Why and how mathematicians read proofs: An exploratory study

Keith Weber, Juan Pablo Mejia-Ramos

Graduate School of Education

Rutgers – The State University of New Jersey

Abstract

In this paper we report a study in which nine research mathematicians were interviewed with regards to the goals guiding their reading of published proofs and the type of reasoning they use to reach these goals. Using the data from this study, as well as data from a separate study (Weber, 2008) and the philosophical literature on mathematical proof, we identify three general strategies that mathematicians employ when reading proofs: appealing to the authority of other mathematicians who read the proof, line-by-line reading, and modular reading. We argue that non-deductive reasoning plays an important role in each of these three strategies.

1. Introduction

In the last thirty years, there has been a great deal of research in mathematics education in the area of justification and proof. Although this research has focused on a wide range of topics, including the construction of proof, the epistemological nature of proof, and the development of proof in mathematical classrooms, only recently has research in mathematics education investigated the *reading* of mathematical proof (Selden & Selden, 2003; Alcock & Weber, 2005; Mejia-Ramos, 2008).

Most researchers who have examined the reading of proofs have sought to understand students' conceptions of proofs by asking them to read different types of arguments and to *evaluate* these arguments against a given set of criteria (e.g. personal preference, persuasiveness, mathematical validity). This literature has produced three main findings. First, students at all levels often find empirical arguments to be personally convincing and representative of the ways that they would justify mathematical statements (e.g., Martin & Harel, 1989; Healy & Hoyles, 2000). Second, mathematics majors often have difficulty distinguishing between valid mathematical proofs and flawed mathematical arguments (e.g., Selden & Selden, 2003; Alcock & Weber, 2005). Third, students may view a deductive proof of an assertion merely as evidence in favor of the assertion rather than necessitating its truth (e.g., Fischbein, 1982). Together, these findings suggest that students become convinced of mathematical assertions for different reasons than mathematicians do (e.g., Harel & Sowder, 1998).

A goal of many research programs is to lead students to think and behave more like mathematicians with respect to proof. Harel and Sowder (2007) explicitly contend that the purpose of mathematics instruction should be "to help students gradually develop an understanding of proof that is consistent with that shared and practiced in contemporary mathematics”. To accomplish this goal, some researchers have conducted teaching experiments designed to have students develop the same standards of conviction as mathematicians (e.g. Harel, 2001; Stylianides & Styilianides, 2008) while others have created learning environments designed to have students engage in proof-related activity that is similar to mathematicians’ practice (e.g., Lampert, 1990; Maher, Muter, & Kiczek, 2007).

If goals of mathematics instruction include having students (1) engage in the same types of proof-related activities that mathematicians do, (2) behave like mathematicians in these activities, and (3) adopt mathematicians’ beliefs regarding proof, then it is necessary to have an accurate understanding of the types of activities that mathematicians engage in, how mathematicians perform them, and what their beliefs about proof actually are. Recently, the RAND Mathematics Study Panel (2003), a multidisciplinary panel commissioned by the United States Institute of Educational Sciences to improve the quality of mathematics education and educational research, concluded that more research on mathematicians’ practice pertaining to justification and proof is needed to form a sufficient basis to design instruction. Consistent with this report, several recent empirical studies and philosophical theses have investigated this topic and yielded surprising findings about mathematicians’ practice with justification and proof (e.g., Rav, 1999; Inglis, Mejia-Ramos, & Simpson, 2007; Weber, 2008; Inglis & Mejia-Ramos, 2009) and, in some cases, these findings have had important implications for the teaching of mathematics (e.g., Hanna & Barbreau, 2008). With regards to reading proofs, Konoir (1993) contended that “getting to know the complex processes and mechanisms of reading text has essential significance for the didactics of mathematics" (p. 251), noting that few such studies have been conducted and arguing that studies of mathematicians' practice when reading proof need to be conducted.

The goal of this paper is to contribute to the mathematics education community's understanding of mathematicians' practice with regard to the reading of mathematical proof by addressing the following questions:

- For what purposes do mathematicians read the proofs of their colleagues?
- How do they read proofs to achieve these aims?
- What role does non-deductive reasoning play in this proof reading?

These questions are addressed by interviews with mathematicians about their practice in section 3. These results are then coordinated with the results of a previous study (Weber, 2008) and the philosophical literature to present a model of the ways in which mathematicians read proofs. This model is presented in section 4. Finally, we use these findings to discuss what the appropriate goals of instruction with respect to proof should be in mathematics classrooms.

2. Theoretical perspective and clarification of terminology

2. 1. The nature of proof

Balacheff (2002) argued that there is not a consensus within the mathematics education community as to what constitutes a proof, or what should be the object of research with respect to proof. He illustrated how some researchers have emphasized the types of reasoning used in mathematical arguments, others have investigated students' perceptions of proof, and others still have examined the functions that proof plays in the mathematics community.

