficiency of evidence and reasoning identified within the explanations they analyzed and critiqued.

In a typical class, PSTs spent about 30-40% of class time solving a variety of problems for which they generated explanations. They spent about the same amount of typical class time on activities that engaged them in evaluating and critiquing mathematical explanations shared in class, including analyzing the instructor-provided samples of written explanations of middle school students. In small groups, they discussed and shared their explanations for homework problems or class problems, which then individual PSTs presented and offered for class discussions. The problems addressed a wide range of mathematical topics providing PSTs with an opportunity to explore various strategies and ways of mathematical reasoning.

In the context of each problem, PSTs were explicitly asked to support claims they generated for each problem with evidence and reasons. The goal of class discussions and presentations was twofold: (a) to heighten PSTs' awareness of reasoning of others in the context of the problems they solved, and (b) help them see how different ways of thinking about a problem generate different evidence and reasons in support of the given problem solution. Class discussions explicitly addressed explanation quality: e.g., What makes this explanation relevant or effective? What evidence is provided to support the given assertion? Is the evidence credible? Does it stand for any challenge? Does it present an unbroken chain of reasoning? Is it generalizable beyond specific examples or problem situations? Does this explanation provide the meaning for mathematics? Does it have any gaps and holes? The purpose of these activities was to strengthen PSTs' ability to write and critique mathematical explanations with a focus on claims, supporting evidence, and sufficiency and validity of reasoning that links the evidence to identified claims.

Data Sources

Data for this study comes from PSTs' written responses to six parallel tasks for which they had to develop problem solutions and provide explanations (DME tasks), and analyze and provide explanations for their critique of mathematical explanations generated by students (AME tasks). Both types of tasks were designed to facilitate generating claims, providing evidence, justifying claims, and thinking about generality. The DME tasks were typical of those found in the elementary or middle school mathematics textbooks. The AME tasks (selected from the existing literature or instructor's own resources) required the PSTs to explain their critique of analyzed sample explanations generated by elementary or middle school students.

For both types of tasks, PSTs were explicitly asked to provide support for their response. For the DME tasks, they were asked to explain and justify their solutions by giving support for why their strategy and their results are correct. The AME tasks prompted PSTs to (a) examine and describe student strategies and reasoning, (b) critique provided explanation(s) and provide support for their critique using evidence from the analyzed work, and (c) suggest revisions that would enhance the strength of the analyzed explanation. The AME tasks included between one and four student responses to expose PSTs to a variation of claims, evidence, and reasoning within student-generated explanations. For the AME tasks with more than one student-generated explanation, we also asked PSTs to rank the set of explanations providing a specific rationale for their ranking.

We presented these tasks in the context of practice-based pedagogical situations, engaging PSTs in "making the case" for their solution or their critique of analyzed explanation. Task examples are presented in Figure 1 and Appendix A. As discussed earlier, our goal was to compare the quality of mathematical explanations PSTs produced in support of their problem solutions and the quality of explanations they generated for their assessment of student explanations. For consistency of comparisons, we then situated all tasks in the same mathematical domain that addressed thinking about fractions, percents, and proportions.

(1) **Sweater:** A clothing store bought a sweater for a certain price and marked it up 70%. The sweater did not sell, so the storeowner took 25% off the marked-up price and it sold at that price. What profit did the store make?

(2) **River Barge:** Below, is a River Barge problem from Sasha's homework: A river barge can transport 1800 small containers or 1500 large containers. The barge is already loaded with 400 small containers. How many large containers could be still loaded?

This is how Sasha reasoned about this problem:

$$\frac{400}{1800} = \frac{4}{18} = \frac{2}{9}$$

So $\frac{2}{9}$ of the barge is already full now. I can find how many large containers can fit

$$\frac{1800}{1500} = \frac{400}{x}$$

$$\frac{18}{15} = \frac{400}{x}$$

$$\frac{6}{5} = \frac{400}{x}$$

$$1.2 = \frac{400}{x}$$

$$x = \frac{400}{1.2} = 333.333$$

$$x=333$$

Figure 1. Sample DME (1) and AME (2) Tasks (see Appendix A for Other Examples)

Data Analysis

Research Questions 1

We first examined PSTs' responses to each DME and AME task and developed task-specific rubrics to assess the four components of interest (see Table 1). For each task, in the rubric development stage, we discussed PSTs' responses to identify three-point criteria for scoring the quality of each of the four components. One of the authors and a trained research assistant applied developed rubrics to independently code PSTs' responses. Cohen's κ was computed to determine the level of agreement between the two raters on each of the four components of explanation quality for each group of tasks. For the DME tasks, the overall level of agreement was 0.92, $p < 0.05$ (range 0.77 – 1); for the AME tasks, the overall level of agreement was 0.93, $p < 0.05$ (range 0.83 – 1). We continued our analyses upon reaching 100% agreement that our coding reliably represents our assessment of PSTs' explanations for both types of tasks. We illustrate our scoring using PSTs' responses to the DME tasks. Consider PST #15's and PST #1's responses, included in Figures 2 and 3, to the "Sweater" task presented in Figure 1. We use these responses to illustrate our rubric and our scoring for the DME tasks. (See Appendix B for AME task scoring rubric).

Sweater

To solve this problem, I will let the price of the sweater be \$10 so I can have a number to work with. Since 70% of \$10 is \$7, the store would originally sell it for \$17. Now they mark it down 25% from the \$17 that they originally tried to sell it at. This means that they now will sell it for 75% of what they were originally asking for. 75% of \$17 is \$12.75. Now I need to find out how much of a markup \$12.75 is of \$10. To find out, I can do $2.75 \div 10$, since \$2.75 is the extra amount they're charging. This results in 27.5%. The store made a 27.5% profit.

