What makes a good proof: Investigating mathematicians' proof revisions

Yvonne Lai (University of Michigan) & Keith Weber (Rutgers University)

Abstract. We observed ten mathematicians producing a proof and then revising two other proofs for pedagogical purposes and then interviewed them about their proof revision processes. We use this data to discuss what mathematicians believe makes a good proof for pedagogical purposes, how they revise proofs to achieve these goals, and the difficulties and tensions that they face when trying to present a proof that will be accessible to all students.

1. Introduction

There is a broad consensus amongst mathematics educators that proving is an essential activity in mathematical practice and that justifying and proving should play a central role in all mathematics classrooms (e.g., Schoenfeld, 1994; Hanna, 1995; NCTM, 2000; Yackel & Hanna, 2003). However, what the activity of proving entails, how mathematicians engage in this activity, and how students should engage in this activity remain important open research questions. Mejia-Ramos (2008) noted that proving contains three sub-activities: constructing a proof, presenting a proof, and reading a proof. Empirical research on proof in mathematics education predominantly focuses on the construction and evaluation of proofs; in contrast, there have been almost no studies on the presentation of proofs (Mejia-Ramos & Inglis, 2009). This paper focuses on this neglected area of research. In particular, we investigate how mathematicians revise the arguments of their colleagues for pedagogical purposes.

An individual may revise an argument with (at least) two different goals in mind. One may focus on the validity or correctness of the argument; this may entail checking the inferences and calculations within the argument and making appropriate corrections when an error is found, as well as amending typos. However, one may also seek to improve the quality of an argument that is already a valid proof. This may involve making alterations that improve the clarity or comprehensibility of the argument, reformulating the argument to highlight essential points, or rewriting the argument so it better conforms to the norms of how a proof should appear. Our investigations in this paper primarily concern this latter aim of revision.

In the study reported in this paper, we presented ten mathematicians with two proofs of the same mathematical assertion and asked them how they would revise the proof if they were to present it to a sophomore-level or junior-level course for a mathematics major. We then asked them to reflect upon their revision processes. We use this data to address the following questions:

- What types of changes do mathematicians make when revising a proof for pedagogical purposes?
- What motivations do mathematicians have for these changes?
- What can these revision processes tell us about mathematicians' beliefs about the communicative goals of proof (at least within a pedagogical setting)?
- What difficulties and tensions do mathematicians face when revising a proof?

2. Related literature

2. 1. The communicative functions of proof

In a seminal paper, de Villiers (1990) argued that proof serves many functions in mathematics beyond providing conviction that a theorem is true. Proof can also be used 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. Providing mathematicians with a shared language and standards of argumentation facilitates debate about sophisticated mathematical ideas. Since de Villiers' article, others have suggested other roles that proof may serve in the mathematical community, including exploration (Hanna, 2000) and illustrating new ways of reasoning (Hanna & Barbeau, 2008; Weber, 2010). Several researchers have lamented that students and teachers only view proof as a means of providing conviction or assessing students' logical abilities (e.g., Healy & Hoyles, 2000; Herbst & Brach, 2006; Knuth, 2002). Consequently proof plays only a limited role in mathematics classrooms.

A particularly important function of proof to mathematics educators is explaining why a theorem is true. Building on the work of Steiner (1978), Hanna (1990) distinguished between proofs that merely convince with proofs that convince but also explain. Many researchers argue that explanatory proofs have significantly more pedagogical value than proofs that merely convince (e.g., Hanna, 1990; Alibert & Thomas, 1991; Hersh, 1993) and should be given greater prominence in mathematics teaching.

This research has made a substantial contribution to the teaching and learning of proof but we believe that it has two limitations. First, we agree with Raman's (2003) claim that there is not a shared understanding amongst mathematics educators on what an explanatory proof is. Hanna (1990) cites Steiner (1978) who defines an explanatory proof as one that "makes reference to a characterizing property of an entity or structure mentioned in the theorem, such that from the proof it is evident that the results depends upon that property" (Steiner, 1978, p. 143), where a characteristic property is defined as a property that is required for the proof to work and, crucially, the proof could be deformed to prove new theorems by holding the proof idea constant and substituting in different characteristic properties. Such a definition is limited because it does not account for all the proofs that mathematics educators or philosophers would wish to call explanatory (such as a proof that reduces the abstract to the familiar) (e.g., Resnik & Kushner, 1987; Mancosu, 2000) and because it treats the issue of how explanatory a proof is as independent of the reader (i.e., it does not address to whom the proof might be explanatory for) (Weber, 2010). In practice, few mathematics educators reference Steiner's (1978) criteria when judging a proof to be explanatory. If teachers are to follow mathematics educators' advice and present more explanatory proofs, a clearer description of explanatory proofs is needed, with specific techniques that teachers can apply to make a proof more explanatory.

The second issue concerns the methodologies used to determine the functions of proof in mathematics. Researchers in this area often base their judgments on philosophical analyses (Steiner, 1978; Rav, 1999) and the introspections of individual mathematicians (e.g., Thurston, 1994), but rarely corroborate their conclusions with systematic empirical studies. While this research has advanced the field by producing plausible supportable hypotheses, it has also led to contradictory findings. For instance, Alibert and Thomas (1991) argue that the primary purpose of proofs is to explain. However, others contend the main purpose of proof is to introduce new methods (e.g., Rav, 1999). As Bressoud (1999) claims, "The value of a proof … should be judged, not by … its explanatory power, but by the extent to which it enlarges our toolbox" (p. 190). Similar disagreements arise considering the conviction that proof provides. Harel and Sowder (1998) claimed that, for mathematicians, proofs turn conjectures into facts (i.e., provide complete certainty that an assertion is true), while Otte (1994) and de Villiers (1990) argued that certainty in the veracity of a theorem cannot be obtained via proof alone. Investigating the behavior of mathematicians completing proof-related tasks could provide insight into what the role of proof actually is within the mathematical community.