Boero (1999) contributed to the debate on what the object of research on proof should be by noting the need to distinguish between the *process* of conjecturing and proving and the *products* of these activities (i.e., written conjectures or proofs). In this paper, when we speak of proofs, we refer to the socially sanctioned written product that results from mathematicians' attempts to justify why a conjecture is true[1]. The goal of this paper is to focus on the different activities of mathematicians as *readers* of published proofs.

2.2. Activities related to proving

Mejia-Ramos (2008) argued that there are three main argumentative activities related to proving: constructing a novel argument, presenting an available argument, and reading a given argument. Mejia-Ramos emphasized that the way individuals engage in these activities depends on their purposes for engaging in these activities. For instance, individuals may read a proof for conviction differently than they would read it for insight.

Regarding the reading of mathematical arguments, Mejia-Ramos (2008) distinguished between *comprehending* and *evaluating* an argument. A typical example of *argument comprehension* may occur when students read a proof in a textbook or an argument in their teachers' lecture notes. In principle, the focus of this reading activity is not solely on checking if the given proof is correct (or evaluating it in any other manner), but on understanding the content of the proof and learning from it. On the other hand, a

[1] This refers to phase V of Boero's (1999) six phases of conjecturing and proving.

typical example of *argument evaluation* occurs when teachers judge the validity of their students' arguments, or when they assess how clear and insightful a mathematical argument is before they present it in class. Similarly, argument evaluation is common among research mathematicians who review the validity of arguments for publication. In this paper, we focus on mathematicians' stated goals for reading a published proof, and what they claim to do in order to reach these goals.

2. 3. Types of supplementary arguments used in a proof reading activity

When reading a proof, the reader may use supplementary arguments to reach a particular goal in the proof reading activity. In this paper we expand the typology of warrant-types described by Inglis, Mejia-Ramos, and Simpson (2007) by distinguishing four types of arguments that an individual may use when reading a proof.

When an assertion claims every element of a set satisfies a given property, one may check that this assertion is true (or attempt to understand the assertion) by verifying that (or, in the case of generic proofs, illustrating how) a proper subset of the given set satisfies this property. We refer to this type of argument as *empirical*. One may attempt to understand or increase one's confidence in the truth of a claim by seeing if this claim is a property or consequence of one's mental models associated to the relevant mathematical concepts (i.e. diagrams or images that one associates with those concepts). We refer to this type of arguments as *structural-intuitive*. One may increase one's confidence that a claim is correct because an authoritative source endorsed that claim. We refer to this type of argument as *authoritative*. Finally, one may produce or observe a deductive argument that derives the claim from assumptions that are agreed upon as established using socially acceptable mathematical techniques. We refer to this type of justification as *deductive*. In this paper, we are interested in the types of supplementary arguments used by mathematicians in different proof reading activities. Boero (1999) notes that empirical and structural-intuitive arguments are often useful in some stages of conjecturing and proving, but generally do not appear in the product of these activities (i.e., empirical and structural-intuitive evidence are not present in a conjecture or a proof of a theorem). In this paper, we examine how these types of evidence are used by mathematicians in the reading of proofs.

2. 4. Normative views about proof reading

It is widely accepted among mathematicians and mathematics educators that empirical, structural-intuitive, and authoritative justifications have significant limitations in mathematical argumentation. Empirical arguments may lead to false conclusions because the assertion might be true for the examples that one happened to consider, but false for an example that was not considered. Structural-intuitive justifications may be deceiving because one's mental models of the mathematical domain being studied might be inaccurate. Authoritative evidence is similarly limited because the authoritative source might be mistaken. Hence, while any of these sources of evidence can help one understand an assertion and increase one's level of belief that it is true, these modes of argumentation should not be regarded as conclusive.

On the other hand, some mathematics educators, as well as some mathematicians, bestow deductive arguments with a more prominent role in obtaining conviction that a mathematical assertion is true. Some mathematics educators believe deductive arguments can guarantee the truth of a mathematical assertion in a given proof. Duval (2007), for instance, contends that, as opposed to other modes of argumentation, the purpose of valid deductive arguments is to change the epistemic value of the statement being proven to "necessary". In other words, a proven statement is guaranteed to be true *on behalf of* a valid deductive argument. Similarly, Harel and Sowder (1998) argue "that an observation remains a conjecture until the person reaches *absolute certainty* in its truth" (p. 241, italics are our emphasis). They further claim that mathematicians hold deductive proof schemes, implying that mathematicians use deductive arguments to obtain this absolute certainty. However, some disagree with such claims. For instance, de Villiers (1990) claimed that mathematicians are usually convinced that a theorem is true *before* attempting to prove it. Also, Otte (1994) argued that gaining absolute certainty in a theorem via a proof is theoretically impossible, since one would have to produce (or read) a proof, then validate the proof, then validate the validation, and so on.

