Running Head: Features of a Good Pedagogical Proof

Mathematicians' Perspectives on Features of a Good Pedagogical Proof

Yvonne Lai

University of Michigan

Keith Weber and Juan Pablo Mejía-Ramos

Rutgers University

Correspondence should be addressed to Yvonne Lai, 610 East University Avenue, 1600 School of Education, University of Michigan, Ann Arbor, MI, USA, 48109. Email: yxl@umich.edu

Abstract

In this paper, we report two studies investigating what mathematicians value in a pedagogical proof. Study 1 is a qualitative study of how eight mathematicians revised two proofs that would be presented in a course for mathematics majors. These mathematicians thought that introductory and concluding sentences should be included in the proofs, main ideas should be formatted to emphasize their importance, and extraneous or redundant information should be removed to avoid distracting or confusing the reader. Study 2 is a quantitative study assessing the extent to which a larger group of mathematicians ($N = 110$) agreed or disagreed with the eight mathematicians interviewed in Study 1. This quantitative study confirmed the findings of Study 1 by demonstrating a high degree of agreement among mathematicians regarding how they would revise proofs for pedagogical purposes.

Providing explanations is fundamental to mathematics instruction (Charalambous, Hill, & Ball, 2011); through explanation, teachers convey mathematical subject matter to their students (e.g., Leinhardt, Putnam, Stein, & Baxter, 1991), establish classroom norms, illustrate productive metacognitive processes, and represent the discipline of mathematics (Larreamendy-Joerns & Muñoz, 2010; Schoenfeld, 2010). For these reasons, Charalambous, Hill, and Ball (2011) contended that "providing instructional explanations lies at the heart of teaching, for it requires transforming the content in mathematically legitimate *and* pedagogically appropriate ways" (p. 443, authors' emphasis). There has been a great deal of research on instructional explanations, largely focusing on teachers' difficulties or inability to provide adequate explanations (e.g., Ball, 1988; Inoue, 2009; Leinhardt, 1989; Lo, Grant, & Flowers, 2004; Thompson & Thompson, 1994, 1996; Thanheiser, 2009), and more recently on improving teachers' abilities to provide high quality explanations (e.g., Charalambous, Hill, & Ball, 2011; Kinach, 2002; Inoue, 2009; Thanheiser, 2010). However, most research has been conducted with elementary mathematics teachers; little work of this type has been done with teachers of tertiary mathematics.

This paper concerns pedagogical practice in communicating advanced mathematics at the tertiary level. In these courses, the predominant way of presenting mathematical subject matter is via mathematical proof. By mathematical proof, we follow the mathematician Griffiths (2000) who defined mathematical proof as "a formal and logical line of reasoning that begins with a set of axioms and moves through logical steps to a conclusion" (p. 2). Other characteristics of this genre include the use of precise definitions rather than informal descriptions of concepts, the (relative) lack of diagrams and other intuitive representations of concepts, and the use of logical syntax (e.g., Weber & Alcock, 2009).

Although some mathematics educators and mathematicians question whether these types of proofs are an appropriate way of conveying mathematics to students (e.g., Davis & Hersh, 1981; Hersh, 1993; Kline, 1973; Thurston, 1994), proof is still the primary means by which mathematics is presented in advanced mathematics classrooms. For instance, Davis and Hersh (1981) asserted that "a typical lecture in advanced mathematics … consists entirely of definition, theorem, proof, definition, theorem, proof, in solemn and unrelieved concatenation" (p. 151). Similarly Dreyfus (1991) claimed teaching in advanced mathematics courses “so often follows the sequence theorem-proof-application” (p. 27). In Weber’s (2004) and Mills’ (2011) studies of mathematics lectures, professors spent, on average, about half of their lecture time presenting proofs to students.

In this paper, we explore mathematicians’ conceptions of what constitutes a good *pedagogical proof*. By a pedagogical proof, we refer to a proof that is presented to students for pedagogical purposes. The goal of this paper is to investigate what features mathematicians think a good pedagogical proof should have. This fills two underrepresented areas of research in mathematics education. First, there is little research in mathematics education on the knowledge and dispositions of mathematicians regarding their pedagogical practice, including practices around proof presentation. Second, based on a systematic review of the mathematics education literature on argumentation and proof, Mejía-Ramos and Inglis (2009) found that empirical research on proof has mainly focused on students’ construction and evaluation of proof; they could find no empirical studies on how mathematical instructors (or for that matter, students) chose to present mathematical arguments or proofs. Research on what mathematicians value in a pedagogical proof can provide researchers with a better lens to interpret mathematicians' pedagogical behavior.

The research questions we will address in this paper are:

(1) What features do mathematicians believe a good pedagogical proof should have?

(2) For what mathematical or pedagogical reasons do they value those features?

To address these questions, we conducted two studies. The first study is a qualitative study in which we observed eight mathematicians revising two proofs to be presented in a course for mathematics majors; these proofs were purportedly generated by a colleague. The second study is a large-scale quantitative study in which we examined the extent that findings from the qualitative study generalized to 110 practicing mathematicians.

RELATED LITERATURE

Communicative Functions of Proof

A significant function of proof in mathematics is to provide conviction that an assertion is true (e.g., Harel & Sowder, 1998). However, in a seminal paper, de Villiers (1990) argued that proof is much more than that. To mathematicians, proof can serve as a tool to explain why a theorem is true, systematize a mathematical theory, or discover new theorems. De Villiers (1990) further argued that an important function of proof is communication: Shared language and standards of argumentation facilitate debate about sophisticated mathematical ideas among mathematicians. Many educators have argued that proof should play analogous pedagogical roles in the mathematics classroom; in addition to convincing students that a theorem is true, proof should provide explanation and facilitate communication (e.g., Alibert & Thomas, 1991; Hanna, 1990; Healy & Hoyles, 2000;

Knuth, 2002). In this paper, we view proof as a means of explanation and communication between instructors and students.

Educational Research on Proof Presentation

Mathematicians and mathematics educators have remarked that students generally find proofs confusing and learn little from reading them (e.g., Davis & Hersh, 1981; Harel, 1998; Hersh, 1993; Leron & Dubinsky, 1995; Porteous, 1986; Rowland, 2001; Thurston, 1994), with many arguing that the linear, formal nature of proof inhibits students' understanding. Specifically, formal proofs often mask the intuitive representations that are needed to generate and understand them, and the jargon, logical syntax, and abstractness of a proof are intimidating barriers to comprehension (e.g., Hersh, 1993; Kline, 1980; Leron, 1983; Rowland, 2001; Thurston, 1994). Consequently mathematics educators have proposed alternatives to formal proofs to present mathematical information to students (e.g., Alcock, 2009; Hersh, 1993; Leron, 1983; Rowland, 2001); the extent to which these formats improve student understanding is an important open research question.

Aside from these instructional recommendations, research on proof presentation in mathematics education has been limited. In particular, we are not aware of any studies on what mathematicians value in proofs for pedagogical purposes, how mathematicians alter proofs to reduce obstacles to understanding, or how mathematicians choose to present proofs to their students. The studies presented in this paper address these issues.

STUDY 1

Rationale

In this study, mathematicians were given two proofs to revise for clarity. They were told that these proofs were to be presented to second- or third-year mathematics majors. The rationale behind this study was that a mathematician's revisions would presumably improve the pedagogical quality of the proof, in his or her view, by eliminating negative features of the proof or introducing positive features. We used the mathematicians' revisions, as well as their comments on why they made these revisions, to form grounded hypotheses on what features they believed a good pedagogical proof should have. We note that the reasoning used to revise an existing written pedagogical proof may inform, but is not synonymous with, the processes of generating a pedagogical proof for lecture; consequently while the aim of this study is to investigate what mathematicians value in a pedagogical proof, the link between these values and mathematicians' in-class pedagogical practice is not addressed.