2. 2. Proof as rhetorical genre

Some researchers have argued that proving can be understood profitably as a new language (Downs & Mamona-Downs, 2005), a particular form of discourse (Volminik, 1990), a literary or rhetorical genre (e.g., Selden & Selden, 2003; Nardi & Iannone, 2005), or an argument within a particular representation system (Weber & Alcock, 2009). Selden and Selden (2003) elaborated on this notion, remarking that there are some things that are rarely seen in published proofs, even though their occurrence would not affect the correctness of the proof. For instance, proofs usually contain little redundancy. When an assertion is made in a proof, it is usually made only once, even though repeating the assertion would not make the proof less valid.

Konoir (1993) and Weinberg (2009) contended that understanding the rhetorical norms used in proof writing are essential for understanding the proofs that appear in journals and textbooks. For instance, Konoir (1993) observed that a paragraph that appears in a different font, or with a different indentation, than the preceding text often indicates the presence that a sub-proof whose argument structure is independent from the rest of the proof. Weinberg (2009) argued that students are often unaware of these types of norms, which likely inhibits their comprehension of proofs in textbooks. Despite the importance of understanding the genre of proof-writing, little systematic research on the genre of proof-writing has been conducted.

2. 3. Proof presentation

In a literature review, Mejia-Ramos and Inglis (2009) analyzed all articles in the ERIC database that listed "argumentation" or "proof", as well as "mathematics", as key words that appeared in seven widely read journals in mathematics education. Of the 131 articles they considered, not a single article investigated the presentation of proof. Leron (1983) has proposed an alternative format for presenting proofs that does not alter the content or logic employed within the proofs, suggesting that the manner in which a proof is presented may influence its comprehensibility and the extent to which a proof will achieve its communicative goals.

3. Methods

3.1. Participants

We interviewed 10 mathematicians. The types of institutions represented include liberal arts schools, large state universities, and private schools. All mathematicians are referred to using masculine pronoun and by alias for anonymity. The mathematicians' years of experience, number of papers refereed, and research areas are listed below.

Mathematician Participants

alias	~years post PhD	~#papers refereed	research area
BT	30	40	analysis, mathematical physics
KT	25	180	combinatorics
CY	20	40	geometric topology
IR	10	5	geometric topology
TR	5	10	algebraic geometry
KM	5	10	algebraic topology
KZ	2	4	combinatorics
TL	2	3	complex dynamics
PV	1	4	representation theory
FR	1	2	operator algebras

3.2. Procedure and materials

We assigned to each participant the Sine Task and the MVT Tasks, described below. Following the tasks, we interviewed the participants about their experience completing the task.

3.2.1. Sine Task

This task was phrased as follows.

What is the clearest way to write down the proof of the following problem?

Show that restrictions of the sine function to intervals of length greater than pi cannot be injective.

Please think about this problem, and then write down a proof that is geared for a sophomore- or junior-level math major, and so that your solution is self-contained.

On the solution sheet provided for this task was a space labeled "Claim" followed by a space labeled "Proof". We asked participants to follow a think-aloud procedure during this task.

The purpose of this task was to record how mathematicians approached the task of revising a proof that they had just written. The statement to prove was chosen as one that would be swiftly evaluated as valid, unlikely to produce a proof that was primarily symbol pushing, and whose proof would likely to involve the process of turning informal intuition into mathematically formal language. The deliberately cumbersome phrasing of the mathematical task was intended to make revising the "Claim" statement part of the work of the mathematician. Participants were given a blue pen with which to write first draft, and then a red pen to make corrections. The interview questions following the Sine Task addressed the nature of their proof and potential factors in revising their proof:

- Could you describe, in a sentence or two, the key idea or ideas behind your proof?
- Did you try to emphasize or bring this out in any way?
- How did your revisions improve the clarity of the proof?
- Is there anything you would like to say about the notation or organization of the proof?
- What was your claim statement? (If it differed from the statement of the task) How does your claim statement improve upon the original statement?
- Would your presentation have differed if you had written it for a lecture instead of a textbook?
- Is there anything you find unsatisfying about your proof, or the way that you wrote it?
- Is there anything else you would like to say about the process of revising a proof for clarity?

3.2.2. Mean Value Theorem Tasks

Following the Sine Task, we introduced the Mean Value Theorem (from hereon MVT) Tasks. In the copies we gave the participants, the tasks labeled A and B below were unlabeled.

Consider the following 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.

You will be given two solutions of this question to revise.

Task A: Solution to Revise

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 $f(x_2) \neq f(x_1)$.

Task B: Solution to Revise

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

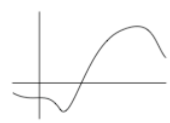
Proof: Let $x_1, x_2 \in \mathbb{R}$ be two distinct elements of $\mathbb{R}$. Then there is a unique line passing through $(x_1, f(x_1))$ and $(x_2, f(x_2))$, two coordinate points along the graph of f.

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.

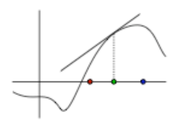
By our hypothesis, $f'(x_3)$ is positive. So either $f(x_1) > f(x_2)$ or $f(x_2) > f(x_1)$. In both cases, $f(x_1) \neq f(x_2)$. Hence f is injective.

Participants were given the following informal statement of the Mean Value Theorem.

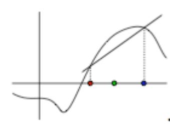
If you have a function from $\mathbb{R}$ to $\mathbb{R}$



that is differentiable, and ● and ● are elements of $\mathbb{R}$, then there's a value ● in between ● and ● so that the slope of the tangent line



is the same as the slope of the line



The purpose of this task was to record how mathematicians approached the process of revising a colleague's writing for a pedagogical purpose.

Features of Task A include a small error in the manipulations of the inequalities, to gauge how finely mathematicians would proofread a colleague's solution; a mathematically valid but extraneous statement at the end, to see how mathematicians would work with redundancy; a thwarting of what we perceived as the usual mathematical convention of expressing the denominator of the slope as the difference between the greater input and lesser input rather than vice versa; and minute arithmetic manipulations that were mathematically correct but we believed would not typically be found in a textbook, to see how mathematicians would treat correct but arguably unnecessary steps.