3. Data for this paper

In section 4, we will present a model of three general strategies that mathematicians employ when reading the mathematical proofs of others. In addition to the writings of mathematicians and philosophers of mathematics, we will use two primary sources of data to support our argument. The first comes from a recent paper (Weber, 2008) in which the first author of this paper (a) analyzed the way in which eight research active mathematicians determined if mathematical arguments were correct and (b) interviewed these eight mathematicians about their professional practice in evaluating the proofs of others. A central result from this paper was that mathematicians occasionally marshaled non-deductive evidence to bridge what they perceived to be inferential gaps in the proof. In particular, these mathematicians often checked if particular statements in a proof were true by verifying that they held for carefully chosen examples. The second source of data, presented in this paper, comes from interviews with nine other mathematicians on their professional practice of reading the published proofs of others. As the data from this study comes from a total of 17 mathematicians, we cannot claim to offer an exhaustive list of all the reasons why mathematicians read proofs or all the strategies they use to do so. However, by examining the commonalities of these mathematicians' comments, we can hypothesize motivations and strategies for reading proofs that are used by many in the mathematical community.

3. 1. Research methods

Nine professional mathematicians participated in this study and agreed to meet individually with the first author for a semi-structured interview. All participants were tenured mathematics professors at a large research university in the northeast United States. All were highly successful researchers in their fields of study, which included analysis, algebra, and differential equations. These mathematicians differed from the participants in Weber's (2008) study in that they worked at a research university, were more experienced, and achieved national prominence in their areas of research.

Each interview was semi-structured and was one to two hours long. Each interview was audiotaped and then transcribed. The goal of the interview was to investigate the reasons why and the ways in which the mathematicians read the published proofs of their colleagues. The analysis in this paper will focus primarily on the participants' responses to the following questions:

- In your own mathematical work, I assume you sometimes read the published proofs of others. What do you hope to gain by reading these proofs?
- What do you think it means to understand a proof?
- What are some of the things that you do to understand proofs better?
- Does considering specific examples ever increase your confidence that a proof is correct?

The first question addresses participants' goals for reading proofs, the second and third are related to their behavior during proof comprehension, and the last was meant to further investigate Weber's (2008) finding that mathematicians use empirical evidence to evaluate the validity of purported proofs.

3. 2. Results

All mathematicians confirmed that reading the published proofs of others was a significant part of their mathematical practice. In fact, M7's initial response to whether he read other mathematicians' proofs was: "What do you think I was doing when you came into my office?"

Do mathematicians check proofs for correctness?

Six participants indicated that when reading proofs in mathematical journals, they would sometimes do so to determine if the proofs were correct (although some of these participants may have been referring to refereeing a paper, rather than reading a published paper). One representative response was included below:

I: What do you hope to gain when you read these proofs?

M4: Okay. Two things. One is I would like to find out whether their asserted result is true, or whether I should believe that it's true. And that might help me, if it's something I'd like to use, then knowing it's true frees me to use it. If I don't follow their proof then I would be psychologically disabled from using it. Even if somebody I respect immensely believes that it's true. More importantly, I want to understand the proof technique in case I can use bits and pieces of that proof technique to prove something that they haven't yet, that the original author hasn't yet proved.

Two of the participants specifically claimed that they did *not* read proofs to check their correctness. M8 emphasizes this point in the excerpt below.

M8: Now notice what I did *not* say. I do not try and determine if a proof is correct. If it's in a journal, I assume it is. I'm much more interested in the ideas of the proof. (Emphasis was M8's).

Similarly, when asked what he hoped to gain out of a proof, M6 did not specifically mention determining if a proof was right, prompting the interviewer to ask why this was not said.

I: One of the things you didn't say was you would read it to be sure the theorem was true. Is that because it was too obvious to say or is that not why you would read the proof?

M6: Well, I mean, it depends. If it's something in the published literature then… I've certainly encountered mistakes in the published literature, but it's not high in my mind. So in other words I am open to the possibility that there's a mistake in the proof, but I… it's not… [pause]

I: But you act on the assumption that it's probably correct?

M6: Yeah, that's right. That's right.

M6 and M8 both act on the assumption that proofs published in journals are probably correct and do not feel the need to personally validate them. Among the nine mathematicians who were interviewed, these two mathematicians were the only ones to make a comment of this type, although other participants may have held similar viewpoints. Several mathematicians interviewed by Weber (2008) also indicated that they would not check a proof for correctness because they trusted the proof if it appeared in a reputable journal.

Reading a proof for insight.

All nine mathematicians claimed they read proofs for ideas that might be useful in their own research. One instance is provided above, in which M4 says that this is more important than checking for correctness. Other excerpts are provided below:

M6: Well, I would say most often is to get some ideas that might be useful to me for proving things myself. [When asked to elaborate] Ok, actually, I'll say two things. One is, it's just to satisfy my innate curiosity as to *what are they doing to get this conclusion*. So that's one aspect. But then the other aspect is… I mean, usually I'm reading something because it seems to have some connection to some problems that I was interested in. *I'm hoping that if that the tools they're using or ideas they're using might connect up to some of these problems that I have thought about.* (Our emphasis).

M1: Theorems as a way of organizing major results are extremely useful, but they involve a decision on the part of the person who's writing the statement of the theorem of what thing to take as a hypothesis, what is the conclusion, and what route to follow to get from hypothesis to conclusion. And often along that route there are techniques that could have been stated as separate theorems but are not, and then you read the proof carefully and *you discover these are things that you can use*. That's certainly from a pragmatic point of view, *that's an important part of reading proofs, that you steal good ideas out of good proofs.* (Our emphasis).