Method

Participants. The first author, a mathematics research post-doctoral fellow during the data collection period, invited colleagues to participate in this study. We strove to obtain participants with a broad range of research interests and levels of experience. We do not have *a priori* reason to believe these participants were more interested or more capable at mathematics teaching than other mathematicians. Eight mathematicians agreed to participate in this study. The participants' areas of research included algebra, analysis, geometric topology, combinatorics, and mathematical physics. The range of teaching experience of the participants also varied, ranging from 1 year to 30 years, with six of the participants having at least 5 years experience working in mathematics departments. All mathematicians currently work, either in post-doctoral positions or in tenure-track positions, at research institutions in the top 25 graduate programs in mathematics in the United States. We refer to all participants by using pseudonym initials and using male pronouns to protect their anonymity.

Materials and Procedures. Each participant met individually with the first author for a task-based interview. Participants were given the exercise, "Prove that a differentiable function from **R** to **R** with strictly positive first derivative is injective. Use the Mean Value Theorem in your solution," but this task is not discussed in this paper. The participants were then given two Revision Tasks and were told they would be given two sample solutions to revise. Participants were then presented with Task A and Task B (see Figures 1 and 2). The instructions informed the participants, "the intended audience [of this proof] is a sophomore- or junior-level mathematics major" and they were told to "improve the clarity" of the proof. Participants were told to "think aloud" as they were making their revisions. Because this task asked participants to improve the clarity, rather than the correctness of this proof, and because a specific undergraduate audience was mentioned, we assumed the participants would treat these proofs as pedagogical proofs.

***** Insert Figure 1 and Figure 2 about here *****

The research team created Task A and Task B for this study. Features of Task A include a small error in the manipulations of the inequalities ("Equivalently, $x_1 - x_2 > 0$"), to gauge how finely the participant read the text; a mathematically valid but extraneous statement at the end ("If for some x_1, x_2 in **R**, we had …"), to see how mathematicians would work with redundancy; a reversal of the convention of expressing the denominator of the slope as the difference between the greater input and lesser input rather than vice versa; and mathematically correct but potentially extraneous arithmetic manipulations.

Features of Task B include an appeal to cases at the end but not the beginning of the argument; terminology more commonly used in courses for non-mathematics majors rather than for mathematics majors (e.g., "coordinate point"); informal language ("the Mean

Value Theorem gives”); and mathematically correct but potentially extraneous non-arithmetic steps (“there is a unique line,” “x_3 is in the domain of f”). Neither Task A nor Task B contained introductory sentences that explicitly described the heuristic used for proving a function is injective (i.e., given two distinct values in the domain, prove that their output values are distinct).

After completing these tasks, participants were asked questions on their experience revising the proof. These questions included:

- Could you describe, in a sentence or two, the main ideas behind these proofs?
- Did you try to emphasize the main idea of the proof in any way?
- How did your revisions improve the clarity of the proof?
- Would the exposition of the proof without the slope formula be improved with the inclusion of the slope formula?

Analysis. To analyze the types of revisions performed, we adopted the methodology used in Weber’s (2008) study of mathematicians’ proof validations, using an open coding scheme in the style of Strauss and Corbin (1990). For each revision that a participant made on the Revision Tasks, we made a general description of the edit, went through the transcript and noted the reason given for the edit (if any), and developed category names. New episodes were placed into existing categories when appropriate, used to create new categories, or modify names or definitions of existing categories. Once these categories were formed, we repeated the same process for each category, coding the reasons that participants provided for making the revision that they did, although we note here that many of the revisions the participants implemented were made without comment.

RESULTS OF STUDY 1

The eight participants collectively made 141 revisions. Of these 141 revisions, we considered 20 revisions to be *trivial*, where a revision was coded as trivial if the participant replaced a phrase with an equivalent one (e.g., substituted “Let *x*” with “Fix *x*” or “Mean Value Theorem” with “MVT”), the participant did not comment on the revision during the experiment, and we believed the revision did not impact the meaning or emphasis of the original proof. Our analysis focuses on the remaining 121 non-trivial revisions.

In coding the data, we separated these revisions into three broad categories:

- *added* text, that does not alter the structure or meaning of the proof, except possibly by adding justification to an existing step of the proof or introducing new notation. There were a total of 35 additions.
- *altered* text, that substitutes original text, or is new text that changes the structure of the proof. There were a total of 49 alterations.
- *deleted* text, that the participant did not replace with new text. There were a total of 37 deletions.

Where we could, we used the mathematician’s words to guide the coding. Some revisions had unstated reasons. In these cases, we sometimes drew inferences about the type and intention of the revision based on the mathematical reasoning of the proof. The purpose of this coding scheme was not to devise precise categories for any possible revision of any mathematical work, but to organize provisionally the subject of our analysis: the motivation and nature of revisions for the tasks that we provided.

Revisions Classified as Added

We categorized five distinct ways in which participants added text to their proofs. These categories are summarized in Table 1 and discussed below.[1]

***** Insert Table 1 About Here *****

Adding an Introductory or Concluding Sentence to Reveal the Proof Framework. Neither Task A nor Task B contained sentences that made explicit the logical structure of the proof. Revisions supplying such sentences were very common, with 11 instances performed by 7 participants. Typical of this type of revision are participant CY's work on Task A and TR's work on Task B, presented in Figure 3. Their added introduction and conclusion sentences, which are boxed, make explicit what was being accomplished in the respective proofs. As TR said of his conclusion sentence (italics are our emphasis):

> TR: … because they didn't say at the beginning *how they were going to use these two distinct elements to show that f is injective*, I felt like there was one sentence left that was needed: since x_1 and x_2 are arbitrary, it follows that f is injective, because the values are different.

***** Insert Figure 3 About Here *****

Unlike TR, most participants did not provide a reason for why they added these introductory or concluding sentences. We suggest that their revisions can be viewed as revealing the implicit logical structure in the two Revision Tasks. Although mathematicians can often infer the logical structure of a proof from the statement being proven, Selden and Selden's (1995, 2003) research revealed that students often are not able to draw this inference. Hence, this type of revision may have pedagogical value by making clear to students what type of proof is being applied, what is being assumed, and what is being concluded. Indeed, two participants suggested that whether or not their added introductory sentence belonged in the revision would depend on the student population reading the proof.

Introducing Additional Algebraic Justification. Each participant introduced at least one formula for the purposes of justification into one of the proofs they were revising, with two categories of stated reasons for these revisions. This was most sharply illustrated with the following seemingly trivial manipulation of the quotient in Task A:

$$f'(x_3)(x_2 - x_1) = f(x_2) - f(x_1).$$

Five of the eight participants introduced this formula. KT and IR, who both added the manipulation, reasoned:

> KT: So to me, the *heart of the matter* is noticing that OK, this is positive, this is positive, so therefore this third quantity is positive. And so I rewrote that statement. (Italics are our emphasis.)
> IR: [pointing to the phrase "So $f(x_1) - f(x_2) < 0$"] And so, *but this 'So' is what they have to justify. And it is not justified out of anything.* I mean, it is justified to *me* out of this equation and this equation, but it is not justified to a student. (Italics are our emphasis.)

Their viewpoints are representative of the motivations stated by participants for adding algebraic justification: the formulas or manipulations were central to the ideas of the proof, or there was a logical gap that needed to be bridged.

Introducing the Slope Formula. Although the "rise-over-run" definition of slope is generally considered common knowledge to advanced mathematics majors, three mathematicians inserted the formula

$$f'(x_3) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

in Task B, where it was not originally present. For example, FR wrote:

> By our hypotheses, $f'(y)$ is positive; so $f'(y) = \dfrac{f(x_2) - f(x_1)}{x_2 - x_1}$ implies $f(x_1) \neq f(x_2)$. Hence f is injective.