Features of Task B include mathematically correct non-arithmetic steps that we believed would typically be left out of a textbook ("there is a unique line", "x3 is in the domain of f", a proof that needed cases at the very end but not in the beginning, terminology that we believed was more commonly used in non-math major classes than in math major classes ("coordinate point"), and an example of less formal language than that of Task A ("the Mean Value Theorem gives").

Neither Task A nor Task B contained a sentence that concisely captured the heuristic used for proving a function is injective (given two distinct values in the domain, prove that their output values are distinct). We were curious to see if mathematicians would provide such a sentence, given that the task was framed as pedagogical.

From our perspective, Task A and Task B were two qualitatively different proofs. Task A appeared to be based on symbol manipulation—in particular, using algebraic manipulations to the statement in the Mean Valued Theorem to attain the desired results. Task B appeared based on geometric reasoning, making reference to the graph of the function and its secant lines. Following Weber and Alcock (2004), we argue that Task A would seem to be the result of a syntactic proof production and Task B would result from a semantic proof production. We were interested in whether mathematicians would see a connection or differences between these tasks.

Following the MVT Tasks, we interviewed the mathematicians on their experience using similar questions to those asked for the Sine Task. In addition, we asked the mathematicians:

- Would the exposition of either proof be improved with a picture?
- Would the exposition of the proof without the slope formula be improved with the inclusion of the slope formula?

3.3. Analysis

The analysis in this paper has two parts. We first look at specific revisions that participants made on the MVT tasks as well as their stated reasons for making these revisions. We then analyzed the interviews in their entirety to discuss broader issues on the process of revision.

3.3.1. MVT Task

In section 4, our analysis focuses predominantly on mathematicians' revisions to the MVT Tasks. We did not analyze revisions to work on the Sine Task, as many of the revisions were on partially constructed proofs with large amounts of the proof not stated. This is interesting in and of itself, but beyond the scope of this article. Two of the participants (KZ, TL) were not included in the analysis because they took over 50 minutes to complete the Sine Task and were forced to rush and only partially complete the MVT tasks.

To analyze the types of revisions that the remaining participants 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 MVT task, we made a general description of the change that was made, went through the transcript and noted the reason given for the change (if any), and developed category names. New episodes were placed into existing categories when appropriate, but also used to create new categories but also to modify the names or definitions of existing categories.

3.3.2. Broader Issues

To identify broader issues in the revision process, we adopted the methodology of Boaler, Ball, and Even (2003) of "moving from the particular to the general". We first noted what we considered to be key incidents in the Sine and MVT Tasks and the participants' corresponding reflections on their interview process and hypothesized about what was generic about that incident (i.e., not particular to the specific task or situation). If the themes that arose seemed evident in multiple participants, we read through the transcripts to analyze to what extent the issues were present in each interview. We discuss these broader issues in section 5.

4. Results

The eight participants collectively made 135 revisions. In Sections 4.1-4.3, we examine the motivations behind "nontrivial" revisions that *added* text, *altered* text, and *deleted* text. In Section 4.4, we discuss implications of the analysis in Sections 4.1-4.3 as well as broader themes that arose in our study of the revisions. (Here, "trivial" describes revisions such as substituting "Mean Value Theorem" with "MVT", or rewording "Let x … " with "Fix x …" when such revisions were made without remark by the participant. In general, a trivial revision was one where a participant replaced a phrase with an equivalent one, the revision passed unremarked in the think-aloud portion and the interview, and we believed the revision did not impact the meaning or emphasis of the original statement. There were 20 trivial revisions).

While coding the data, we found it useful to delineate the characteristics of each category; we list them here to help frame the analysis:

- *added* represents new 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* represents new text that substituted the original text, or new text that changes the structure of the proof. There were a total of 46 alterations.
- *deleted* represents eliminated text that the participant did not replace with new text. There were a total of 34 deletions.

Where we could, we used the mathematician's words to guide the coding. Some revisions had reasons were unstated. In these cases, we made a judgment call based on our interpretation of the role of the revision in the mathematical reasoning of the proof. The purpose of this coding scheme is not to devise precise categories for any possible revision of any mathematical work, but to yield a provisional way to organize the subject of our analysis: the motivation and nature of revisions for the tasks that we provided.

4.1. Revisions classified as *Added*

In this category, our participants discussed motivations for: adding introduction or conclusion sentence, introducing the slope formula or manipulating the slope formula, introducing notation, inserting a picture, and adding a word for emphasis.

4.1.1. Added introduction or conclusion sentence. This type of revision was very common, with 11 instances performed by 7 participants. To the participants, the purpose of these sentences was describing or summarizing how the proof accomplished its goals. Consider, for example, CY's revision of Task A and TR's work on Task B (emphasis ours):

CY's revision of 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 $x_2 > x_1$. $\boxed{\text{We must prove that } f(x_1) \neq f(x_2).}$

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, $f(x_2) \neq f(x_1)$. $\boxed{\text{So } f \text{ is not injective.}}$

TR's revision of 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 $f(x_2)$. $\boxed{\text{Since } x_1 \text{ and } x_2 \text{ are arbitrary, it follows that } f \text{ is injective.}}$

The boxed sentences typify the role of the introduction and conclusion sentences added by participants: making explicit the work of the proof. As TR said of his conclusion sentence (emphasis ours):

> 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.

Thus the participants' revisions can be viewed as revealing the logical structure that had previously been implicit in the two MVT Tasks. Although the logical structure of a proof can often be inferred by mathematicians 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. Indeed, two participants suggested that whether or not their added introductory sentence belonged in the revision would depend on the audience.

4.1.2. Manipulated or introduced slope formula. Five of the eight participants wrote out the following seemingly trivial manipulation of the quotient in Task A:

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

When explaining the motivation behind this manipulation, the participants' reasons fell into two categories: the formulas or manipulations were intrinsic to the ideas of the proof, or that there was a logical gap that needed patching. KT and IR, who both added the manipulation, represent these two viewpoints (emphasis ours):

> 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.