Figure 2. PST #15's Explanation ("Sweater" DME Task)

the problem. We rated an explanation as (2) if the explanation revealed that the PSTs' understanding of the problem, its solution, relevant mathematical ideas, strategies, or relationships within the problem were limited to only a specific problem case. Finally, we rated an explanation on the CR-component as (1) if, within the provided explanation, the PST did not articulate an understanding of mathematical concepts, procedures, or relationships needed to "make the case" for the solution to the problem.

As illustrated in PST #15's and PST #1's explanations, both PSTs demonstrated a conceptual understanding of the problem and the embedded relationships. PST #15 demonstrated this understanding within only one specific problem case. In her explanation, PST #15 motivates selecting \$10 for the initial price by saying, "so I can have a number to work with." She does not explicitly articulate that the percent profit is independent of the initial price, to show her conceptual awareness that her specific choice for the initial price will not affect her conclusion for *any* price. Thus, on the CR-component, we scored PST #15's explanation as (2). PST #1 conceptualized the problem situation broadly by considering *all* possible classes of problem situations, as exemplified in her statement "take some original price, x ." Her statement suggests that she conceptualized the problem solution independently of any specific price. On the CR-component then, we assigned PST #1's explanation a score of (3).

In this round of the analysis, we also examined and scored PSTs' responses to the AME tasks which asked PSTs to provide explanations for their critique of student-generated explanations. For the AME tasks, to answer RQ 1 we analyzed explanations PSTs generated in support of their assessment of student explanations, together with their interpretations of student strategies and reasoning. (See AME tasks prompts (a) and (b). In the context of the AME task responses, we also scored each identified component of PSTs' explanations (see Table 1) on a 3-point scale. The Scoring rubric used for scoring PSTs' responses to the AME tasks is included in Appendix B.

Measuring Explanation Quality. We defined the strength of PSTs' explanations on each task (in the context of providing problem explanations and in the context of providing explanations for their critique of student-generated explanations) with a focus on each explanation component as a ratio of the raw score the PST received for that component (range 1–3) and a maximum possible score (max 3). Similarly, we defined the *overall* quality of PSTs' explanations for each DME and AME task as a ratio of the sum of the scores a PST received for each explanation component and the maximum possible score for the four components (max 12). To answer our first research question, we compared mean scores across the DME and AME tasks for the Overall Explanation Quality (O), Conclusion (C), Supporting Evidence (SE), Reasoning (R), and Conceptual Reference (CR) using the repeated measures ANOVA test. We then conducted correlation analysis (Pearson correlation test) to determine a possible association between PSTs' overall competency in generating mathematical explanations and their competency in analyzing and critiquing mathematical explanations.

Research Question 2

To answer our second research question, we further examined PSTs' responses to all AME tasks. Specifically, we analyzed PSTs' explanations of their assessment of analyzed explanations, together with possible revisions they proposed, and their ranking of analyzed explanations. See the description of the AME task prompts (b) and (c). This time, our focus was on specific features of student-generated explanations that PSTs noticed and addressed in their analyses. We used qualitative methods and open coding (Miles & Huberman, 1994) to identify different ways in which PSTs perceived and examined the strength of student-generated explanations. We parsed each response into meaning segments (words, phrases, or sentences) that conveyed PSTs' interpretations and judgments of student-generated explanations. This stage of data analysis comprised multiple passes through the data, during which each response was carefully annotated. We illustrate this process in Figure 4 with an excerpt from PST #15's analysis for the "Iced Tea" AME task (see Appendix A). For this task, the PSTs were asked to analyze and critique a sample of four student-generated explanations. In the excerpt below, PST #15 discusses the explanations presented by two of the four students.

Annotations were systematically compared and contrasted within and across PSTs' responses. Our goal was to delineate the meaning segments, identify their similarities, revise and collapse segment descriptions into codes, establish definitions for codes, and check for overlaps between emergent codes. Emergent themes were grouped into common criteria that discerned how the PSTs perceived and critiqued explanations. The initial descriptive codes and the resulting final coding categories are presented in Appendix C. In the subsequent rounds of analysis, we grouped those 27 codes into the final six categories that represented criteria (lenses) through which PSTs analyzed and critiqued student explanations. We then tabulated and compared code frequencies across all AME tasks to identify the overall patterns across PSTs' responses (z test for proportions). The six criteria

(lenses) through which PSTs examined student explanations, along with our discussion of observed patterns, are presented in the results section.

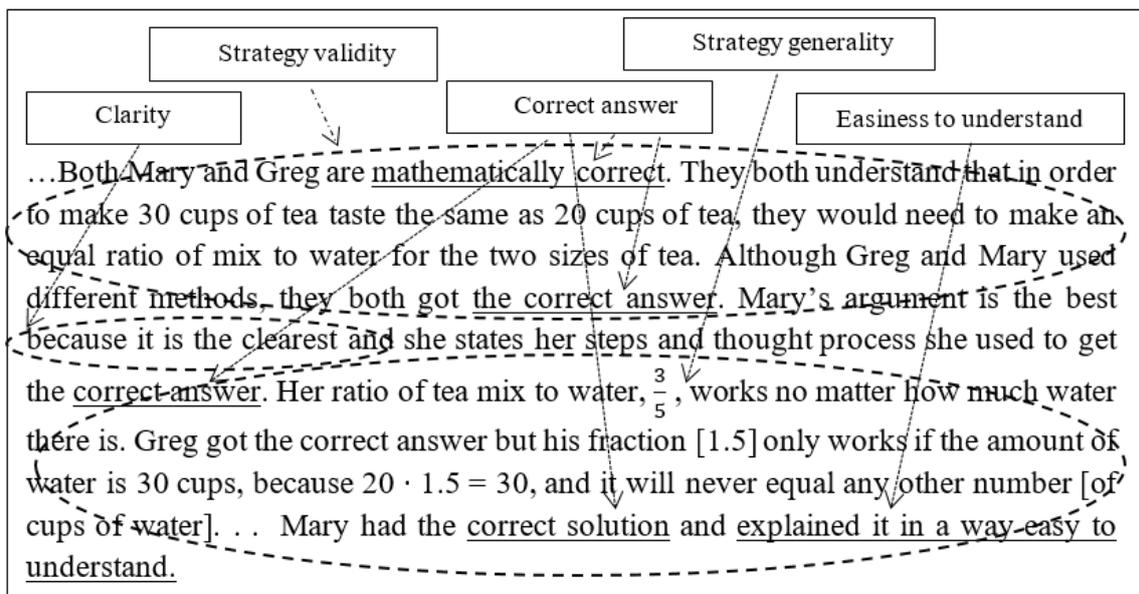


Figure 4. Excerpt of PST #15's Solution Analysis (Iced Tea AME task)