Despite the fact that he did not do anything with the slope formula other than stating it, he defended the inclusion by citing its importance to the ideas of the proof:

> FR: *Because the [main] idea as I see it is this equation right here*, that $f'(y) = \dfrac{f(x_2) - f(x_1)}{x_2 - x_1}$. And I feel like *as long as that equation appears, everything else is just words*. … So yeah, I did make sure that equation appeared in the text of this proof. (Italics are our emphasis.)

We hypothesize that the connection of a step to the main ideas of the proof can have as much to do with its inclusion as the difficulty of a maneuver.

Revisions Classified as Altered

The six types of alterations that participants made are summarized in Table 2.

***** Insert Table 2 About Here *****

Altered the Main Assumptions. As written, Task B begins by setting x_1 and x_2 to be distinct real numbers. Each participant strengthened the assumption of distinctness to the condition $x_1 < x_2$. While this does not affect the validity of the proof, it does simplify the calculations in the proof, as noted by CY. He attributed the revision to concreteness:

> CY: I think … you really do want $x_2 > x_1$, because when you don't say that, then when you say $f'(x_3)$ is positive, you're left with this either/or. And this either/or is exactly dependent on whether $x_2 > x_1$ or $x_1 > x_2$ and that's just confusing. And you want to be able to draw the picture, and when you draw the picture, x_1 is

> going to be on one side of x_2. So I think you have to make x_2 bigger than x_1 for the picture to make sense.

In this excerpt, CY observed that the condition $x_1 < x_2$ obviated the need for the cumbersome proof by cases that appeared later, making the subsequent proof more clear and direct. Similar sentiments were expressed by FR, KT, IR, and KM.

Altered Reference to Mean Value Theorem. Task B states, "The Mean Value Theorem gives an x_3 in the real numbers such that.…" Six participants revised this reference to the Mean Value Theorem; IR and CY revised the statement of the Mean Value Theorem in Task A as well. Participants offered two reasons for these revisions: mathematical accuracy and student perception. Five participants objected to the use of the word "give," as the Mean Value Theorem is not constructive (i.e., the statement of the Mean Value Theorem guarantees the existence of an x_3, but does not provide a method for determining a specific value for such an x_3). BT reacted particularly strongly when asked why he made this revision.

> BT: Because nobody knows where x_3 is. A theorem does not 'give' anything. I mean, it doesn't give a number. Again you want to be clear—I mean it is colloquial talk, people could say that. But I would say something like, 'By the Mean Value Theorem, there exists.…' OK, that's what I would say. You don't know what it is, but it exists."

Although BT was the only participant to raise such a thorough objection to the mathematical inaccuracy of the sentence, four other mathematicians revised the reference similarly. FR, who replaced "gives" with "says there exists," mentioned in passing that the statement was "a slight lie about what the Mean Value Theorem says, or does."

Although it did not affect the validity of the proof, participants objected to the omission of the interval between x_1 and x_2. As CY explains:

> CY: I mean, it's a pedagogical choice.… If I were writing a textbook, I would write it this way, because I want to remind them what the Mean Value Theorem says. And my pedagogical purpose is served by reminding them that x_3 is in between x_1 and x_2.

For CY, the reason for the revision concerns student perception of the Mean Value Theorem. TR remarked the author of Task B was "missing the point" of the Mean Value Theorem by not specifying that x_3 would be located between x_1 and x_2.

Rephrased for Clarity. The following phrases received extensive alterations. In Task A, participants altered the phrase "We may assume without loss of generality…" and in Task B, participants altered the sentence beginning with "There is a unique line.…" Though participants tended to alter these phrases without stating their reasons, IR and CY described their motivation in terms of pedagogy. For instance, IR worried about the phrase "we may assume without loss of generality":

> IR: So I don't know how comfortable are people using the sentence, "We may assume without loss of generality." OK? So, I mean in my class I would say, "Say [$x_1 < x_2$]."

These revisions reveal uneasiness about mathematically correct ideas expressed in ways unfamiliar to students.

Reformulated an Algebraic Expression. Six participants revised Task A so the numerator and denominator of the slope formula were positive. Their reasons concerned logical structure as well as aesthetics. With respect to logical structure, FR altered the formula so that "we're making statements in the form that we want to be making them." Although he does not explicitly explain his intent here, we suggest that his edit can be interpreted as an intent to fit more closely with the next step of his proof: "By hypotheses, $f'(y) > 0$, and $x_2 - x_1 > 0$, so $f(x_2) - f(x_1) > 0$."

Two mathematicians cited aesthetics. For instance, CY supported his aesthetic judgment with a comment about mathematical writing and the meaning of slope.

> CY: I guess I tend to pick x_2 bigger than x_1. I think that's fairly normal in mathematics, x_2 is usually the bigger number than x_1. Then it's phrased as a rise-over-run expression. And it works better to say $x_2 - x_1$ is positive, and I have all these terms are positive. In the previous incarnation, $x_1 - x_2$ is negative. It's just cleaner.

Altered the Typesetting of the Proof. Four participants revised the proof by taking important formulas or equations in the proof and displaying these on their own centered line, as in Figure 3. CY describes this process below, referring to applying a "double dollar sign" to the equation. (Placing the characters "$$" before and after an expression in LaTeX, a common word processing software package used by mathematicians, would center the contents of that expression.)

> CY: I might put this [$x_1 < x_3 < x_2$ and the slope formula] out, you know, like double dollar sign … are we talking typesetting here?
> I: Sure.
> CY: Because that's the key here. You need to refer to it later. One way to highlight it is to double dollar sign it. So double dollar sign it.

Hence these participants used formatting to emphasize the main ideas of a proof. We note that Konior (1993) discussed indentations as a common contextual cue in standard mathematical proof presentations, where one of the aims was to add emphasis to parts of the proof.

The alterations regarding formatting and the choice of x_1 and x_2 so that $x_2 > x_1$ may be regarded as expressions of "meta-representational competence" (diSessa, 2002)—skills and judgments concerning how best to design external representations. In this case, the mathematicians are making judgments concerning how best to facilitate students' comprehension of a written proof by selecting particular ways to display it. In particular, the elementary arithmetic fact that 2 > 1 mnemonically supports the choice of $x_2 > x_1$, and choosing to work with the positive difference $x_2 - x_1$ allows more direct connection to the "rise-over-run" characterization of slope.

Revisions Involving Deletions

There were a total of 37 deletions. There were 13 instances in which participants deleted statements for irrelevance (e.g., IR deleted the expression "$x_1 - x_2 < 0$" from Task B because "it brings us nowhere… so it should go away") or because they felt the statement should be obvious to the readers of the proof (e.g., CY deleted "$f(x_2) - f(x_1) < 0$" because it was "not necessary for students. They shouldn't be struggling with number line stuff. Although I would be prepared to be surprised").

Five participants thought the awkward phrasing of some statements might confuse students. For instance, six participants deleted the sentence "If for some x_1, x_2 in **R**, we had $f(x_1) = f(x_2)$, then by contrapositive, $x_1 = x_2$." Of those six, three participants noted they deleted this because it could be very confusing to students.

Four participants removed statements that they believed (correctly) were mathematically incorrect. Three of these participants deleted "$x_2 - x_1 < 0$" from Task A for this reason, as it did not follow validly from the assumption $x_1 < x_2$. (The other five participants deleted this assertion as well, either for unstated or other reasons.)

Broader Issues

The Role of Audience. In the open-ended interview questions given to participants after they completed the two revision tasks, all eight participants remarked on the importance of the audience when presenting or revising a proof. All participants said the audience determined what statements could be treated as established knowledge and what statements would need to be justified; five participants claimed they strove to present proofs using techniques familiar to students.