> 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.

CY echoed KT, characterizing the move as the "actual key arithmetic step that makes it work."

Similarly, although the definition of the slope as rise-over-run is arguably a formula that a student should know, three mathematicians inserted the slope 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 $f'(y) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$ 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 backed up his reasons for including it by citing its importance to the ideas of the proof:

> FR: *Because the key idea as I see it is this equation right here*, that $f'(y)$=$f(x_2)$-$f(x_1)$/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).

From these revisions, we deduce 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.

4.1.3. Introduced notation. Three participants introduced notation for referencing the Mean Value Theorem in Task B, either by assigning a name to the secant line or a variable to be equal to the slope of the secant line. There are several plausible pedagogical reasons for assigning variables to refer to these quantities, such as stressing the importance of the secant line in the proof or helping the reader reify these quantities as mathematical objects. However, none of the three participants a response of this type.

One participant, TR, apparently considered the notation as support for his exposition: "I want to explain why that slope being strictly positive implies that $f(x_1)$ does not equal $f(x_2)$. And so I gave the line a name." Another reason offered by one participant was efficiency. The cumbersome phrase "the slope between/the line through the points $(x_1,f(x_1))$ and $(x_2,f(x_2))$" appears twice in Task B. By defining this slope with the variable s, PV avoided the need to write this phrase twice. As he introduced his notation, PV commented, "I need to refer to this slope twice."

4.1.4. Added picture. We chose to use Task B since it appeared to be based on graphical reasoning. Only one participant, CY, spontaneously drew a picture during the MVT task:

> CY: … So this one needs … if you're going to say something like that, [draws axes] you're definitely going to need a picture [sketches two points, then a line through those points].

CY then actively referred to this picture throughout the rest of the task, even using the picture as an argument for introducing the hypothesis $x_2>x_1$ in his revision. The reasons why participants did not add a graph in their revisions of Task B are discussed in section 5.

4.1.5. Added word for emphasis. One participant, TR, added the modifier *strictly*, resulting in the sentence, "By our hypothesis, $f'(x_3)$ is strictly positive", giving the reason. "I wanted to emphasize – sometimes people say positive when they mean nonnegative." TR's motivation was to make the meaning of the statement clearer, a theme that we discuss further in Section 4.4.

4.2. Revisions classified as *Altered*

In this category, our participants discussed reasons that related to concreteness, accuracy, and ease of reading. To examine these reasons, we look at the cases of introduced hypothesis, altered reference to the Mean Value Theorem, and altered notation, phrases, and expressions.

4.2.1. Introduced hypothesis. As written, Task B contains the condition that x_1 and x_2 are distinct. Three participants revised Task B to consider a condition equivalent to $x_1<x_2$, even though the proof was valid with or without this added hypothesis. KT and CY attributed their revision to concreteness:

> KT: Without knowing the relationship between x_1 and x_2, you can't really say anything about where x_3 is. Without, again, further awkwardness. But now that you know where it is, you can tell them that x_3 is somewhere in between. And again, it also makes the conclusion that you're drawing easier to state.

> 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.

To KT and CY, being able to specify the relative locations of x_1 and x_2 make it easier to move the mathematics of the proof along. Concreteness, according to KT, supports the transparency of the logical structure: he wanted to be able to state the conclusion in a cleaner manner.

4.2.2. Altered reference to Mean Value Theorem. Upon reading the reference to the Mean Value Theorem in Task B, four mathematicians raised issues of inaccuracy. For example, the reference provoked a strong reaction from BT, who characterized the phrase "the Mean Value Theorem gives" as erroneously suggesting that the Mean Value Theorem is a constructive result rather than an existence statement.

> 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 this 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." It is not clear if the participants did this for improving precision (BT's reason) or so that the argument would conform to the sociomathematical norms of proof.

Another objection that mathematicians raised was the omission of the interval between x_1 and x_2, even though the omission does not detract from the validity of the proof. CY characterized the imprecision in terms of student perception:

> 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.

As TR remarked, the author of Task B was "missing the point" of the Mean Value Theorem.

Neither the phrase "gives" nor the omission of the interval was necessary for the proof in Task B. However, the reference casted the Mean Value Theorem in ways that participants found unfaithful to the proof of the Mean Value Theorem. This lack of acknowledgement to the central ideas in the background material caused the mathematicians to revise the reference.

4.2.3. Altered paragraph or phrase. Besides the reference to the Mean Value Theorem, the following phrases received extensive alterations:

- "We may assume without loss of generality…" (Task A)
- "There is a unique line … coordinate points along the graph of f." (Task B)

Though participants tended to alter these phrases without stating their reasons, two participants described their motivation in terms of pedagogy. IR expressed concern about the phrase "we may assume without loss of generality" and CY about the phrase "coordinate points":

> 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 is less than x_2]".

> CY: I don't know *what* is added by saying 'two coordinate points along the graph of *f*' [crosses it out]. Because students are going to sit there and say, 'What's a coordinate point?' They know that they are points on the graph, and if you say this, you'll just confuse them, I think.

These revisions capture the pedagogical concern that students may be confused about a concept that they know how to manage (let $x_1<x_2$, points along a graph) if it is stated in a mathematically correct but unfamiliar way.

4.2.4. Reformulated algebraic expression or shifted notation. Six mathematicians revised Task A so the slope formula's numerator and denominator were positive. The reasons cited dealt with the 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." FR used the same phrase to describe a revision from the Sine Task, when he performed an algebraic reformulation that aligned his notation more closely to the logical structure of the proof. Thus, although he does not explicitly explain his intent here, we extrapolate that he altered the slope formula so that it fit more closely with the next step of her proof: "By hypotheses, $f'(y)>0$, and $x_2-x_1>0$, so $f(x_2)-f(x_1)>0$."