Results

RQ 1. How Does the Quality of Explanations PSTs Formulate to Support their Problem Solutions Compare and Relate to the Quality of Explanations they Provide to Support their Critique of Student-Generated Explanations?

A comparison of the overall quality of PSTs' problem explanations (DME tasks) and explanations for their critiques of student-generated explanations (AME tasks) revealed that the AME scores were significantly lower while compared to the DME scores. The repeated measures ANOVA test showed that the group means for DME and AME tasks overall, and on each component of explanation strength were statistically significantly different; $F_O(1, 36) = 55.496, p < 0.00$; $F_C(1, 36) = 5.505, p < 0.05$; $F_{CR}(1, 36) = 33.754, p < 0.01$; $F_{SE}(1, 36) = 113.970, p < 0.01$; $F_R(1, 36) = 34.736, p < 0.01$. A summary of group means overall, and for each component of explanation strength for DME and AME tasks are included in Table 2. For each measure, the mean difference was statistically significant at the 0.05 level.

Table 2. Comparison of Group Means across all DME and AME Tasks and for each Component of Explanation Quality

Measure	DME $\bar{M}, (SE)$	AME $\bar{M}, (SE)$	Mean Difference	Significance
Overall	0.739 (0.022)	0.572 (0.023)	0.167	$p < 0.01$
C	0.747 (0.022)	0.676 (0.028)	0.072	$p < 0.05$
SE	0.773 (0.025)	0.532 (0.022)	0.241	$p < 0.01$
R	0.665 (0.029)	0.523 (0.025)	0.142	$p < 0.01$
CR	0.773 (0.025)	0.580 (0.031)	0.193	$p < 0.01$

Across all tasks, PSTs' scores on generating mathematical explanations were positively correlated with their scores on critiquing mathematical explanations. The results were statistically significant at the 0.05 level, $r = 0.501$. PSTs who were more proficient in constructing explanations for the problems they solved were also more proficient in supporting their critiques of mathematical explanations they analyzed.

Moreover, for both types of tasks, PSTs' (CR) component scores were positively correlated to their (C) scores, (SE) scores, and their (R) scores. For the DME tasks, the respective correlations at the 0.01 level were: CR and C scores, $r = 0.838$; CR and SE scores, $r = 0.902$; and CR and R scores, $r = 0.594$. For the AME tasks, the respective correlations at the 0.01 level were: CR and C scores, $r = 0.644$; CR and SE scores, $r = 0.654$; and CR and R scores, $r = 0.514$. This result suggests that while generating problem explanations and generating explanations in support of the assessment of student explanations, PSTs with stronger conceptual understanding were overall more proficient in generating evidence, making accurate conclusions, and providing supportive reasons linking the provided evidence to stated conclusions.

RQ 2. What Specific Criteria Do PSTs Use as They Evaluate and Critique Students' Written Explanations?

Our analysis revealed six criteria that PSTs used while asked to examine and critique student-generated explanations, modify, or rank student-generated explanations. These criteria are (1) Correctness, (2) Organization, (3) Mathematical Foundations, (4) Communicative Power, (5) Justification, and (6) Generality. These criteria capture the specific aspects of student-generated explanations that PSTs identified and addressed in their explanation analyses. Below we discuss each criterion and illustrate with excerpts from PSTs' responses.

Attention to Correctness of Results or Strategy. This criterion of PSTs' analyses was discerned from PSTs' comments about the correctness of results, solution strategy, or execution of steps of a procedure implemented to solve a problem that was included in the analyzed explanation. An excerpt from PST #20's response provides an example:

Dan has all of the correct work until he tries to figure out the percent profit. They make a profit of \$55 but this is not the percent profit made. He would have to divide 255 by 200 to get 1.275. The 1 stands for the amount they paid, so their profit was .275. You have to multiply by 10 [sic] to get it into a percent, so it is 27.5%...Dan's argument could be challenged by looking at the percent increase he found at the end. If you multiply 200 by [0].55 (which is the % increase he found) it gets you 110. If you add it to 200 that means the sweater was sold for \$310 which is incorrect. But, if you multiply 200 by .275 you get 55 which if you add that to 200 it gets you \$255 which is what the sweater was sold for. So, this disproves his argument. (Sweater AME).

Attention to Organization. Another criterion identified within PSTs' explanation critiques was the overall organization of explanation. PSTs critiqued student-generated explanations commenting on the step-by-step flow of analyzed explanations, or the overall logical progression in which information within the analyzed explanation was shared. Excerpts from PST #5's and PST #2's responses serve as an illustration. PST #5 wrote:

Dan's steps are in a very logical order. He had to find the markup price before he could find the sales price." PST #2 observed: "[Dan's response] has a good flow from one step to the next. I am not feeling confused as to how he went from one step to the next.