Three participants indicated they found the revision task ill-defined, as they did not know exactly who would be reading the proof. When FR was asked about the most difficult part about revising a proof, he replied:

> FR: For me what was hard was that this is a sort of imaginary audience. I felt like if I had actually been preparing this for a class, I would have kind of known what things I could gloss over and what things needed justification. So I wasn't really quite sure where to tailor it and what level of detail should be there.

At least for these participants, this type of revision task would not be consistent with how they would prepare a proof for their lectures since they presumed they would have a better sense of what was covered in their class and what the students knew.

The Importance of Medium. A second issue that emerged in the open-ended interview questions concerned the medium in which the proof was presented. Five participants discussed how proof revision and presentation might differ between the contexts of lecture and textbook. In general, these participants noted that in lecture, an important goal of proof is to present the big picture or intuitive idea behind the proof. Details about the proof matter less, as an instructor can base expository decisions on student questions or facial expressions. In contrast, textbooks must be more complete and precise because, as TR comments, "you put it out there and then you have absolutely no control over how anyone responds or understands it, so you want to be super careful what you're writing in a textbook." The participants in this study treated the task of proof revision as if they were revising a textbook, consequently the task does not necessarily indicate how they would prepare for their lectures.

Discussion of Revisions

Many of the participants' revisions satisfied one of three broad purposes. First, participants made the logical structure of the proof more explicit by adding introduction and conclusion sentences. Second, participants strove to make the main ideas of the proof more transparent to the reader by typesetting to emphasize important steps in the proof and introducing important equations that get at the "heart" or "crux" of the proof. Third, participants aimed to assuage students' confusion with a proof by adding justifications for steps that students might find non-trivial, rewording phrases with unfamiliar terminology, emending confusing computations, and eliminating irrelevant or unnecessary statements that might be distracting. We note that these changes do not affect the proofs' validity but rather their pedagogical quality.

STUDY 2

Rationale for Study 2

Based on the analysis in the previous section, we generated the following hypotheses for features mathematicians would value in a pedagogical proof:

(H1) The quality of a pedagogical proof can be improved by adding introductory and concluding sentences that highlight the proof's framework.

(H2) The quality of a pedagogical proof can be improved by highlighting the main ideas of a proof with the judicious use of formatting.

(H3) The quality of a pedagogical proof can be improved by adding justifications for statements that undergraduates may have difficulty justifying.

(H4) The quality of a pedagogical proof is lessened by adding unnecessary or irrelevant assumptions or computations.

These hypotheses were generated based on the results from Study 1, which employed a small sample size and asked participants to revise proofs rather than judge the effectiveness of a particular revision. Study 2 was designed to test these four hypotheses with a larger sample size and specified revisions.

Method

Participants. We recruited mathematicians to participate in this study as follows. Thirty secretaries from mathematics departments in the United States were contacted and asked to distribute an e-mail to the mathematics faculty, post-docs, and PhD students of that department. These mathematics departments had strong national reputations. A total of 110 mathematicians (20 faculty, 13 post-docs, and 77 PhD students) agreed to participate. The text of the e-mail to the participants was:

> I am currently running a new internet study on the ways in which mathematical proofs may be modified in order to make them more understandable to second or third year undergraduate math students. In this study participants are shown a short proof and a sequence of variations of this proof. The task is to evaluate

the extent to which each new variation is an improvement of the original version for pedagogical purposes.

Validity of Internet-Based Studies. We followed the methodology employed by Inglis and Mejía-Ramos (2009) to maximize our sample size by collecting data through the internet. Reips (2000) notes one practical threat to the validity of internet-based studies is multiple submissions from the same individual. We adopted the strategy advocated by Reips, and implemented by Johnson-Laird and Savary (1999), and logged the IP address of each participant and the time they submitted their response. Under the assumption that each IP address was associated with a unique individual, these data were used to screen for possible cases of multiple submissions.

The validity of internet-based experiments was studied by Kranz and Dalal (2000), who compared 20 internet-based studies with their laboratory equivalents and found a "remarkable degree of congruence" between the methodologies. A similar conclusion was reached by Gosling, Vazire, Srivastava, and John (2004). These findings suggest that internet data has comparable validity to traditional data. Given these findings, our adherence to Reips' (2000) guidelines, and the impracticality of obtaining large samples of research-active mathematicians in any other fashion, we believe our methods were justified.

Materials. Participants were shown a Master Proof (presented in Figure 4) and five proofs (M1, M2, M3, M4, and M5—presented in Figures 5, 6, 7, 8, and 9) with minor modifications to the Master Proof highlighted in blue (these parts are boxed in the figures; they were not boxed in the experiment). The participants were asked to judge whether each of these modifications increased or decreased the pedagogical value of the proof.

***** Insert Figures 4, 5, 6, 7, 8, and 9 here *****

The first proof, M1, was used to test our first hypothesis, H1. It added an introductory and concluding sentence to the Master Proof to make explicit the proof framework being employed. M2 was used to test H2 by reformatting two important formulas in the proof to highlight their significance. M3 was used to test H3 by adding a justification in the proof that we felt might be difficult for students to infer. If H1, H2, and H3 were correct, we would expect mathematicians to view M1, M2, and M3 respectively as improvements to the proof. M4 and M5 were used to test H4. M4 added an irrelevant calculation to the proof, while M5 added an unnecessary assumption before applying the Mean Value Theorem. If H4 were correct, we would expect mathematicians to view M4 and M5 as lessening the pedagogical quality of the proof. We included two items, M4 and M5, to test the possibility that participants would view any additions or changes to the text positively.

Procedure. When participants visited the study website, they were first asked to indicate whether they were a PhD student, a post-doc, or mathematics faculty, as well as their level of undergraduate teaching experience. They were then shown the Master Proof and told that they would be shown modifications to the Master Proof; they would then be asked to judge whether the changes made the proof "less or more understandable to a second- or third-year undergraduate student." Next, the participant was presented with a screen containing the Master Proof at the top of the screen and a modified proof beneath that. They were asked "How do the changes in blue affect the pedagogical quality of the proof?" and were given the options "significantly better," "somewhat better," "the proofs were the same," "somewhat worse," or "significantly worse." We coded responses as 2, 1, 0, -1, and -2 respectively. This process was repeated until the participants evaluated all five modified proofs. The order in which the modified proofs were presented was randomized by participant. Along with their evaluations, the participants were given the option of commenting on their choices via a free response text box. For each modified proof, we performed an open coding of the participants' responses in the same manner in which we coded participants' revisions in Study 1.

RESULTS OF STUDY 2

A repeated measures ANOVA revealed a main effect on participants' evaluations by which proof the participants evaluated ($F(327, 4) = 231.7$, $p < 0.001$), indicating participants did not all believe the modifications were of equal quality. ANOVAs comparing the status (faculty, post-doc, PhD) and degree of teaching experience of the participants with their response patterns did not yield significant effects ($F < 1$, $p > 0.5$) in both cases, indicating that there was not a significant difference between how mathematics faculty members, post-docs, and PhD students performed on this task. Table 3 summarizes the main quantitative results from this study.

***** Insert Table 3 About Here *****

M1: Making Explicit the Proof Framework

The results summarized in Table 3 confirm H1—participants overwhelmingly believed that adding an introductory sentence that makes the framework of a proof explicit improves the pedagogical clarity of the proof. A summary of participants' comments is presented in Table 4. Of the 36 participants who evaluated the proof positively and volunteered to leave comments, 30 cited the benefit of giving students a "roadmap" to the proof and communicating explicitly to the students what was being proved. Eight participants argued that the first sentence was pedagogically good because it reminded the participants of the definition of injectivity. These comments also reveal that some participants did not view the concluding sentence as an improvement, suggesting the perceived benefit of the changes came from adding the introductory sentence.