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

> CY: "I guess I end 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… which is just … it's just cleaner."

Aesthetic reasons drove shifts in notation as well. Two mathematicians altered the notation, finding the inequality $x_1<x_3<x_2$ jarring because the subscripts were out of order.

4.2.5. Replaced word with symbolic expression. CY replaced "strictly positive first derivative" with "$f'(x)>0$ for all x in R" for pedagogical reasons: "students might grab onto '$f'(x)>0$' faster than they would grab onto the phrase 'strictly positive first derivative'." Students may find the symbolic expression easier to use than the verbal phrasing.

4.3 Deleted

Many phrases that were deleted were also phrases that were altered. The stated reasons for deletions, as illustrated below, primarily concerned not including statements in a proof that were not necessary; the arguments against necessity included a characterization of the content as "pretty standard stuff", "confusing", and "it did not add anything." The vast majority of deletions (~30) were made for unstated reasons.

Combining the counts for alterations and deletions, the mathematicians were fairly consistent. Every mathematician reworked or removed the phrase, "If they were equal, we'd gain no information about injectivity" (Task A), and "If for some x_1, x_2 in R, we had $f(x_1)=f(x_2)$, then …"(Task A). Every mathematician reworked or removed the phrase "So either $x_2>x_1$ or $x_1>x_2$" (Task B), with KT explaining that the author had "missed opportunities to simplify the proof". Six mathematicians reworked or removed the phrase "coordinate points". Five mathematicians removed the phrase "The point x_3 is in the domain of f", characterizing it as "unnecessary".

4.4 Summary of Revisions

In this section, we extract common themes among the motivations behind revision. Five themes emerged from our readings of the transcripts and successive coding of the data: making meaning of statements clear, making logical structure explicit, emphasizing the crux of the proof, adding justification, and correcting inaccuracies. The number of instances we coded as representing these motivations is shown in the table below.

Theme	Most common instances of theme
Making meaning of the statements or the proof clear	Eliminating “unnecessary”, “worthless”, or “extraneous” pieces (no further elaborations given) (10), eliminating content that was not used substantively in the proof (3), rewording phrases with confusing terminology (6), simplifying the delivery of the content (3), introducing a hypothesis (3)
Making logical structure explicit	Adding introductory or concluding sentences (7)
Highlighting the crux of the proof	Adding an algebraic formula or manipulation that captured the “key piece” (4), using typesetting to “display” an equation central to the idea of the proof (4)
Adding justification	Providing arguments to complete a missing step (4)
Correcting inaccuracies	Revising reference to Mean Value Theorem (3), correcting mathematical logic (1)

4.4.1. Making meaning of statements or the proof clear

As illustrated in the preceding table, our participants devoted a significant portion of their revision work to clarify the meaning of statements in the proof. Much of the work involved eliminating “unnecessary” and “extraneous” phrases. For example, when TL first encountered the algebraic moves in Task A, he described them as “too much information that it’s distracting.” TR too asserted that redundancy can foster distraction:

> TR: Yeah, so I guess there’s some judgment as to how much can one person hold in their mind at once. Is this still in active memory, or is this written on the hard drive somewhere? And if it’s written on the hard drive, then sure it’s redundant. But bring it back into memory before using it.
> I: What’s wrong with using something that’s in active memory? […]
> TR: It distracts you.

We hypothesize that distraction may be one reason for eliminating other steps characterized as unnecessary or of low difficulty level.

If distractions can also be interpreted as places where a reader could linger unproductively, then the revisions involving terminology are folded into eliminating distractions as well. CY objected to the terminology “coordinate points”, as he worried that a reader may spend time wondering why the author had chosen that phrase (as quoted in Section 4.2.3), and KM criticized the phrase “If they were equal, we’d gain no information about injectivity” because “it doesn’t seem like that’s going to help someone to understand what’s going on. They’re either going to get no information from it, or be confused trying to figure out what it’s trying to say.” Making the meaning of statements clear helps the reader move along.

Other instance of making the meaning of statements clear is predicting misconceptions. For example, TR inserted the modifier “strictly” in front of positive to ward against possible confusion between nonnegative and positive.

Thus far, we have described considerations that anticipate how the audience will read the text. Indeed, by nature, it should seem that a central part of clarifying statements is reflecting for whom the statements may be unclear or clear. However, there was one revision concerning clarity that was fundamentally more about mathematical tangibility than audience perception: clarifying the meaning of statements was simplifying the explanation. IR, CY, and KT introduced the hypothesis $x_2 > x_1$ in Task B to complete the proof without vacillating between two cases. The reasons that were given for this were "concreteness" (IR), avoiding "awkwardness" (KT), and to help draw the picture (CY).

4.4.2. Making logical structure explicit

The way in which participants made the logical structure explicit was through introductory and concluding sentences, as argued in Section 4.1.1. TR demonstrated this through his introductory sentence in Task A and his concluding sentence in Task B.

> TR (Task A): "I like to say at the beginning what I'm going to prove and how I'm going to prove it."
> TR (Task B): And then, in this case, 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 once sentence left that was needed – since x_1 and x_2 are arbitrary, it follows that f is injective, because the values are different.

No sentence in the original versions of Task A or Task B fulfilled the same mathematical and expository role as these beginning and concluding sentences: stating succinctly the heuristic employed for showing injectivity. The sentence closest to this role is the original conclusion of Task A, "If for some x_1, x_2 in R, we had $f(x_1)=f(x_2)$, then by contrapositive, $x_1=x_2$. If not, then the above argument implies $f(x_2)$ is not equal to $f(x_1)$". No participants retained this sentence in their revisions, though FR altered the sentence into his concluding sentence. That so many of the participants added similar introductory or concluding sentences in the face of the absence of such sentences from the original texts, suggests that in the proof genre, the role of such sentences is to clarify the logical structure.