Attention to Foundations. This criterion of PSTs' analyses was identified from PSTs' comments about the transparency with which the explanation articulated the problem and its interpretation, problem-solution, or solution processes. Using this criterion, PSTs examined and critiqued student explanations observing and commenting on whether or not student explanation is transparent about the meaning of terms, symbols, or variables. PST #9's response illustrates this lens of critique:

...Looking at his work, he seems to have the correct steps but one thing that could really help would be to write down all the important information first so that we know it...He has the value of x in his solution, well, what does the x stand for? He should write down what does the x is representing [sic]. (River Barge AME).

Attention to Explanation's Communicative Power. PSTs who critiqued analyzed explanations with a focus on these explanations' communicative power examined the clarity with which mathematical ideas, processes, calculations, or strategies were articulated within the student's explanation. Some PSTs emphasized how well an explanation communicated what a student did to solve a given problem. Others focused on completeness of analyzed explanations, examined, and commented on whether or not an explanation was comprehensive and included enough details. Some PSTs also critiqued explanations analyzing whether the explanation was concise or easy to understand. PST #9's response provides an example:

Most people would be confused by reading this [explanation]. I am saying this because when he is writing out his equations there are no labels so most people would have no clue as to what these numbers

represent and mean. He also didn't give much explanation as to how he got there [to the answer] and how he even started for that matter. (Sweater AME).

Attention to Justifications. PSTs demonstrated their focus on justification by identifying mathematical procedures or results in the analyzed explanations and reflecting on whether or not these results or procedures were adequately justified. Excerpts from PST #3 illustrate this focus of explanation critiques:

He does not do a good job of including why he did what he did. He should explain why he subtracted 340 – 85, (because 85 dollars was taken off with the 25% discount). He should also justify why he got 55% from 55-dollar profit, this is not correct and cannot be correctly justified. (Sweater, AME).

Attention to Generality. This criterion was operationalized as PSTs' awareness of the limitations of case-specific arguments for the solution to the problem which they recognized in student explanations. PSTs who critiqued student explanations with attention to generality explicitly articulated that conclusive evidence needs to include all possible cases in the problem domain. They also considered whether or not analyzed explanation can be applied to broader classes of problems and used this observation describing why a given explanation could be challenged. We illustrate this lens of analysis with an excerpt from PST #21's response:

The challenge is that since he picked a random number, he is not explaining that it could work for all other prices. (Sweater, AME)

Table 3 illustrates the PSTs' overall focus on each of the six criteria:

Table 3. Distribution of Criteria Identified in PSTs' AME Responses

1	2	3	4	5	6
Correctness	Organization	Foundations	Communicative Power	Justifications	Generality
# PSTs (%) (n = 37)					
34 (91.9%)	34 (91.9%)	21 (56.8%)	37 (100%)	10 (27%)	13 (35.1%)

All PSTs critiqued student-generated explanations with a focus on explanations' communicative power. Almost all (92%) also examined sample explanations with a focus on organization and correctness of conveyed results or ideas. Less frequently, PSTs critiqued student-generated explanations by attending to mathematical foundations (57%), justifications conveyed in the context of explanation (27%), or generality of presented results or procedures (35%).

Because our AME tasks were open-ended, i.e., we did not explicitly direct PSTs to comment on, or modify, any *specific* aspects of student explanations, the results reflect what PSTs spontaneously pay attention to as they examine and critique student-generated explanations. The identified criteria can also be interpreted as a window into the complexity of PSTs' explanation analyses and critiques. For example, only about 29% of PSTs used five of the above criteria as a lens of their explanation analysis and critique. About a third of PSTs (32%) used four different criteria while analyzing sample explanations, and about 27% of our PSTs analyzed and critiqued provided explanations with a focus on only two of the above criteria, (4) & (1) or (4) & (2).

The proportion of PSTs who focused their critique on mathematical foundations of analyzed explanations was significantly lower compared to the proportion of PSTs who addressed organization, correctness, or communicative power in their explanation critiques, (Foundations vs. Organization or Correctness, 56.8% vs. 91.9%, $z = 3.727$, $p < 0.01$; Foundations vs. Communicative Power, 56.8% vs. 100%, $z = 5.237$, $p < 0.01$). The same was true while comparing the proportions of PSTs who critiqued explanations with attention to justification and generality of conveyed results or procedures; significantly less PSTs used the lens of justification or generality while examining sample explanations compared to the lens of Organization, Correctness, or Communicative Power, (Justifications vs. Organization or Correctness, 27% vs. 91.9% respectively, $z = 7.466$, $p < 0.01$; Justifications vs. Communicative Power, 27% vs. 100% respectively, $z = 9.86$, $p < 0.01$; Generality vs. Organization or Correctness, 35.1% vs. 91.9% respectively, $z = 6.193$, $p < 0.01$; and Generality vs. Communicative Power, 35.1% vs. 100% respectively, $z = 8.152$, $p < 0.01$).

Discussion and Implications

Communicating one's understanding through explanation is an essential part of participating in mathematics (Hiebert et al., 1997; McClain, 2009) and listening to, and analyzing mathematical thinking (e.g., when a student provides mathematical explanation) is one of the central tasks of mathematics teaching (Hoover et al., 2016; NCTM, 1991). In this study, in the context of a problem-solving course, we examined explanations PSTs generated to support their problem-solutions and explanations they provided in support of their critique of student-generated explanations. We also identified criteria PSTs used while assessing and critiquing student-generated explanations.

We aimed to facilitate PSTs' learning to write and critique mathematical explanations with a focus on claims, supporting evidence, and sufficiency and validity of reasoning that links the evidence to identified claims. We planned our instruction intending to increase PSTs' awareness of how different ways of thinking about a problem generate different evidence and reasons in support of the problem solution. In the context of each problem PSTs solved, and each problem explanation they critiqued, we explicitly asked them to support claims they generated with evidence and reasons. We also provided them with multiple opportunities to read and reflect, individually, during small group or whole-class discussions, on weaknesses and strengths of explanations generated by others.