***** Insert Table 4 About Here *****

M2: Displaying Significant Equations

The results in Table 3 confirm H2—most participants believed that re-formatting the proof by centering significant equations on their own line improved the pedagogical quality of the proof. A summary of participants' comments is presented in Table 5. Of the 31 participants who evaluated the proof positively and volunteered to leave comments, 21 cited the benefits of making the proof less cluttered and easier to read, while 10 cited the benefits of emphasizing the important equations of the proof.

***** Insert Table 5 About Here *****

M3: Supplying Justification for Concluding Inference

The results summarized in Table 3 failed to confirm H3. As Table 3 illustrates, many participants ($N = 41$) believed adding the extra justification improved the Master Proof. However, nearly an equal number of participants ($N = 40$) felt the addition of the extra justification made the proof worse. A summary of the comments that participants left for both the positive and negative evaluations is presented in Table 6. Eight participants commented that they approved of the modification, believing the extra justification might prevent confusion. However, seven participants who disapproved of the modification believed that the justification unnecessarily lengthened the proof; they argued that undergraduates should be able to make this inference on their own. This suggests that there may not be general agreement among mathematicians as to what types of inferences undergraduates are capable of making.

***** Insert Table 6 About Here *****

A post-hoc analysis reveals that participants' position and experience did not have an effect on their evaluation of M3. Table 7 presents participants' evaluation of M3 as a function of their position, and reveals that roughly an equal number of participants in each status group (PhD students, post-docs, and faculty) thought M3 increased and decreased the pedagogical quality of the proof. Post-hoc between-group ANOVAs comparing the status of the participants and their levels of teaching experience on their evaluations of M3 did not yield significant effects ($F < 1$, $p > 0.5$ in both cases). One plausible interpretation of these results is that tertiary teaching experience does not affect mathematicians' perceptions of what types of inferences a generic undergraduate is capable of making nor does it lead to a consensus on such judgments.

***** Insert Table 7 About Here *****

M4: Additional Algebraic Justification

The results summarized in Table 3 confirm H4—participants overwhelmingly believed the inclusion of the extra justification made the proof worse. A summary of the 40 comments left by participants who evaluated M4 negatively are presented in Table 8. Most of these participants noted that the included change made the proof longer unnecessarily, while 15 participants commented that the new equation was likely to be distracting or confusing for the reader.

***** Insert Table 8 About Here *****

M5: Adding a Reference to an Assumption

These results, presented in Table 3, also confirm H4—most participants viewed adding the assumption that f was a real-valued function in the proof diminished its pedagogical quality. A summary of the 40 comments left by participants who evaluated M5 negatively are presented in Table 9. Twenty of these comments note that adding this extra assumption could distract or confuse the students because it might lead them to try and make sense of why it was included.

***** Insert Table 9 About Here *****

Summary of Results

Study 2 revealed three results about the pedagogical perspectives of mathematicians. First, the results confirmed H1, H2, and H4, supporting our claims that mathematicians believe proofs for undergraduates should make the proof frameworks explicit, format important equations to highlight their importance, and avoid adding unnecessary calculations and assumptions. Second, the comments left by participants deepened our understanding of why mathematicians valued these features. For instance, although some participants commented that the formatting of equations in M2 highlighted their importance and illustrated the main ideas of the proof, even more participants preferred the formatting because it made the proof less condensed and easier to read, suggesting that spacing, and more generally, visual appearance are important aspects that mathematicians value in pedagogical proofs. The comments on M4 and M5 illustrate that one reason mathematicians prefer pedagogical proofs without unnecessary calculations or assumptions is that these may distract the reader from the main ideas of the proof and confuse the reader by encouraging him or her to try to understand why the calculations or assumptions were included.

Finally, the results of M3 illustrate the complexity of deciding when to add an extra justification to a proof. In a sense, the results of Study 1 and Study 2 are consistent. In Study 1, several participants suggested adding justifications to a proof to bridge logical gaps that students might find challenging. In Study 2, a sizeable minority of participants (41 out of 110 participants, or 37%) viewed the added justification in M3 as improving the pedagogical quality of the proof. However, a nearly equal number of participants believed adding this justification diminished the pedagogical quality of the proof, with some commenting that this inference would be obvious to the audience. This suggests that mathematicians may strongly disagree about what level of detail is optimal in the proofs that they present to students in part because they do not agree on what would be obvious to a generic student.

DISCUSSION

While the teaching of elementary mathematics has been widely analyzed and the knowledge and dispositions of elementary mathematics teachers have been the focus of many research studies, there has been limited research on the pedagogical practices of mathematicians. Recently, researchers have been aiming to fill this void by interviewing mathematicians about their beliefs on proof and pedagogy (e.g., Alcock, 2010; Harel & Sowder, 2009; Hemmi, 2010; Nardi, 2008; Weber & Mejía-Ramos, 2011; Yopp, 2011). This paper contributes to the literature by using qualitative and quantitative studies to investigate the features mathematicians value in a pedagogical proof. Main findings from this study are that mathematicians believe the quality of a pedagogical proof can be improved by adding an introductory sentence that makes transparent the proof framework being employed, highlighting the main ideas of the proof via typesetting, and eliminating irrelevant information that might distract or confuse students.

These findings raise two interesting issues for future research. First, it would be worthwhile to connect the mathematicians' revisions to their potential influences on student learning. Mathematics educators have yet to develop evidence-based frameworks to assess the pedagogical quality of proof presentations, making it impossible for us to assess whether the changes that the participants in our studies represent good pedagogy.[2] For example, it would be interesting to examine the ways in which the revisions suggested by mathematicians might affect students' comprehension or use of the proof.

Second, the participants in these studies revised written proofs that were given to them or evaluated changes generated by researchers. However, in preparing to teach, mathematicians are more likely to generate a proof that they will present verbally, often from scratch. Although the data in this paper reveal useful insights into what features mathematicians value in pedagogical proofs, the disconnect between this research and mathematicians' practice leaves open the question of how these mathematicians' values influence their teaching. Examining the relationship between desirable features of pedagogical proofs and classroom practice would fill a gap in the literature of studying pedagogical practice in situ (see Speer, Smith, & Horvath, 2010).

In Study 2, there was no consensus as to whether adding a specific justification improved the quality of a pedagogical proof, in part due to disagreement on how difficult it would be for students to make this justification on their own. Further, participants' judgments did not vary as a function of their experience or professional status, which suggests that mathematicians might not learn about students' abilities from their teaching experience. This is perhaps not surprising, as recent research on learning through teaching finds that teachers learn about students' thinking when their curricula is conceptually-oriented, allows for student creativity, and encourages student contributions. The lectures that dominate collegiate mathematics classes do not provide teachers with the opportunity to observe student contributions that could challenge inaccurate conceptions the mathematician may have about students' abilities (see Leikin & Zazkis, 2010; Weber & Rhoads, 2011).

The participants in Study 1 admirably noted the important role of the audience in deciding how to revise or present a proof. The results of Study 2 suggest that mathematicians may have differing beliefs about what inferences are easy or difficult for students to make. Accurate perception of student knowledge is part of mathematical knowledge for teaching, a construct shown to have positive correlation with student outcomes in the elementary level (Hill, Rowan, & Ball, 2005; Rockoff, Jacob, Kane, & Staiger, 2011). In the context of secondary mathematics, Baumert et al. (2010) studied the relationship between instruction and pedagogical content knowledge, which includes knowledge of students. Although only one component of their conception of quality of instruction concerned explanation, they did find a positive relationship between pedagogical content knowledge, quality of instruction, and student achievement. Analogously, studies exploring mathematicians' views about what students know and how they reason, as well as the accuracy of these beliefs and how these beliefs may evolve, are important avenues for future research.