The goal of these mathematicians' reading seems to go beyond comprehending the proof. They seek to find ideas and techniques in their proofs that will help them address their own research questions. In the following excerpt, I asked M5 to go into more depth about what he meant with respect to looking at the ideas of a proof and seeing if they could apply elsewhere.

M5: As a researcher, I want to understand the idea of the proof and to see if that idea could be applied elsewhere.

I: The second point that you made, the one about ideas, is something that I've been hearing from your colleagues too. Can you elaborate on that?

M5: Sure. Sometimes when a mathematician answers a hard question, he has a new way of looking at the problem or a new way of thinking about it. As a researcher, when you see this, sometimes you can use this idea to solve problems that you are working on. Let me give you an example. We were having trouble showing bounds for approximation techniques on this space with an unusual norm. Someone realized that you could use this particular partial differential equation to find these bounds. This new idea made a lot of the other problems easier. The idea wasn't easy. It wasn't obvious at all that this partial differential equation was relevant. That was a great insight. But once we had the idea, it allowed us to approach questions that were inaccessible before.

I: So after this theorem came out, a lot of other theorems were proved using this idea?

M5: Oh yeah. But it doesn't always have to be big things, although this one was. Sometimes when I read a proof, I get an idea that helps me get around a little thing that I was stuck with.

Not only did the participants consistently emphasize that ideas were the primary reason that they read proofs, but two participants went so far as to question whether a proof of a new theorem is of value if it does not contain new ideas.

> M9: [As editor of a journal,] I occasionally get papers where the author took an idea from a new proof that just came out, took the idea from the proof, and applied it in a straightforward fashion to prove some new theorem. I'm reluctant to publish these types of papers.

> M1: When I was on the editorial board for one of the journals, one of the instructions we had was "it's not allowed to just publish a paper where you've taken somebody else's proof and simply made a different statement of what we should get out of the proof". That if there is a really good theorem and you come up with an original, alternate proof of that theorem, that could be publishable. But just to take the same proof and say well, we can state the conclusion differently. That's not considered professionally acceptable as a result … Mathematics has this very high, perhaps unrealistically high standard for what is admissible as a claim of research.

These excerpts (which speak to reviewing a proof rather than reading a published proof) suggest that the result of a mathematical piece of research does not depend solely, perhaps not even primarily, on the significance of the theorem being proved and the validity of its proof. The originality and utility of the ideas in the proof is of crucial importance. This notion is endorsed by Rav (1999), who emphasizes that proofs, not theorems, are the bearers of mathematical knowledge (p. 20). Like M1, Rav suggests that what a mathematician claims a proof establishes is somewhat arbitrary; it is the proof method that is of primary importance: "Think of proof as a network of roads in a public transportation system, and regard statements of theorems as bus stops; the site of the stops is just a matter of convenience" (p. 20-21).

Understanding a proof.

When asked what it means to understand a proof, all nine participants indicated that understanding did not solely consist of knowing how each step followed logically from previous steps. In fact, several participants distinguished between understanding a proof logically and understanding the central ideas of a proof.

> M8: There are different levels of understanding. One level of understanding is knowing the logic, knowing why the proof is true. A different level of understanding is seeing the big idea in the proof. When I read a proof, I sometimes think, how is the author really trying to go about this, what specific things is he trying to do, and how does he go about doing them. Understanding that, I think, is different than understanding how each sort of logical piece fits together.

In fact, M5 explained that although a full understanding of a proof involved verifying that each particular instance within a proof was valid, this was a process that he often did not engage in.

> M5: [To understand a proof] means to understand how each step followed from the previous one. I don't always do this, even when I referee. I simply don't always have time to look over all the details of every proof in every paper that I read. When I read the theorem, I think, is this theorem likely to be true and what does the author need to show to prove it's true. And then I find the big idea of the proof and see if it will work. If the big idea works, if the key idea makes sense, probably the rest of the details of the proof are going to work too.

Three of the participants mentioned organizing the proof in terms of different sections or modules was a useful tool for building an understanding of proof, two excerpts of which are provided below.

> M6: Another tool [for understanding proofs] is properly encapsulating the pieces of the proof … I have one particular example that I spent a lot of time on this, where there was this very technical lemma in a paper and the lemma had a bunch of hypotheses. And, you know, what I was trying to do was just strip away the superfluous detail… really this lemma held in a much more general context where there were many fewer hypotheses, and then you get something which it

just reduced the technical mess of the proof a lot because, you know, he's carrying along all these hypotheses which were really unnecessary.

M9: [Understanding a proof involves] understanding how the proof is structured. A good proof often has a number of interesting lemmas and corollaries and sub-theorems and the like. Longer proofs can get pretty complicated. One of the things I try to do when I read a proof is to see how all these things, these lemmas and such, fit together.