4.4.3. Highlighting the crux of the proof

The most popular way for the participants to highlight what they considered as "key pieces" of the proof was by adding an algebraic manipulation or through typesetting. Six mathematicians manipulated the slope formula in Task A, and all these manipulations were equivalent to multiplying both sides by the denominator x_2-x_1. Given the emphasis on reducing distractions for making meanings clear, it may seem strange that these mathematicians almost uniformly added a relatively trivial mathematical move. However, this step, in contrast to the manipulations that in IR's words "bring us nowhere" in Task A, was seen as critical to the proof's punch line, as described in Section 4.1.2. These moves suggest that the judgment of what is important to include in a proof relies not only the level of difficult of the maneuver, but also how closely tied it is to the main ideas.

Two mathematicians saw the layout of the text as part of revision, and used centering an equation or expression on a separate line to emphasize a concept, or in CY's terminology, to "double dollar sign" an equation.

> 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? [A double dollar sign would center the equation, which might draw a reader's attention to it].
> 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.

It is interesting that CY and TR saw typesetting as part of their revision process. Although PV and IR did not explicitly mention typesetting, they too "double dollar signed" equations in their revisions (i.e., indented the formula). Konoir (1993) describes typesetting as crucial to the proof genre in emphasizing key ideas and the logical structure of the proof. The data here illustrate how some mathematicians would use typesetting for these purposes.

4.4.4. Adding justification and correcting inaccuracies

To some of the mathematicians in the study, what counted needing correction or justification depended on the author and the intended audience. For example, as IR asserted of Task A:

> IR (Task A): And, so, but this "So" here is what they have to justify. And it is not justified out of anything. I mean, so. It is justified out of this equation and this equation and this equation for **me**, but not for students. So this is all … bad. (Italics are IR's emphasis).

IR noted that the proofs in Task B provided might pass muster as something that a student wrote – but would not constitute good exposition in a textbook or published work.

> IR (Task B): I mean, if this were that some student, obviously a smart student, wrote that, so presents that solution in class. Then he would get an A. But if it is intended to explain why it is true, then I think it's a little bit not so good.

Similarly, in describing his revision of the reference to the Mean Value Theorem to specify the interval $[x_1, x_2]$, CY said, "If a student wrote this, I wouldn't mark it wrong." But as mentioned in Section 4.2, CY also believed that citing the interval served a pedagogical purpose.

Thus there can be a tension between sheer validity – which is considered when grading student papers – and good explanation and pedagogy, perhaps because explanation and pedagogy involve understanding the mathematical ideas in the proof and predicting how the reader will perceive the text. In the data that we collected, the revision moves around justification and clarification primarily concern the audience's interpretation, whereas the rationales for moves around the crux and logical structure of the proof were more driven by the mathematicians' interpretation of the mathematics.

5. Comments from participants about the process of revision

5. 1. The importance of audience when revising a proof

When discussing how they revised the proof, all ten participants remarked about the importance of the audience of the proof. Indeed, eight participants mentioned the audience without any prompting or direct questions from the interviewer to do so. Further, three participants felt their task was undefined because they did not know who would be reading the proof. For instance, 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.

The most common reason for considering the audience was to determine what was established knowledge and what needed justification. Eight made comments of this sort. For instance, PV remarked that in presenting a proof, "you justify things that you think your reader won't know and you otherwise just state the ideas. Different levels of different things should be justified [depending on who the audience is]". Similarly, CY commented, "The big thing for me would be deciding … what can be asserted and what needs backup. And that would be the real issue, in writing this proof". It is interesting to note that the mathematicians in Weber's (2008) study on proof validation also mentioned the importance of audience in determining if a proof was correct and that, in some cases, they considered asking whether a proof was valid without stating its intended audience as an ill-defined task.

Five participants considered what type of statements and proof techniques would be most comprehensible to the students they were teaching. For instance, IR remarked that he prefers to prove by contradiction but some students find the technique confusing, so he aims to present direct proofs to students. Similarly, KM expressed the desire to use proving techniques that students were comfortable with, rather than present a more efficient proof using an idea they might not have seen before. TR expresses a similar viewpoint:

> TR: If my audience is sophomore/junior mathematics majors, I'm thinking to myself, "what do they know, what have they seen over the years, what have they seen most recently, and what's fresh in their minds?' Let me relate things as much as possible to things that are fresh in their minds.

Five participants discussed the difficulties they had in presenting proofs to a large student population with different backgrounds, abilities, and ways of thinking. KM described this in the context of providing details within a proof, saying "people's brains absorb details at different rates, so if you tried to give details you'd be boring some people and maybe confusing others". KZ remarked that students prefer to think of mathematics in different ways and the same proof might not be best for all students.

> KZ: I think the picture would be most helpful … because that would be closer to how I think about it. But that depends on whether the student likes to think in terms of pictures or in terms of formulas or whatever way.

5. 2. The medium of the presentation

It was surprising to us that five participants discussed how their proof revisions and presentation depended upon whether they were preparing the proof for a textbook, a lecture, or a one-on-one interaction with the student. Three participants argued that textbook presentations needed to be complete and precise. An illustrative comment of this type is presented below:

> TR: I think a lecture is so much more flexible because you can talk and watch the student's faces and see if they're getting it or they're not getting it. Whereas a textbook, you put it out there and then you have absolutely no control over how anyone responds or understands it, so yeah, you want to be super careful what you're writing in a textbook.

IR made a similar observation: "In a textbook, no one is speaking. So everything you have is the written thing". The participants felt that they had more flexibility in a lecture and their job was to present a big picture, or the intuitive idea, behind the proof, rather than focus on the logical details. This is illustrated with KM in the transcript below:

> KM: [In a lecture] the point of the instructor is that the instructor can give, maybe a big picture kind of view of things that is harder to get from a textbook, than details.
> I: Gotcha, OK. So in other words, there might be a difference between the sort of arguments that you might present as an instructor versus as an author of a textbook?
> KM: I think so. I think, well, if you're writing a textbook, one of your goals is to have someone be able to read your book and understand it in complete detail, what you're talking about. Whereas for many things, this is impossible in a classroom setting because you don't have time to go into all of the details… What you can do much more effectively at the board is, say, draw pictures in real time and point at things.