Research Question 1. A comparison of PSTs' DME and AME scores revealed that explanations PSTs generated in support of their critique of student-generated explanations were overall weaker while compared to explanations they generated for the problems they solved on their own. Given the instructional emphasis on providing *and* critiquing mathematical explanations, we wondered whether some other dynamics were obscuring this outcome. First, we believe that PSTs might already have had more practice in generating mathematical explanations before coming to our course. We also believe that the course activities focused on analyzing and critiquing student-generated mathematical explanations constituted considerably new experiences for our PSTs. Research on K-12 curricular materials and students' learning opportunities to generate and evaluate arguments generally documents limited emphasis on these practices in school mathematics textbooks (e.g., Bieda et al., 2014; Thompson, Senk, & Johnson, 2012). The curricular emphasis on problems or activities that engage students in *analyzing and critiquing* mathematical arguments is even more diminished (e.g., Bieda et al., 2014).

Second, our course was a first in a 3-course mathematics sequence for PSTs. All PSTs were at the beginning of their teacher education program and had overall limited experiences in analyzing student mathematical thinking and strategies. Before they could critique student-generated explanations, they needed to understand how students reasoned about the problem situation. Research shows that PSTs' ability to analyze and make sense of student thinking and reasoning is closely related to their professional noticing skills (Zambak & Magiera, 2018). While in our work with PSTs during class discussions we placed an explicit emphasis on mathematically significant aspects of student explanations; we did not directly focus on supporting PSTs' professional noticing skills, which could also explain to some extent the overall weaker scores on the AME tasks.

Our analysis revealed a positive relationship between PSTs' overall competency in generating mathematical explanations (DME scores) and their competency in critiquing explanations (AME scores). PSTs who were more proficient in generating mathematical explanations for the problems they solved were also more proficient in providing explanations in support of their critique of student explanations. Moreover, our study brings attention to the role of conceptual understanding (of the problem or student thinking and strategies embedded in student-generated explanations). In the DME and AME tasks alike, CR scores were positively correlated with C scores, SE scores, and R scores. PSTs with higher CR scores were more proficient in formulating well-articulated and valid conclusions, generating supportive evidence, and articulating why their evidence supports their claims.

We are not aware of any research that focused on the relationship between PSTs' competency in generating explanations of the problems they solve and generating explanations to support their critique of student explanations. Within the past research in the area of argumentation skills (broadly defined), competencies in generating mathematical arguments and critiquing mathematical arguments were mostly studied in isolation from one another, and primarily focused on understanding and conceptions of proofs (e.g., Knuth, 2002; Bleiler, Thompson, & Krajčevski, 2014). Given the complexity of teacher knowledge, finding ways to integrate the essential skills and aspects of teacher knowledge is critical to support the development of PSTs' professional knowledge and skills (Zambak & Magiera, 2018). Our study suggests that concurrent attention to both competencies might be beneficial for supporting PSTs' skills in generating and critiquing mathematical

explanations. Our work also shows that conceptual understanding (of the problem at hand or student explanation) significantly relates to the overall quality of explanations one generates because it is positively associated with the ability to generate evidence, formulate accurate conclusions, and provide reasons that link the evidence to stated conclusions.

Research Question 2. We uncovered six criteria that PSTs used to critique student explanations: *Correctness, Organization, Mathematical Foundations, Communicative Power, Justification, and Generality*. These criteria describe the nature of the evidence which PSTs identified in student-generated explanations they examined, and on which they drew in their critiques. Our PSTs most frequently examined student explanations with a focus on Correctness, Organization, and Communicative Power. Less frequently they attended to mathematical Foundations (which may also include any assumptions that underlie the solution to the problem), Justifications, and Generality of analyzed explanations. Even though our PSTs were systematically engaged in class discussions that engaged them in reflections on the quality of justifications and generality of discussed solutions, most of the PSTs did not address these aspects in their analyses and critiques of student-generated explanations. As discussed in the “Data Sources” section, for the AME tasks, we asked PSTs to carefully examine and describe student solution strategy and reasoning, and analyze and critique provided explanation(s) suggesting revisions to improve it. As stated earlier, the AME tasks did not explicitly direct PSTs to comment on any *specific* aspects of student explanations. The results reflect what PSTs spontaneously pay attention to as they analyze student-generated explanations.

In our study, all PSTs analyzed explanations with a focus on explanation’s communicative power and almost all used correctness and organization as criteria while examining analyzed explanations. It might be more difficult for PSTs to recognize and learn to assess student explanations with a focus on generality and the quality of justifications than, for example, with a focus on correctness or organization. It is also possible that after PSTs identified certain aspect(s) of explanations which provided a focus for their critiques (e.g., correctness, organization, communicative power), they simply stopped conducting further analyses and seeking any additional support for their explanation critiques. Bleiler et al. (2014) studied written feedback on proofs that secondary school PSTs provided to students. They reported that upon identifying first, most noticeable error in student’s proof, PSTs frequently failed to look for any additional aspects of proof on which they could provide feedback to the student. It might seem plausible to think that, likewise, upon identifying certain aspect(s) of explanations many of our PSTs simply stopped conducting further analyses and did not think about any additional aspects of student explanations to address in their critiques.

However, as our analysis revealed, about a third of our PSTs used four of the six identified criteria in their explanation critiques, and about an additional third critiqued analyzed explanations with a focus on five of the six identified criteria. This result shows that our PSTs’ explanation critiques were rather complex while considering the number of aspects they addressed. While the set of criteria we identified might serve as a tool for assessing the growth in PSTs’ ability to analyze and critique mathematical explanations, it also provides insights into aspects of students’ explanations that PSTs might not routinely consider. Even though our PSTs were systematically engaged in class discussions centered on heightening their attention to the quality of justifications and generality of discussed solutions, most of the PSTs did not attend to these aspects spontaneously when asked to analyze and critique student-generated explanations. Our finding that PSTs do not routinely attend to justification and generality while analyzing student-provided explanations is consistent with past research. For example, Lo, Grant, and Flowers (2008) reported that many PSTs tend to describe what was or needs to be done, rather than think about *why* it is valid. Morris (2007) similarly found that PSTs who were asked to analyze and critique student arguments in the context of classroom discussions used a wide range of criteria but rarely focused on the logical validity of analyzed arguments.