The use of examples in understanding a proof.

When asked what they did to understand a proof better, six of the participants claimed that they would consider how the proof related to specific examples.

M4: Commonly, if I'm really befuddled and if it's appropriate, I will keep a two-column set of notes: one in which I'm trying to understand the proof, and the other in which I'm trying to apply that technique to proving a special case of the general theorem.

M1: I'm doing a reading course with a student on wallpaper groups and there is a very elegant, short proof on the classification of wallpaper groups written by an English mathematician. So in reading this… so this is one where he's deliberately not drawing pictures because he wants the reader to draw pictures. And so I'm constantly writing in the margin, and trying to get the student to adopt the same pattern. Each assertion in the proof basically requires writing in the margin, or doing an extra verification, especially when an assertion is made that is not so obviously a direct consequence of a previous assertion.

I: So you're writing a lot of sub-proofs?

M1: I write lots of sub-proofs. And also I try to check examples, especially if it's a field I'm not that familiar with, I try to check it against examples that I might know.

When asked if they ever used examples to increase their confidence that a proof is correct, all nine participants emphatically answered yes. Indeed, to some participants, this question was almost meaningless since they claimed that they never read a proof without considering examples. Many of the participants discussed using examples so they could view the proof as a generic proof, as M4 does in the excerpt above. When asked if he used examples to increase his confidence in the correctness of a proof, M5 responded:

M5: Always. Always. Like I said, I never just read a proof at an abstract level. *I always use examples to make sure the theorem makes sense and the proof works.* I'm sure there are some mathematicians that can work at an abstract level and never consider examples, but I'm not one of them. *When I'm looking through a proof, I can go off-track or believe some things that are not true.* I always use examples to see that makes sense. (Our emphasis)

The italicized portions of the excerpt above illustrate how some mathematicians claim not to be able to work on an abstract, or purely logico-deductive, level. Due to human error, even professional mathematicians can "go off track and believe some things that are not true". As Thurston (1994) notes, mathematicians "are not good at checking *formal correctness* of proofs, but they are quite good at detecting potential weaknesses or flaws in proofs" (p. 169, emphasis is the author's). Checking the logic with examples and other forms of background knowledge (M1 mentions the construction of diagrams above) appears crucial for some mathematicians to reliably understand and validate a proof.

4. A model for how mathematicians read proofs

Based upon the interview data and the philosophical literature, we posit that mathematicians may understand a proof in three different ways: as *a cultural artifact*, a *sequence of inferences*, or as *the application of methods* (Rav, 1999). The way in which a proof is viewed influences how the proof is read and evaluated. We do not claim this model to be a hierarchy; in some cases, we believe that the reader may switch between these different conceptions of the proof while reading it.

4. 1. Proof as cultural artifact

In order to *evaluate the validity of a given argument*, at least some mathematicians rely on evidence that is not directly related to the content of the argument, but rather the contextual history the argument underwent to be sanctioned as a proof by the mathematical community. For instance, in this paper, we present two mathematicians who claimed that when they read a proof in a journal, they act on the assumption that the proof is correct. Similarly, in Weber (2008), one mathematician claimed, "to be honest, when I read papers, I don't read the proofs … if I'm convinced that the result is true, I don't necessarily need to read it, I can just believe it". In these cases, the mathematicians appear to be saying that since *other mathematicians* checked the proof and claimed it was valid, they were prepared to believe the proof was valid.

This does not appear to be atypical. For instance, Jackson (2006) described one mathematician who believed Perelman's proof of the Poincare conjecture "must be right" because (i) if it was not, the collective expertise of the mathematical community would have found the mistake and (ii) Perelman's work had been reliable in the past (p. 899). His evaluation of the validity of Perelman's proof did not come from the deductive process that Perelman employed, but from the authority of the mathematical community and Perelman himself (cf., Inglis & Mejia-Ramos, 2009). Note that we are not claiming that mathematicians will believe a mathematical assertion is true because an expert in their field *claimed* it was so; rather they will believe a proof is valid because it was *validated* by mathematicians who presumably have the expertise to locate a fault in the proof if one existed.

4. 2. Proof as a sequence of inferences

As one would expect, in order to *comprehend* or *evaluate the validity* of a given argument, mathematicians seem to use what is commonly described as line-by-line reading (see Weber, 2008). Rav (1999) and Arzarello (2007) conceptualize a proof of a theorem as a series of claims of the form $A_0 \rightarrow A_1 \rightarrow A_2 \rightarrow A_{3\ldots} \rightarrow A_n$, where A_n is the theorem. Such an argument consists of n inferences, where the i^{th} inference is $A_{i-1} \rightarrow A_i$.

Mathematicians seem to read proofs line-by-line in order to *comprehend* each of their inferences. As Rav (1999) notes, when mathematicians do not understand how a statement A_i in a proof follows from previous inferences, they may construct a sub-proof of A_i, i.e. a series of valid inferences $A_{i-1} \rightarrow A_{(i-1)1} \rightarrow A_{(i-1)2} \rightarrow A_{(i-1)3} \rightarrow_{\ldots} \rightarrow A_i$. This would amount to using deductive evidence to verify that A_i follows validly from previous assertions. However, mathematicians also seem to employ non-deductive arguments to reach this goal. Six of the nine mathematicians in this study claimed to use examples to help them understand a proof. Indeed, several of them described a process by which the line-by-line reading of the proof would be accompanied by a parallel study of one or more specific examples.