5. 3. Central ideas

After completing the Sine Task, participants were asked if they thought their proof brought out the key idea used to create it. The participant who drew a picture (TR) claimed he did include the key idea in his proof. Three participants realized their proofs did not contain a key idea and found this to be a shortcoming of their proofs. For instance, when asked if he tried to bring the key idea out in his proof, KT replied, "No. And you're right. For pedagogical purposes, that is probably the more important thing to do". Similarly, TL commented:

> TL: But I did not do that [make the key idea, the application of the Intermediate Value Theorem, clear]. I just wrote out the string of ideas that the logical construction of the arguments and sort of treated them all equally. It might not be clear… where the meatiest part was. But it might not be clear to someone who was just reading it that that was really the important part.

Three participants were not sure if it was necessary to emphasize the key idea in their proof. IR claimed this depended upon "how good the textbook should be". KM didn't think "much emphasis [on the key idea] was required in this case, because the proof just consists of implementing the idea" and CY made a similar comment. In summary, although many participants valued presenting the main ideas of their proofs in their teaching, only one of the participants took this into account when creating and revising their proof.

5. 4. Proofs and pictures

The second proof in the Mean Value Task was based on the graph of an increasing function, yet when revising the proof, only one participant suggested adding a graph to improve the comprehensibility of the proof. One other participant, TR, when asked if there was a difference between the two proofs of the Mean Value Task, commented "they [the authors of the second proof] drew a picture somewhere and I would have encouraged them to draw the picture in their solution". The remaining participants did not mention a picture until the interviewer asked whether adding a picture would remove the proof.

When asked by the interviewer, six of these eight participants agreed that adding a picture would help and thought their revision would be improved if a graph had been added. PV claimed he would not draw a graph to avoid redundancy.

> PV: I guess I did not draw the graphic because there is a graphic in the explanation of the mean value theorem. And that's where it should be. I mean if the context is for a textbook, I would be presenting this proof. Here this picture points on the same page as this. I would not draw the picture again.

KZ claimed he would not draw a graph because he feared students would find the proof unnecessary once they inspected the graph. In summary, most participants agreed that a graph would make the second proof in the Mean Value Task clearer and that adding the graph would be a useful revision for the proof, but most participants did not make this revision without prompting from the interviewer.

5. 5. Tensions in proof revision

All participants expressed a desire to be concise in their proofs, but four participants noted that sometimes this goal was in conflict with the goal of presenting an understandable proof. This is illustrated in the excerpts below.

> KT: That's always important in my opinion, is to have a proof that only uses what's necessary. And then, sometimes, you have to trade that off against increasing complexity… Often, the shortest proofs are not the clearest. Sometimes they're too clever. Or they take advantage of special features o the problem, even when there might be a more robust technique that takes slightly longer but is more widely applicable.

> TR: You want to be precise, clear, but also, I feel like you want to present things … the way that a person can absorb it. And so something can be concise and clear and written out Russell/Whitehead style. All symbols—there exists, for all, double brackets, all this crazy stuff… even though a proof written in that language and notation might be a lot shorter, it takes me a lot longer to parse it.

In these excerpts, KT discusses how the desire to have the shortest proof possible may be in conflict with having a clear proof or a proof whose techniques might be applicable in other settings. TR also notes how the desire for a concise proof can be at variance with one that is understandable.

6. Discussion

6. 1. Limitations

Section 4 examined the behavior of eight mathematicians revising two proofs for pedagogical purposes. Consequently we cannot claim to characterize all of mathematicians' revision goals and all the processes that mathematicians use to achieve those goals. It is likely that conducting a similar study with more mathematicians and different tasks would have yielded additional types of revisions. Similarly, other types of revisions may have been elicited if the proof were intended for a different age group of students, or for professional rather than pedagogical purposes.

6. 2. Summary of findings

Our data suggest that the participants in this study believed that a good proof for pedagogical purposes had the following three features: (a) the proof was concise, (b) the meanings of the statements of the proof, and the proof as a whole, were clear and comprehensible to its intended audience, and (c) the main idea of the proof would be clear to the reader. In their revision process, to achieve the first aim, the participants would remove extraneous and unnecessary statements. They would also logically restructure the proof to make the argument in the proof more direct, such as by reducing the number of cases that needed to be considered.

The participants discussed a number of strategies that they could use to improve the clarity of a proof to its audience. In revising the proof, they would add justifications to statements that might not be obvious to its audience, revise statements that were technically correct to improve their clarity, and add additional words or phrases to statements to avoid misinterpretations. These revisions refer to how the participants would amend the details of the arguments that they presented. However, the interview data suggests that considerations of the audience might influence the *type* of argument they presented. Some participants expressed a desire to use forms of argumentation that students were familiar and comfortable with, and to base their arguments on facts that students had recently seen.

While most participants expressed the desire to bring out the main ideas of their proofs in their presentations, they did not always attempt to do so in constructing or revising a proof. For instance, for the Sine Task, only one participant explicitly brought out the key idea in his final proof. For MVT Task, only one participant added a proof when revising Task B, yet later in the interview, eight participants claimed that Task B would have been improved with the addition of a picture. The participants did attempt to bring out the main idea of MVT Task A by restating the Mean Value Theorem so the reader could see how it was applied, emphasizing its use within the argument (including using judicious formatting), and restructuring the argument so it was more in the spirit of the application of the Mean Value Theorem. The participants also added beginning and concluding sentences for the MVT task to make explicit what the proof was showing and why it would prove the claim.

It is interesting to note that participants claimed that differential emphasis on the three goals above depended upon the medium that the proof was being presented in. If the argument was in a textbook, the participants would emphasize clarity and correctness. If the argument was presented in a lecture, the participants claimed to emphasize the larger ideas of the proof while avoiding minor details. To justify this distinction, the participants cited both the functions of the medium (e.g., textbooks are meant to serve as references) and the affordances of the medium (e.g., lectures allow for a more dynamic presentation, yet as opposed to textbooks, students have a limited amount of time to comprehend the argument presented in real time).