The results suggest that PSTs might benefit from activities that explicitly help them recognize different features of student-generated explanations. In our current work with PSTs, when we now engage PSTs in analyzing and critiquing their mathematical explanations or mathematical explanations of students, we are explicitly directing their attention to the *specific* aspects of explanations that we want them to examine and reflect on (e.g., justifications, generality, foundations—which may also include any assumptions that underlie the solution to the problem). Our goal is to increase PSTs’ awareness of a broad range of features that contribute to explanation strengths, or weaknesses. By engaging PSTs in analyzing and critiquing student-generated explanations and being explicit about features of analyzed explanations they examine, our goal is also to increase PSTs’ awareness of the strength and weaknesses of their explanations they generate, giving them tools to self-critique explanations they generate.

Final Remarks

To effectively promote the growth of students' reasoning and mathematical understanding, PSTs need the ability to generate mathematical explanations and critically unpack explanations of their students (Hoover et al., 2016). In this study, we concurrently explored PSTs' competencies in generating problem explanations and the explanations they provide in support of their critique of student-generated problem explanations. Our selection of participants was limited to PSTs from one institution, and our selection of problems was also limited to problems typical of the elementary mathematics curriculum. Building a further understanding of PSTs' competencies in generating and analyzing mathematical explanations requires future research focusing on a broader range of problems and diverse groups of PSTs. To continue unpacking PSTs' competencies in generating problem explanations and critiquing student-generated explanations, future research might also explore different instructional sequences of activities and their effect on the development of PSTs' skills in these areas. For example, in addition to analyzing and critiquing student-generated explanations, PSTs could be asked to reflect on the pedagogical values of their critiques. The goal would be to increase PSTs' awareness of different aspects of explanations that contribute to an explanation's overall strength. At the same time, the goal would be to heighten PSTs' awareness of characteristics of explanations which they might not routinely consider or choose to address in problem explanations they produce or critique (e.g., statement of mathematical foundations, assumptions that give rise to the explanation, explanation generality, presence, and validity of justification).

Funding Information

The research described in this article was supported by the National Science Foundation's grant to Marta T. Magiera for the project entitled "CAREER: L-MAP: Pre-service Middle School Teachers' Knowledge of Mathematical Argumentation and Proving." Grant No. DRL-1350802. Opinions, findings, conclusions, or recommendations are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

- Australian Curriculum and Assessment Reporting Authority [ACARA], 2015. Australian curriculum: Mathematics, Version 8.1 Retrieved from www.australiancurriculum.edu.au/Mathematics/Curriculum/F-10
- Balacheff, N. (2010). Bridging knowing and proving in mathematics. A didactical perspective. In G. Hanna, N. H. Jahnke, & H. Pulte (Eds.), *Explanation and proof in mathematics* (pp. 116-135). New York, NY: Springer Science + Business Media.
- Banes, L. C., López, G., Skubal, M., & Perfecto, L. (2017). Co-constructing written explanations. *Mathematics Teaching in the Middle School*, 23(1), 30-38.
- Bicknell, B. (1999). The writing of explanation and justification in mathematics: Differences and dilemmas. In J. Truran & K. M. Truran (Eds.), *Proceedings of the 22 Annual Conference of the Mathematics Education Research Group of Australasia*, (pp. 75-83). New South Wales, Australia: Mathematics Education Research Group of Australasia.
- Bieda, K., Ji, X., Drwencke, J., & Pickard, A. (2014). Reasoning-and-proving opportunities in elementary mathematics textbooks. *International Journal of Educational Research*, 64, 71-80.
- Bleiler, S. K., Thompson, D. R., & Krajčevski, M. (2014). Providing written feedback on students' mathematical arguments: proof validation of prospective secondary mathematics teachers. *Journal of Mathematics Teacher Education*, 17(2), 105-127.
- Bostic, J. (2016). Fostering justification: A case study of preservice teachers, proof-related tasks, and manipulatives. *Journal of Mathematics Education at Teachers College*, 7(1), 35-43.
- Boyle, J. D., Bleiler, S. K., Yee, S. P., & Ko, Y. (2015). Transforming perceptions of proof: A four-part instructional sequence. *Mathematics Teacher Educator*, 4(1), 32-70.
- Chi, M. (2000). Self-explaining: The dual processes of generating inference and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science* (vol. 5, pp. 161-238). Mahwah, NJ Lawrence Erlbaum Associates.
- Chi, M., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13(2), 145-182.
- Choppin, J. (2007). Teacher orchestrated classroom arguments. *Mathematics Teacher*, 101(4), 306-310.