Similarly, when *evaluating the validity* of a given proof, a mathematician may (intuitively and implicitly) assign a probability p_i to his or her level of confidence that the i^{th} inference of the proof is correct. The probability that every inference in a proof is correct—that is, that the proof is fully valid—is then the product of each of the probabilities assigned to each inference, that is ($p_1p_2p_3\ldots p_n$). If the proof is short and in a relatively simple domain, it is possible to become (nearly) absolutely certain that every step within a proof is correct. However, as DeMilo, Liptus, and Perlis (1979) argue, for lengthy proofs in complex domains, there will be a non-trivial probability that some of the assertions in the proof do not follow validly from previous claims. Indeed, Davis (1972) and Hanna (1991) claim that half of the published proofs contain logical errors and many proofs are rife with errors.

One way to increase one's confidence that a proof is correct is to find inferences in the proof that are problematic (i.e., have a probability value below a certain threshold) and examine them more closely to increase one's confidence in them. For instance, suppose that one only assigned a value of 0.9 that the third assertion of a proof was correct. A mathematician might construct a sub-proof with the third assertion as claim (as described above), where his or her confidence in each of the sub-proof's inferences is very high. This would be using deductive evidence to increase one's overall probability that the original proof does not contain any logical flaws. However, this is not always how mathematicians behave. Weber (2008) observed that mathematicians occasionally used empirical evidence to increase their confidence about a particular inference within a proof. For instance, one participant was uncertain about a particular assertion within a proof, but gained enough confidence to judge that assertion (and the proof in its entirety) as valid after verifying the assertion held for a single example. There were other instances in Weber's study where mathematicians increased their conviction of particular assertions within a proof because they noticed a pattern, searched for a counterexample but failed to find one, or produced a generic proof. These results are corroborated by the data reported in this paper. Although no mathematician claimed to verify an assertion only by looking at an example, all claimed that inspecting examples increased their confidence that a proof is correct. As M5 noted, "When I'm looking through a proof, I can go off-track or believe some things that are not true. I always use examples to see that makes sense."

4. 3. Proof as the application of methods

We do not believe that mathematicians' comprehension of a proof amounts to their comprehension of each step in the proof, or that their confidence in a theorem is equivalent to their confidence that every step within its proof is valid. Extracting ideas that could be useful in their own research (a process that mathematicians in this study considered to be crucial) seems to be more sophisticated than what we have described as line-by-line reading. Similarly, it seems unlikely that proofs are normally checked solely by appeals to authority and line-by-line reading. Davis (1972) and Hanna (1991) estimate that half of the published proofs in mathematics contain logical errors, but also argue that most of the theorems published in the literature are true. This would imply for an arbitrary theorem T and published proof P, we could believe P was completely valid (i.e., each inference was valid) with probability 0.5 yet believe T was true with a probability of greater than 0.9.

In the previous sub-section, we argued that mathematicians may attempt to understand and increase their overall confidence in a proof by viewing the proof as a series of inferences and looking carefully at each inference within a proof. Figuratively speaking, we say the mathematicians attempt to reach these goals by *zooming in* on the problematic parts of the proof. We conjecture that mathematicians also reach these goals by *zooming out* and looking at the high-level structure of the proof and thinking carefully about, not the individual inferences, but the *ideas* or *methods* in the proof.

Rav (1999) discusses how mathematicians do not focus on the logical details of the argument they are reading or even the theorem being proved. Rather, they look at the mathematical machinery being used to deduce new results from established ones. We believe that mathematicians read proofs to locate these methods, and that it is these methods that are useful to participants in their own professional work. Rav (1999) further argues that the reliability of proof does not stem from its *logical components*, but from its *methodological components* (p. 29). In addition to, or perhaps instead of, viewing proofs as a lengthy sequence of derivations $A_0 \rightarrow A_1 \rightarrow A_2 \rightarrow A_3 \ldots \rightarrow A_n$, we hypothesize that mathematicians might encapsulate strings of derivations into a short collection of methods and determine whether these methods would allow one to deduce the claim that was proven. In other words, mathematicians viewed the proof as the application of a sequence of methods. Konoir's (1993) analysis of the structure of written proofs supports our claim. He found that proofs were often written with cues indicating to the reader how the proof should be partitioned and what methods were being applied in each partition. Similarly, in describing how he reads a proof in a field that he is familiar with, Thurston (1994) made the following remarks:

> I concentrate on the thoughts that are between the lines. I might look over several paragraphs or strings of equations and think to myself, 'oh yeah, they're putting in enough rigamorole to carry out such-and-such idea'. When the idea is clear, the formal set-up is usually unnecessary and redundant—I often feel that I could write it out myself more easily than figuring out what the authors actually wrote (p. 367).