The participants noted several difficulties with creating a good proof for pedagogical purposes. First, classrooms often have students that vary considerably in mathematical background, ability, and learning styles. An assertion that is obvious to one student may be baffling to another. Spending time justifying the first assertion may bore or distract the first student, yet glossing over it will lose the second student. Similarly, it can be difficult to base a proof on ideas that students are familiar with if students are not all comfortable with the same mathematics. Second, the goal of presenting a concise proof can be at variance with a proof that is comprehensible or brings out the main idea. We suggest there may be a third difficulty in presenting a good proof for the classroom. The participants may overestimate what students may infer when reading a proof. For instance, three participants did not believe it was necessary to emphasize the key idea of their proof in the Sine Task because they thought it would be transparently obvious to the student, in part because the proof was simply an application of this idea. In our experience, students do not always recognize this. We also hypothesize that one reason that participants did not draw graphs when revising MVT Task B is that it was obvious to them what type of graph the proof was referring to and they had little difficulty coordinating the details of the proof with a mental image of the graph. We suspect that students would not be able to do this so easily. Several mathematicians suggested typesetting the Mean Value Theorem to emphasize its importance in the MVT Tasks, a rhetorical device that Konoir (1993) argues is common in the genre of proof. However, Weinberg (2009) cautions that students might not be aware of such rhetorical cues when reading proof. Having a better understanding of what students attend to when reading proofs, and how they attend to them, could provide useful insight into how proofs can be presented better to achieve their pedagogical aims. Such research is limited (Mejia-Ramos & Inglis, 2009) and we suggest more research in this area.

References

Alibert, D. and Thomas, M. (1991). Research on Mathematical Proof. In D. O. Tall (Ed.), *Advanced Mathematical Thinking*, (p. 215-230). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Boaler, J., Ball, D. L., & Even, R. (2003). Preparing researchers for disciplined inquiry: Learning from, in, and for practice. In A. Bishop & J. Kilpatrick (Eds.), *International Handbook of Mathematics Education* (491-521). Dordrecht: Kluwer.

Bressoud, D.M. (1999). *Proofs and confirmations: The story of the alternating sign matrix conjecture*. Cambridge University Press: Cambridge.

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

Downs, M. L. N. & Mamona-Downs, J. (2005). The proof language as a regulator of rigor in proof, and its effect on student behavior. In *Proceedings of the 4th Conference on European Research in Mathematics Education*. Sant Feliu de Guixols, Spain.

Hanna, G. (1990). Some Pedagogical Aspects of Proof. *Interchange*, 21(1), 6-13.

Hanna, G. (1995). Challenges to the importance of proof. *For the Learning of Mathematics*, 15(3), 42-50.

Hanna, G. (2000). Proof, explanation, and exploration: An overview. *Educational Studies in Mathematics*, 44, 5-23.

Hanna, G. & Barbeau, E. (2008). Proofs as bearers of mathematical knowledge. *ZDM, 40*, 345-353.

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

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

Herbst, P. & Brach, C. (2006). Proving and doing proofs in high school geometry: What is it that is going on for the students. *Cognition and Instruction*, 24, 73-122.

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

Konoir, J. (1993). Research into the construction of mathematical texts. *Educational Studies in Mathematics*, 24, 251-256.

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

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

Mancosu, P. (2000). Mathematical explanation. In E. Grosholz & H. Breger (eds.) *The growth of mathematical knowledge*. 103-119.

Mejia-Ramos, J.P. (2008). *The construction and evaluation of arguments in undergraduate mathematics: A theoretical and a longitudinal multiple-case study*. Unpublished Ph.D. dissertation. University of Warwick, U.K.

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.

Nardi, E., & Iannone, P. (2005). To appear and to be: Acquiring the 'genre speech' of university mathematics. In *Proceedings of the 4th Conference on European Research in Mathematics Education*. Sant Feliu de Guixols, Spain. Available for download at: http://cerme4.crm.es/Papers%20definitius/14/NardiIannone.pdf.

National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.

Otte, M. (1994). Mathematical knowledge and the problem of proof. *Educational Studies in Mathematics*, 26, 299-321.

Raman, M. (2003). Key ideas: What are they and how can they help us understand how people view proof? *Educational Studies in Mathematics*, 52, 319-325.

Rav, Y. (1999). Why do we prove theorems? *Philosophia Mathematica, 7*, 5-41.

Resnick, M. & Kushner, D. (1987). Explanation, independence, and realism in mathematics. *British Journal for the Philosophy of Science*, 49, 603-624.

Schoenfeld, A. (1994). What do we know about mathematics curricula? *Journal of Mathematical Behavior*, 13, 55-80.
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
Selden, J. & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, 29(2), 123-151.
Steiner, M. (1978). Mathematical explanation. *Philosophical Studies*, 34, 135-151.
Strauss, A. & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. London: SAGE.
Thurston, W.P. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society*, 30, 161-177.
Volminik, J.D. (1990). The nature and role of proof in mathematics education. *Pythagoras,* 23, 7-10.
Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39*, 431-459.
Weber, K. (2010). Proofs that develop insight. *For the Learning of Mathematics*, 30(1), 32-36.
Weber, K. & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics*, 56, 209-234.
Weber, K. & Alcock, L. (2005). Using warranted implications to understand and validate proofs. *For the Learning of Mathematics, 25*(1), 34-38.
Weber, K. & Alcock, L. (2009). Proof in advanced mathematics classes: Semantic and syntactic reasoning in the representation system of proof. In D. Stylianou, M. Blanton, & E. Knuth (eds.) *Teaching and learning proof across the grades: A K-16 perspective*. Routledge: New York.
Weinberg, A. (2009). How students use their textbooks: Reading models and model readers. In *Proceedings of the 12th Conference for Research in Undergraduate Mathematics Education.* Available for download: http://rume.org/crume2009/proceedings.html