ACKNOWLEDGEMENTS

We are grateful to the editor and the anonymous reviewers for helpful comments, to Aron Samkoff for transcribing interviews in the first study, and to the mathematicians who participated in the two studies. This work was supported by grants from the National Science Foundation (#EHR-1008317, #DUE-081736, and #DRL-0643734) and by an Investigating Student Learning Grant from the Center for Research on Learning and Teaching at the University of Michigan.

References

Alcock, L. (2009). e-Proofs: Student experience of online resources to aid understanding of mathematical proofs. *Proceedings of the 12th Conference for Research in Undergraduate Mathematics Education*. Retrieved from http://sigmaa.maa.org/rume/crume2009/proceedings.html

Alcock, L. (2010). Mathematicians' perspectives on the teaching and learning of proof. *Research in Collegiate Mathematics Education*, *6*, 63-92.

Alibert, D. & Thomas, M. (1991). Research on mathematical proof. In D. Tall (Ed.) *Advanced mathematical thinking* (pp. 215-230). Dordrecht, The Netherlands: Kluwer.

Ball, D. L. (1988). *Knowledge and reasoning in mathematical pedagogy: Examining what prospective teachers bring to teacher education* (Unpublished doctoral dissertation). Michigan State University, East Lansing.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., …Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Education Research Journal*, *47*(1), 133-180.

Charalambous, C. Y., Hill, H. C., & Ball, D. L. (2011). Prospective teachers' learning to provide instructional explanations: How does it look and what might it take? *Journal of Mathematics Teacher Education*, *14*, 441-463.

Davis, P. J. & Hersh, R. (1981). *The mathematical experience*. New York, NY: Viking Penguin Inc.

de Villiers, M. D. (1990). The role and function of proof in mathematics. *Pythagoras, 24*, 17-24.

diSessa, A. A. (2002). Students' criteria for representational adequacy. In K. Gravemeijer, R. Lehrer, B. van Oers, & L. Verschaffel (Eds.), *Symbolizing, modeling and tool use in mathematics education* (pp. 105-129). Dordrecht, The Netherlands: Kluwer.

Dreyfus, T. (1991). Advanced mathematical thinking processes. In D. Tall (Ed.) *Advanced Mathematical Thinking* (pp. 25-41). Dordrecht, The Netherlands: Kluwer.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six studies about internet questionnaires. *American Psychologist*, *59*, 93-104.

Griffiths, P. A. (2000). Mathematics at the turn of the millennium. *American Mathematical Monthly*, *107*, 1-14.

Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange*, *21*(1), 6-13.

Harel, G. (1998). Two dual assertions: The first on learning and the second on teaching (or vice versa). *American Mathematical Monthly*, *105*, 497-507.

Harel, G. & Sowder, L. (1998). Students proof schemes. *Research in Collegiate Mathematics Education, 3*, 234-282.

Harel, G. & Sowder, L. (2009). College instructors' views of students vis-a-vis proof. In M. Blanton, D. Stylianou, & E. Knuth (Eds.) *Teaching and learning proof across the grades: A K-16 perspective* (pp. 275-289). New York, NY: Routledge/Taylor & Francis.

Healy, L. & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education*, *31*, 396-428.

Hemmi, K. (2010). Three styles characterising mathematicians' pedagogical perspectives on proof. *Educational Studies in Mathematics*, *75*, 271-291.

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

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction, 26*, 430-511.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, *42*, 371-406.

Inglis, M. & Mejia-Ramos, J. P. (2009). The effect of authority on the persuasiveness of mathematical arguments. *Cognition and Instruction*, *27*, 25-50.

Inoue, N. (2009). Rehearsing to teach: Content-specific deconstruction of instructional explanations in preservice teacher training. *Journal of Education for Teaching*, *35*(1), 47-60.

Johnson-Laird, P. N. & Savary, F. (1999). Illusory inferences: A novel class of erroneous deductions. *Cognition, 71*, 191-229.

Kinach, B. M. (2002). Understanding and learning-to-explain by representing mathematics: Epistemological dilemmas facing teacher educators in the secondary mathematics ‘‘methods’’ course. *Journal of Mathematics Teacher Education*, *5*, 153-186.

Kline, M. (1973). *Why Johnny can’t add: The failure of the New Mathematics*. New York, NY: St. Martin’s Press.

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

Kranz, J. & Dalal, R. (2000). Validity of web-based psychological research. In M. Birnbaum (Ed.) *Psychological experiments on the internet* (pp. 35-60). San Diego, CA: Academic Press.

Larreamendy-Joerns, J., Muñoz, T. (2010). Learning, identity, and instructional explanations. In M. K. Stein & L. Kucan (Eds.), *Instructional explanations in the disciplines* (pp. 23-40). New York, NY: Springer.

Leinhardt, G. (1989). Math lessons: A contrast of novice and expert competence. *Journal for Research in Mathematics Education*, *20*, 52-75.

Leinhardt, G., Putnam, R. T., Stein, M. K., & Baxter, J. (1991). Where subject knowledge matters. In J. Brophy (Ed.), *Advances in research on teaching* (Vol. 2, pp. 87-113). London, UK: JAI Press Inc.

Leikin, R. & Zazkis, R. (2010). *Learning through teaching mathematics*. Dordrecht, The Netherlands: Springer.

Leron, U. (1983). Structuring mathematical proofs. *American Mathematical Monthly*, *90*(3), 174-184.

Leron, U. & Dubinsky, E. (1995). An abstract algebra story. *American Mathematical Monthly*, *102*(3), 227-242.

Lo, J., Grant, T., & Flowers, J. (2004). Developing mathematics justification: The case of prospective elementary school teachers. In D. McDougall & J. Ross (Eds.), *Proceedings of the Twenty-Sixth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 1159–1166). Toronto, Ontario, Canada: University of Toronto.

Mejia-Ramos, J. P. & Inglis, M. (2009). Argumentative and proving activities in mathematics education research. In F.-L. Lin, F.-J. Hsieh, G. Hanna & M. de

Villiers (Eds.), *Proceedings of the ICMI Study 19 Conference: Proof and Proving in Mathematics Education* (Vol. 2, pp. 88-93). Taipei, Taiwan.

Mills, M. (2011). Mathematicians' pedagogical thoughts and practices in proof presentation. In S. Brown, S. Larsen, K. Marrongelle, & M. Oehrtman (Eds.), *Proceedings of the 14th Conference for Research in Undergraduate Mathematics Education* (Vol. 2, pp. 283-297). Portland, OR.

Nardi, E. (2008) *Amongst mathematicians: Teaching and learning mathematics at university level*. New York, NY: Springer Science.

Porteous, K. (1986). Children's appreciation of the significance of proof. In Univ. of London Institute of Education (Eds.), *Proceedings of the Tenth International Conference of the Psychology of Mathematics Education* (pp. 392-397). London, England.

Reips, U. D. (2000). The web experiment method: Advantages, disadvantages, and solutions. In M. H. Birnbaum (Ed.) *Psychological experiments on the internet*. San Diego, CA: Academic Press.

Rockoff, J. E., Jacob, B., Kane, T. J., & Staiger, D. O. (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy 6*(1), 43-74.

Rowland, T. (2001). Generic proofs in number theory. In S. Campbell and R. Zazkis (Eds.) *Learning and teaching number theory: Research in cognition and instruction* (pp. 157-184). Westport, CT: Ablex Publishing.

Samkoff, A., Lai, Y., & Weber, K. (in press). On the different ways that mathematicians use diagrams in proof construction. *Research in Mathematics Education.*

Schoenfeld, A. H. (2010). How and why do teachers explain things the way they do? In M. K. Stein & L. Kucan (Eds.), Instructional explanations in the disciplines (pp. 83-106). New York, NY: Springer.

Selden, J. & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, *29*(2), 123-151.