M5 indicates that, due to time constraints, he sometimes behaves similarly. Several participants cited the benefits of breaking a proof into modules as a tool they used to understand proofs better. We believe that much of proof comprehension may consist of coordinating inspections of the logical details of a proof while zooming out to see these logical details in the context of the methods they are being used to support.

Our hypothesis that mathematicians validate proofs by examining whether the general methods in the proof would work rather than if each step in the proof was valid is also supported by Manin's (1977) and Hanna's (1991) claim that mathematicians evaluate a proof more by the plausibility of the argument in the proof than by its logical details and Thurston's (1994) claim that mathematicians can detect errors in a proof by thinking carefully about the ideas in the proof rather than validating its formal correctness. If this hypothesis is correct, it explains how theorems that appear in journals are usually true even if the proofs that accompany them are often flawed. High-level arguments could be correct despite logical flaws in the detail in the proof, and mathematicians are much better at evaluating the former than the latter; in fact, some may not spend much time on the latter task at all.

It is difficult to say how mathematicians "zoom out" from a proof by encapsulating particular strings of inferences of a proof into methods and then determine if those methods are valid. We suspect that this process gets at the heart of mathematical reasoning and is as cognitively complex as any task in mathematics. We imagine this process would have to involve structural-intuitive evidence, using one's intuition about the mathematics being discussed to see how the mathematical methods being used would work in one's mental models. Kreisel (1985) refers to the process of comparing a formal mathematical argument with one's mathematical background knowledge as "cross-checking" and Kreisel (1985), Thurston (1994), Otte (1994), and Rav (1999) argue that this process is essential in evaluating a proof.

5. Discussion

5. 1. Summary

In this paper, we present a model for how mathematicians read the proofs of others, both to check for correctness and to gain insight. Mathematicians' confidence in a proof can be increased in three ways: (a) if the proof appears in a reputable source, some mathematicians will believe the proof is likely to be correct, (b) mathematicians will verify that each line in a proof validly follows from previous assertions, and (c) mathematicians may evaluate the overarching methods used in the proof and determine if they are appropriate, in addition to, or perhaps instead of, inspecting each line of the proof. Also, mathematicians attempt to understand proofs by studying not only (b) how each assertion in a proof follows from previous assertions, but also (c) what overarching methods in a proof could be useful to them to prove conjectures in their own work. What is interesting is that each of these ways seems to rely, in part, on non-deductive evidence. For (a), trusting that a proof is reliable because it appears in a reputable source involves deferring to the authority of the editor of the journal and the reviewers who certified the proof as valid. While (b) can involve marshalling deductive evidence (i.e., the construction of sub-proofs), the data presented in Weber (2008), and supported by the data presented in this paper, show mathematicians sometimes rely on empirical evidence also. We conjecture that (c) also involves the use of structural-intuitive evidence. As the model was derived largely on the basis of introspective interviews with nine mathematicians, more research is needed to corroborate the claims made by this model, both in terms of examining more mathematicians and perhaps using different methodologies, such as cognitive or ethnographic studies of mathematicians' practices when reading proofs.

Boero (1999) describes a six-stage (non-linear) process from which an individual initially encounters a mathematical situation and concludes by producing a proof of a theorem. We argue that in the mathematical community, the process does not end there. The proof is submitted for review, published in a reputable source (if it is judged to be of sufficient quality), and read by mathematicians as a means to obtain conviction and insight. Just as Boero (1999) observes that non-deductive argumentation plays a role in many of the stages of forming a conjecture and proof (although not in the actual written proof), we argue that non-deductive argumentation plays an interesting role in the evaluation and comprehension of proof.

5. 2. Suggestions for revising the goals of instruction and directions for future research

As noted in the introduction, many researchers argue that the teaching of mathematics should be informed by the beliefs and values of mathematicians, especially with respect to proof (e.g,. Harel & Sowder, 2007; RAND, 2003). We conclude this paper by offering two tentative suggestions for the goals of instruction and posing research proposals based on these suggestions. First, it is often lamented that students believe a proof is correct because a teacher said it was or because it appeared in a textbook (e.g., Harel & Sowder, 1998). However, we believe the students' behavior is consistent with the practice of mathematicians. The data presented in this paper suggest that mathematicians are likely to accept an argument as valid if it appears in a reputable source with a reliable reviewing process. It would seem that students who believe a proof is correct because it appears in a textbook and is sanctioned by their teacher are acting rationally. Proofs in textbooks are usually correct. In our view, the problem with relying *only* on an authoritative source is *not* that students might believe some things that aren't true, but that they will fail to gain comprehension or insight from the proof. One potential implication for research is that students might come to better appreciate a proof not by leading them to see that authoritative sources may be mistaken, but by seeing the extra value or insight that understanding a proof may provide for them (cf., Harel, 2001).