- Cross, D. I. (2009). Creating optimal mathematical learning environments: Combining argumentation and writing to enhance achievement. *International Journal of Science and Mathematics Education*, 9, 905–930.
- Dolev, S., & Even, R. (2015). Justifications and explanations in Israeli 7th-grade mathematics textbooks. *International Journal of Science and Mathematics Education* 13, 309-327.
- Esmonde, I. (2009). Explanations in mathematics classrooms: A discourse analysis. *Canadian Journal of Science, Mathematics and Technology*, 9(2), 86-99.
- Evens, H., Houssart, J. (2004). Categorizing pupils' written answers to a mathematics test question: 'I know but I can't explain'. *Educational Research*, 46(3), 269–282.
- Felton, M. (2007). Context and preservice teachers' conceptions of proof. In (T. Lamberg & L. R. Wiest, Eds.). *Proceedings of the 29th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, (pp. 85-92). Stateline (Lake Tahoe), NV: University of Nevada, Reno.
- Forman, E. A., McCormick, D. E., & Donato, R. (1997). Learning what counts as a mathematical explanation. *Linguistics and Education*, 9(4), 313-339.
- Francisco, J. M. (2013). Learning in collaborative settings: Students building on each other's ideas to promote their mathematical understanding. *Educational Studies in Mathematics*, 82, 417–438.
- Francisco, J. M., & Maher, C.A. (2005). Conditions for promoting reasoning in problem solving: Insights from a longitudinal study. *Journal of Mathematical Behavior*, 24(3-4), 361–372.
- Hafner, J., & Mancosu, P. (2005). The varieties of mathematical explanation. In P. Mancosu (Ed.). *Visualization, explanation and reasoning styles in mathematics*, (pp. 215-250). Springer.
- Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational Studies in Mathematics*, 44(1-3), 5–23.
- Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Wearne, D., ... Human, P. (1997). *Making sense: Teaching and learning mathematics with understanding*. Portsmouth, NH: Heinemann.
- Hodds, M., Alcock, L., & Inglis, M. (2014). Self-explanation training improves proof comprehension. *Journal for Research in Mathematics Education*, 45(1), 62-101.
- Hoover, M., Mosvold, R., Ball, D., & Lai, Y. (2016). What does it take to develop assessment of mathematical knowledge for teaching? Unpacking the mathematics work of teaching. *The Mathematics Enthusiast*, 13(1/2), 3–34.
- Kline, S., & Ishii, D. K. (2008). Procedural explanations in mathematics writing. *Written Communication*, 25(4), 441-461.
- Knipping, C. (2008). A method for revealing structure of argumentation in classroom proving process. *ZDM Mathematics Education*, 40, 427–441.
- Knuth, E. J. (2002). Secondary mathematics teachers' conceptions of proofs. *Journal for Research in Mathematics Education*, 33(5), 379-405.
- Krummheuer, G. (1995). The ethnography of argumentation. In P. Cobb, & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 229-269). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lepak, J. (2014). Analyzing students' written arguments (2014). *Mathematics Teaching in the Middle School*, 20(4), 213-219.
- Levenson, E. (2013). Exploring one students' explanations at different ages: the case of Sharon. *Educational Studies in Mathematics*, 83(2), 181–203.
- Levenson, E., & Barkai, R. (2013). Exploring the functions of explanation in mathematical activities for children ages 3-8 year old: The case of the Israeli curriculum. In B. Ubuz, C. Haser, & M. Mariotti (Eds.), *Proceedings of the Eighth Congress of the European Society for Research in Mathematics Education* (pp. 2158-2167). Ankara, Turkey: Middle East Technical University.
- Lo, J.J., Grant, T.J., & Flowers, J. (2008). Challenges in deepening prospective teachers' understanding of multiplication through justification. *Journal of Mathematics Teacher Education*, 11(1), 5–22.
- McClain, K. (2009). When is an argument just an argument? In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades. A K-16 perspective* (pp. 222-234). New York, NY: Routledge.
- Martin, W. G., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education*, 20(1), 41-51.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed., pp. 10–12). Newbury Park, CA: Sage.
- Morris, A. K. (2007). Factors affecting pre-service teachers' evaluations of the validity of students' arguments in classroom contexts. *Cognition and Instruction*, 35(4), 479-522.

- Morselli, F., & Levenson, E. (2014). Functions of explanations and dimensions of rationality: Combining frameworks. In P. Liljedhal, C. Nicol, S. Oesterle, & D. Allan (Eds.), *Proceedings of the joint meetings of PME 38 and PME-NA 36* (Vol. 4, pp. 249-256). Vancouver, Canada: PME.
- Nardi, E., Biza, I., & Zacheriades, T. (2012). 'Warrant' revisited: Integrating mathematics teachers' pedagogical and epistemological considerations into Toulmin's model for argumentation. *Educational Studies in Mathematics*, 79(2), 157-173.
- NGA & CCSSO (National Governors Association Center for Best Practices & Council of Chief State School Officers). (2010). *Common Core State Standards for Mathematics*. Washington, DC: Authors.
- National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- National Council of Teachers of Mathematics (NCTM). (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.
- Pedemonte, B., & Balacheff, N. (2016). Establishing links between conceptions, argumentation and proof through the κ -enriched Toulmin model. *Journal of Mathematical Behavior*, 41, 104-122.
- Prediger, S., & Erath, K. (2014). Content, interaction or both? Synthesizing two German traditions in a video study on learning to explain the mathematics classroom microcultures. *Eurasia Journal of Mathematics, Science and Technology Education*, 10(4), 313-327.
- Reid, D., & Zack, V. (2009). Aspects of teaching proving in upper elementary classroom. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proofs across the grades: A K-16 perspective*, (pp. 133 - 146). New York, NY: Routledge.
- Stacey, K., & Vincent, J. (2009). Modes of reasoning in explanations in Australian eight-grade mathematics textbooks. *Educational Studies in Mathematics*, 72(3), 271-288
- Stylianides, A. J., & Stylianides, G. J. (2009). Proof construction and evaluation. *Educational Studies in Mathematics*, 72(2), 237-253.
- Stylianides, G. J., Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education*, 40(3), 314-352.
- Stylianides, G. J., Stylianides, A. J., & Phillippou, G. N. (2007). Pre-service teachers; knowledge of proof by mathematical induction. *Journal of Mathematics Teacher Education*, 10(3), 145-166.
- Thompson, D. R., Senk, S. L., & Johnson, G. (2012). Opportunities to learn reasoning and proof in high school mathematics textbooks. *Journal for Research in Mathematics Education*, 43(3), 253-295.
- Toulmin, S. E. (1958/2003). *The uses of argument*. Cambridge, UK: Cambridge University Press.
- Wagner, P. A., Smith, R. C., Conner, A., Singletary, L. M., & Francisco, R. T. (2014). Using Toulmin's model to develop prospective secondary mathematics teachers' conceptions of collective argumentation. *Mathematics Teacher Educator*, 3(1), 8-26.
- Wall, J., Selmer, S., & Bingham Brown, A. (2016). Assessing elementary prospective teachers' mathematical explanations after engagement in online mentoring modules. *Contemporary Issues in Technology and Teacher Education*, 16(3), 388-414.
- Weber, K., & Alcock, L. (2005). Using warranted implications to understand and validate proofs. *For the Learning of Mathematics*, 25(1), 34-51.
- Yackel, E. (2004). Theoretical perspectives for analyzing explanations, justification and argumentation in mathematics classrooms. *Journal of the Korea Society of Mathematical Education Series D: Research in Mathematical Education*, 8(1), 1-18.
- Yopp D. A. (2015). Prospective elementary teachers' claiming in responses to false generalizations. *The Journal of Mathematical Behavior*, 39, 79-99.
- Zambak, V. S., & Magiera, M. T. (2018). Pre-service K-8 Teachers' Professional Noticing and Strategy Evaluation Skills: An Exploratory Study. *Eurasia Journal of Mathematics, Science and Technology Education*, 14(11), em1572. <https://doi.org/10.29333/ejmste/92021>