Selden A. & Selden J. (2003) Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 34*(1) 4-36.

Speer, N., Smith, J. P., III, & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *Journal of Mathematical Behavior, 29*(2), 99-114.

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

Thanheiser, E. (2009). Preservice elementary school teachers' conceptions of multidigit whole numbers. *Journal for Research in Mathematics Education*, *40*, 251-281.

Thanheiser, E. (2010). Investigating further preservice teachers' conceptions of multidigit whole numbers: Refining a framework. *Educational Studies in Mathematics*, *75*, 241-251.

Thompson, P. W., & Thompson, A. G. (1994). Talking about rates conceptually, part I: A teacher's struggle. *Journal for Research in Mathematics Education*, *25*, 279-303.

Thompson, A. G., & Thompson, P. W. (1996). Talking about rates conceptually, part II: Mathematical knowledge for teaching. *Journal for Research in Mathematics Education*, *27*, 2-24.

Thurston, W. P. (1994). On proof and progress in mathematics, *Bulletin of the American Mathematical Society*, *30*, 161-177.

Weber, K. (2004). Traditional instruction in advanced mathematics classrooms: A case study of one professor's lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior*, *23*, 115-133.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39*(4), 431-459.

Weber, K. & Alcock, L. (2009). Semantic and syntactic reasoning and proving in advanced mathematics classrooms. In M. Blanton, D. Stylianou, and E. Knuth (Eds.) *The Teaching and Learning of Proof across the K-16 Curriculum* (pp. 323-338). New York, NY: Routledge/Taylor & Francis.

Weber, K. & Mejía-Ramos, J. P. (2011). Why and how mathematicians read proofs: An exploratory study. *Educational Studies in Mathematics, 76*(3), 329-344.

Weber, K. & Rhoads, K. (2011). Review of learning through teaching mathematics. *Journal for Research in Mathematics, 42*(5), 521.

Yopp, D. (2011). How some research mathematicians and statisticians use proof in undergraduate mathematics. *Journal of Mathematical Behavior*, *30*, 115-130.

FOOTNOTES

1. For purposes of brevity, we did not discuss “introducing a variable for the slope of a line,” “adding a word for emphasis,” or “adding a picture” for added revisions or “renaming a variable” for altered revisions. Mathematicians’ use of pictures in Study 1 is discussed in Samkoff, Lai, and Weber (in press).

2. This situation differs from elementary mathematics in which there are frameworks that researchers use to analyze the quality of instruction (e.g., Hill et al, 2008).

Captions for Tables and Figures

TABLE 1 Summary of Data for Added Revisions from Study 1

Category Name	Number of revisions (Number of participants making such a revision)
Added introductory or concluding sentence	11 (7)
Included an extra formula or justification	16 (8)
Introduced slope formula	3 (3)
Introduced variable for slope	3 (3)
Added picture	1 (1)
Added a word for emphasis	1 (1)

TABLE 2 Summary of Data for Altered Revisions from Study 1

Category name	Number of revisions (Number of participants making such a revision)
Altered main hypothesis	9 (8)
Changed reference to Mean Value Theorem	8 (6)
Altered phrasing	13(8)
Reformulated algebraic expression	8 (6)
Altered the typesetting of the proof	4 (4)
Renamed variables	7 (5)

TABLE 3 Summary of Data from Study 2

Condition	Mean score	Number of participants who thought proof was better	Number of participants who thought proof was worse
M1	1.29*	97	4
M2	1.05*	88	2
M3	0.02	41	40
M4	-1.66*	6	98
M5	-1.12*	7	94

* Indicates a mean score statistically different than zero with $p < 0.001$.

Note. Participants who claimed the original and modified proofs had the same pedagogical value were not included as either participants who thought the proof was better or who thought the proof was worse.

TABLE 4 Comments from Participants Who Evaluated M1 Positively

Type of response	Number of responses (36 total)	Representative responses
It is beneficial to state the objectives of the proof	30	“Stating what is to be proven is almost always helpful to students,” “I think any proof benefits from a summary of what needs to be shown,” “Roadmaps. Great.”
The second sentence was less helpful than the first	5	“The last blue part is too much, I think,” “The last sentence is superfluous”
It is useful to remind students of the definition of injectivity	8	“Yes! Repeat definitions like mad! Drill them into their heads!”; “It’s always nice to repeat the definition”

Note. Some responses were assigned to more than one category.

TABLE 5 Comments from Participants Who Evaluated M2 Positively

Type of response	Number of responses (31 total)	Representative responses
The added spacing makes the proof less cramped and easier to read	21	“Formulae on their own lines makes reading proofs easier,” “Easier to read,” “The original seemed cramped,” “This is easier on the eyes, making the proof psychologically more appealing, I think”
The formatting draws the reader’s attention to the important ideas of the proof	10	“I think this makes the proof much clearer. Giving these equations their own line emphasizes that this is where the main points in the proof are happening.”; “Displayed equations draw the eye to important claims”
Other	5	“Now the proof, once understood, can be recalled at a glance”

Note. Some responses were assigned to more than one category.

TABLE 6 Comments from Participants Who Evaluated M3 Positively or Negatively

Favorable evaluation comments (N = 10)

Type of response	Number of responses	Representative responses
Including the extra calculation may prevent potential confusion about the last step	8	“I can imagine a student being confused by the last step and this change would make it clearer,” “The new statement is redundant, but it saves time in thinking about why it works”
Other	2	“The extra explanation helps a bit, but it makes the proof bulkier”

Unfavorable evaluation comments (N = 15)

Type of response	Number of responses	Representative responses
The audience is capable of making this calculation so it unnecessarily lengthens the proof	7	“Maybe ‘worse’ sounds harsh, but it adds more words and if you have proved the Mean Value Theorem, then a deduction like this should be straightforward,” “Clutter by repetition”
The added justification is not written well	6	“I think you should say f(x2) – f(x1) is greater than zero, therefore f(x1) <> f(x2),” “Not a good way to conclude the proof”
Other	2	“I suspect it takes more to read the inserted line than to realize what is being pointed out. It would take a study of actual undergrads to make sure.”

Note. Some responses were assigned to more than one category.

TABLE 7 Participants' Response to M3 Sorted by Position

Position	Mean score	Number of participants who thought proof was better	Number of participants who thought proof was worse
Faculty (*N* = 20)	-0.10	6	7
Post-doc (*N* = 17)	-0.08	5	4
PhD student (*N* = 77)	0.04	30	29

Note. Participants who claimed the original and modified proofs had the same pedagogical value were neither included as participants who thought the proof was better nor who thought the proof was worse.

TABLE 8 Comments from Participants Who Evaluated M4 Negatively

Type of response	Number of responses (40 total)	Representative responses
The extra calculations unnecessarily lengthen the proof	31	"The additional formula makes it longer but not clearer," "That isn't the manipulation used at the end, why do it?"; "The addition seems irrelevant"
Including the extra calculation will distract or confuse the student	15	"The blue material just adds another equation for the poor reader to parse," "Students may lose time and energy wondering why this line is there and/or develop misunderstandings"
The added expression divides by $f'(x_3)$ before justifying that $f'(x_3) \neq 0$	4	"This comes before $f'(x_3) > 0$ in the argument so you shouldn't divide through without indicating reasoning"
Other	3	"This takes the argument on a weird path"

Note. Some responses were assigned to more than one category.

TABLE 9 Comments from Participants Who Evaluated M5 Negatively

Type of response	Number of responses (40 total)	Representative responses
The assumption is superfluous for this part of the proof	16	"The inserted detail is not pertinent to that area of the proof so I don't see any benefit to pointing it out," "This is unnecessary," "That's not the issue here, honestly"
Including this assumption is distracting and potentially confusing	20	"The *f* was real-valued distracts from the main point," "This statement is distracting to the reader," "The blue phrase diverts attention away from the salient hypotheses"
It was implicitly understood throughout the proof that *f* was real-valued	21	"We are presumably referring to f being real-valued to appeal to properties of an ordered field. But in this context, it's quite obvious so the change just adds words." "The students most likely assume all functions are real-valued"

Note. Some responses were assigned to more than one category.