Similarly, many researchers lament that students frequently use empirical evidence to increase one's confidence in a proven theorem. In a famous study, Fischbein (1982) found that, after reading a proof of a theorem in number theory and finding it valid, many high school students still wished to check the theorem held for several examples. Several authors remarked that students failed to understand the generality of proof—checking examples after reading a proof is superfluous because proven theorems cannot have counterexamples (e.g., Fischbein, 1982; Harel & Sowder, 1998). However, we suspect that just as mathematicians cannot gain full confidence that a complicated proof is completely correct solely by inspecting the logic of the proof, many students who are in the process of learning about proof will also be unable to tell with certainty if a proof is correct. If so, it would make sense for students to verify the theorem with examples, as it supplies additional evidence that the theorem is true. As de Villiers (1990) notes, and as our data suggest, such actions are consistent with the practice of mathematicians. While we agree that students should recognize the limitations of empirical reasoning and the generality of a deductive proof, we also think students and teachers should be aware of the role of empirical reasoning in gaining conviction of a theorem and checking the validity of a proof. One implication for the design of learning environments is that students might be taught how to use examples to increase their conviction in, or understanding of, a proof in the same way that the mathematicians in this paper described the ways that they read proofs.

Acknowledgments.

The research in this paper was supported by a grant from the National Science Foundation (NSF #DRL0643734). The views expressed here are not necessarily those of the National Science Foundation. We would like to thank Sean Larsen for helpful insights on developing our theoretical model, as well as the anonymous reviewers for their comments on earlier drafts of this manuscript.

References

Alcock, L. & Weber, K. (2005). Proof validation in real analysis: Inferring and evaluating warrants. *Journal of Mathematical Behavior, 24*(2), 125-134.

Arzarello, F. (2007). The proof in the 20th century: From Hilbert to automatic theorem proving. In P. Boero (ed.), *Theorems in schools: From history, epistemology, and cognition to classroom practice.* Rotterdam, The Netherlands: Sense Publishers.

Balacheff, N. (2002). The researcher epistemology: A deadlock from educational research on proof. In F.L. Lin (ed.) *2002 International Conference on Mathematics Education- Understanding proving and proving to understand*. Taipei: NCS and NUST.

Boero, P. (1999, July/August). Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education. *International Newsletter on the Teaching and Learning of Mathematical Proof.* Retrieved August 2, 2010, from http://www.lettredelapreuve.it/OldPreuve/Newsletter/990708Theme/990708ThemeUK.html

Davis, P. (1972). Fidelity in mathematical discourse: Is one and one really two? *American Mathematical Monthly, 79*, 252-263.

De Milo, R., Liptus, A., & Perlis, A. (1979). Social processes and proofs of theorems and programs. *Communications of the ACM, 22*, 271-280.

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

Duval, R. (2007). Cognitive functioning and the understanding of the mathematical processes of proof. In P. Boero (ed.) *Theorems in schools: From history, epistemology, and cognition to classroom practice.* Rotterdam, The Netherlands: Sense Publishers.

Fischbein, E. (1982). Intuition and proof. *For the Learning of Mathematics, 3*(2), 9-18.

Hanna, G. (1991). Mathematical proof. In D. Tall (ed.) *Advanced Mathematical Thinking.* Dordrecht: Kluwer.

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

Harel, G. (2001). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. Campbell & R. Zazkis (Eds.) *Learning and teaching number theory*. New Jersey: Ablex Publishing Company.

Harel, G. & Sowder, L. (2007). Towards a comprehensive perspective on proof. In F. Lester (Ed.), *Second handbook of research on mathematical teaching and learning*. Washington, DC: NCTM.

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.

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

Inglis, M., Mejia-Ramos, J. P., & Simpson, A. (2007). Modelling mathematical argumentation: The importance of qualification. *Educational Studies in Mathematics, 66,* 3–21.

Jackson, A. (2006). Conjectures no more? Consensus forming on the proof of the Poincare and geometrization conjectures. *Notices of the American Mathematical Society, 53*, 897-901.

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

Kreisel, G. (1985). Mathematical logic: Tool and object lesson for science. *Synthese, 62*, 139-151.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Education Research Journal, 27*, 29-63.

Maher, C., Muter, E., & Kiczek, G. (2007). The development of proof making by students. In P. Boero (Ed.), *Theorems in Schools*. Rotterdam: Sense Publishers.

Manin, Y. (1977). *A course in Mathematical Logic*. New York: Springer Verlag.

Martin, W. G. & Harel, G. (1989). Proof frames of pre-service elementary teachers. *Journal for Research in Mathematics Education*, 20(1), 41-51.

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.

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

RAND Mathematics Study Panel (2003). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education*. Santa Monica, CA: RAND corporation.

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

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

Stylianides, A. & Stylianides, G. (2008). Enhancing undergraduates students' understanding of proof. In *Proceedings of the 11th Conference for Research in Undergraduate Mathematics Education*. Last downloaded March 18, 2009: http://www.rume.org/crume2008/Stylianides_LONG(21).pdf.

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

Weber, K. & Alcock, L. (2005). Using warranted implications to understand and validate proofs. *For the Learning of Mathematics, 25*(1), 34-38.

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