Author Information

Marta T. Magiera

Marquette University
 Department of Mathematical and Statistical Sciences
 United States of America
 Contact e-mail: marta.magiera@marquette.edu

Vecihi S. Zambak

Monmouth University
 Department of Curriculum and Instruction
 United States of America

Appendix A. Additional Task Examples

AME Task: Iced Tea

Mary made some iced tea from a mix, using 12 tablespoons of ice tea mix and 20 cups of water. Chris and Greg thought it tasted great, but they needed 30 cups of tea for their party. Frank arrived, and they found they disagreed about how to make 30 cups that tasted just the same:

Chris: It's easy: Just add 10 tablespoons of tea and 10 cups of water. Includes everything by 10.

Greg: Wait a minute. 30 is just 1 and a $\frac{1}{2}$ times of 20, so since you add $\frac{1}{2}$ as much water, you should add $\frac{1}{2}$ the tea: Add 10 cups of water and 6 tablespoons of tea.

Mary: I think about it this way: We used 12 tablespoons for 20 cups, so $12/20=3/5$ tablespoons for 1 cup, so for 30 cups we should use $30 \cdot 3/5=18$ tablespoons.

Frank: Wait: $20-12=8$, so you want to keep the difference between water and tea at 8. Since there are 30 cups of water, we should use $30-8=22$ tablespoons of tea. That will keep everything the same.

AME Task: Sweater

A clothing store bought a sweater for a certain price and marked it up 70%. The sweater did not sell so the store owner took 25% off the marked-up price and sold it for that price. What percent profit did the store make?

This is how Dan argued about this problem:

If the store paid for the sweater, for example, \$200, the 70% mark-up would be

$$\$200 \cdot \frac{70}{100} = 140$$

So, the store is first trying to sell the sweater for $\$200 + \$140 = \$340$.

The 25% discount on the store price would be

$$340 \cdot \frac{25}{100} = \frac{340 \cdot 25}{100} = \$85$$

That means that they actually sold the sweater for $340 - 85 = \$255$. Since the store bought it for \$200 and sold it for \$255, there is \$55 profit. They made 55% profit.

Appendix B. Explanation Critique Scoring Rubric

Explanation Component	Scoring
C	Claims about analyzed explanation are: (3) well-articulated, clear and valid (2) valid but lack clarity (1) no valid claims about analyzed explanation
SE	Explanation critique includes: (3) comprehensive, specific and valid evidence in support of one's claims about analyzed explanation (2) specific but incomplete evidence for one's claims about analyzed explanation (1) does not provide, or includes vague or incorrect evidence in support of one's claims about analyzed explanation
R	Explanation critique: (3) provides comprehensive clear links for <i>why</i> the identified evidence supports ones' assessment/critique of analyzed explanation (2) partially articulates <i>why</i> the provided evidence supports ones' assessment/critique of analyzed explanation (1) fails to provide, or includes unclear reasons for <i>why</i> the identified evidence supports ones' assessment/critique of analyzed explanation
CR	Explanation critique reveals that PST has: (3) strong conceptual understanding of mathematical ideas and student's reasoning that provide the basis for analyzed explanation (2) some understanding of mathematical ideas and student's reasoning that provides the basis for analyzed explanation (1) no evidence of understanding of mathematical ideas or student's reasoning that provides the basis for analyzed explanation

Appendix C. PSTs' Explanation Analyses and Critiques: Initial Codes and Codes Grouping

Overall Lens of Explanation Analysis	Specific Focus of Explanation Analysis
The notion of validity	(1) Correctness of final results (2) Correctness of intermediate results (3) Correctness of solution strategy (4) Appropriateness of solution strategy (5) Correctness of mathematical steps
The notion of organization	(6) Flow of steps (7) Completeness; step-by-step (8) Logical order of steps
The notion of providing foundations	(9) Assumptions about the problem (10) Meanings of symbols (11) Meanings of variables (12) Meanings for representations used (13) Articulation of what is/was known
The notion of communicating meaning	(14) Contextualization of results (15) Contextualization of strategies (16) Contextualization of procedures (17) Clarity for understanding (18) Easiness for understanding (19) Easiness for following (20) Conciseness (21) Consistency in mathematical notation (22) Consistency in language used (23) Comprehensiveness
The notion of justification	(24) Justification for selected strategy (25) Justification for results
The notion of generality	(26) Generality of presented strategy (27) Generality of presented results