FIGURE 1 Task A for revision.

Claim: If $f : \mathbb{R} \to \mathbb{R}$ is a differentiable function with strictly positive first derivative on its domain, then f is injective.

Proof: Let $x_1, x_2 \in \mathbb{R}$. We may assume without loss of generality that $x_2 > x_1$. If they were equal, we'd gain no information about injectivity.

By the Mean Value Theorem, there exists $x_3 \in \mathbb{R}$ with $x_1 \leq x_3 \leq x_2$ and $\frac{f(x_1)-f(x_2)}{x_1-x_2} = f'(x_3)$. By our hypotheses, $f'(x_3) > 0$ and $x_2 - x_1 > 0$. Equivalently, $x_1 - x_2 > 0$. So $f(x_1) - f(x_2) < 0$, in other words, $f(x_2) - f(x_1) > 0$.

If for some $x_1, x_2 \in \mathbb{R}$, we had $f(x_1) = f(x_2)$, then by contrapositive, $x_1 = x_2$. If not, then the above argument implies $f(x_2) \neq f(x_1)$.

FIGURE 2 Task B for revision.

Claim: If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and for all $x \in \mathbb{R}$, $f'(x) > 0$, then f is injective.

Proof: Let $x_1, x_2 \in \mathbb{R}$ be two distinct elements of $\mathbb{R}$. Then there is a unique line passing through $(x_1, f(x_1))$ and $(x_2, f(x_2))$, two coordinate points along the graph of f.

The Mean Value Theorem gives an $x_3 \in \mathbb{R}$ so that the slope of the line from $(x_1, f(x_1))$ to $(x_2, f(x_2))$ is given by $f'(x_3)$. The point x_3 is in the domain of f.

By our hypothesis, $f'(x_3)$ is positive. So either $f(x_1) > f(x_2)$ or $f(x_2) > f(x_1)$. In both cases, $f(x_1) \neq f(x_2)$. Hence f is injective.

FIGURE 3 (a) CY's revision for Task A (b) TR's revision for Task B

(a) CY's revision for Task A

Claim: If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $f'(x) > 0 \; \forall x \in \mathbb{R}$, then f is injective.
Proof: Let x_1 and x_2 be distinct points in $\mathbb{R}$. We may assume w/o loss of generality that $x_2 > x_1$. We must prove that $f(x_1) \neq f(x_2)$.
The Mean Value Thm implies that there exists $x_3 \in [x_1, x_2]$ such that

$$f'(x_3) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Since, by hypothesis, $f'(x_3) > 0$ and $x_2 - x_1 > 0$,

$$f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0.$$

Therefore, $f(x_2) \neq f(x_1)$. So f is not injective.

(b) TR's revision for Task B

Claim. If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and for all $x \in \mathbb{R}$, $f'(x) > 0$, then f is injective.
Proof. Let $x_1, x_2 \in \mathbb{R}$ be two distinct elements of $\mathbb{R}$. Then there is a unique line passing through the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$ on the graph of f.
By the Mean Value Theorem, there is some x_3 in the interval $[x_1, x_2]$ so that the slope of the line L from $(x_1, f(x_1))$ to $(x_2, f(x_2))$ is equal to $f'(x_3)$.
By our hypothesis, $f'(x_3)$ is strictly positive, and in particular, nonzero. Therefore the slope of L, which is given by

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2}$$

is nonzero, and hence $f(x_1)$ is not equal to $f(x_2)$. Since x_1 and x_2 are arbitrary, it follows that f is injective.

FIGURE 4 Master proof for quantitative study.

Proposition. *If* $f : \mathbb{R} \to \mathbb{R}$ *is differentiable and* $f'(x) > 0$ *for all* $x \in \mathbb{R}$*, then* f *is injective.*

Proof. Let $x_1, x_2 \in \mathbb{R}$, where $x_2 > x_1$. The Mean Value Theorem implies there exists $x_3 \in [x_1, x_2]$ such that $f'(x_3) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$. Since, by hypothesis, $f'(x_3) > 0$ and $x_2 - x_1 > 0$, then $f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0$. Therefore $f(x_2) \neq f(x_1)$. □

FIGURE 5 Modified proof M1 (making explicit the proof framework).

Proposition. *If* $f : \mathbb{R} \to \mathbb{R}$ *is differentiable and* $f'(x) > 0$ *for all* $x \in \mathbb{R}$, *then* f *is injective.*

Proof. Let $x_1, x_2 \in \mathbb{R}$, where $x_2 > x_1$. To show f is injective, we must show that $f(x_1) \neq f(x_2)$. The Mean Value Theorem implies there exists $x_3 \in [x_1, x_2]$ such that $f'(x_3) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$. Since, by hypothesis, $f'(x_3) > 0$ and $x_2 - x_1 > 0$, then $f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0$. Therefore $f(x_2) \neq f(x_1)$. It follows that f is injective. □

FIGURE 6 Modified proof M2 (displaying significant equations).

Proposition. *If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $f'(x) > 0$ for all $x \in \mathbb{R}$, then f is injective.*

Proof. Let $x_1, x_2 \in \mathbb{R}$, where $x_2 > x_1$. The Mean Value Theorem implies there exists $x_3 \in [x_1, x_2]$ such that

$$\boxed{f'(x_3) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}}.$$

Since, by hypothesis, $f'(x_3) > 0$ and $x_2 - x_1 > 0$, then

$$\boxed{f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0}.$$

Therefore $f(x_2) \neq f(x_1)$. □

FIGURE 7 Modified proof M3 (supplying justification for concluding inference).

Proposition. *If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $f'(x) > 0$ for all $x \in \mathbb{R}$, then f is injective.*

Proof. Let $x_1, x_2 \in \mathbb{R}$, where $x_2 > x_1$. The Mean Value Theorem implies there exists $x_3 \in [x_1, x_2]$ such that $f'(x_3) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$. Since, by hypothesis, $f'(x_3) > 0$ and $x_2 - x_1 > 0$, then $f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0$. As $f(x_2) - f(x_1) \neq 0$, it follows that $f(x_2) \neq f(x_1)$. □

FIGURE 8 Modified proof M4 (additional algebraic justification).

Proposition. *If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $f'(x) > 0$ for all $x \in \mathbb{R}$, then f is injective.*

Proof. Let $x_1, x_2 \in \mathbb{R}$, where $x_2 > x_1$. The Mean Value Theorem implies there exists $x_3 \in [x_1, x_2]$ such that $f'(x_3) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$, so $x_2 - x_1 = \frac{f(x_2)-f(x_1)}{f'(x_3)}$. Since, by hypothesis, $f'(x_3) > 0$ and $x_2 - x_1 > 0$, then $f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0$. Therefore $f(x_2) \neq f(x_1)$. □

FIGURE 9 Modified proof M5 (adding a reference to an assumption).

Proposition. *If* $f : \mathbb{R} \to \mathbb{R}$ *is differentiable and* $f'(x) > 0$ *for all* $x \in \mathbb{R}$*, then* f *is injective.*

Proof. Let $x_1, x_2 \in \mathbb{R}$, where $x_2 > x_1$. The Mean Value Theorem implies there exists $x_3 \in [x_1, x_2]$ such that $f'(x_3) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$. Since, by hypothesis, f is a real valued function, $f'(x_3) > 0$ and $x_2 - x_1 > 0$, then $f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0$. Therefore $f(x_2) \neq f(x_1)$. □