

Lessons Learned from an Instructional Intervention on Proof Comprehension

Aron Samkoff

10 Seminary Place, New Brunswick NJ 08901 USA

aron.samkoff@rutgers.edu

Keith Weber (corresponding author)

10 Seminary Place, New Brunswick, NJ 08901 USA

keith.weber@gse.rutgers.edu

(848) 932 0804

Corresponding author's institution: Rutgers University

Abstract. In previous research, we proposed a set of proof reading strategies that we hypothesized can help students better understand the proofs that they read. The goal of the present paper is to report lessons that we learned from two instructional interventions in which students were taught to apply these strategies. We found suggestive evidence that implementing these strategies helped students understand the proofs that they read, but also found students' implementation of these strategies to sometimes be problematic. We present instructional modifications, as well as refinements to the strategies themselves, that enabled the students to implement the strategies more effectively.

Keywords: Proof, proof comprehension, proof reading strategies

Introduction

Mathematical proof plays a fundamental role in mathematicians' practice. Proof is the primary means by which mathematicians demonstrate that a theorem is true. Further, proofs can serve as bearers of knowledge, both by providing insight for why a theorem is true, and by illustrating problem solving methods that can be used to prove other statements (Rav, 1999; Steiner, 1978). Proofs also play an important role in advanced mathematics courses-- i.e., the upper-level proof-oriented university classes for mathematics majors. In these courses, proofs are the dominant form of pedagogical explanation (cf. Lai & Weber, 2014; Lai, Weber, & Mejia-Ramos, 2012), with Mills (2011) estimating that half of the time in mathematics classrooms is spent dealing with proofs. Interviews with mathematics professors indicate that the proofs presented in classrooms are presented so that students might gain techniques and understanding, and not merely conviction that the theorem statement is true (Weber, 2012; Yopp, 2011).

Despite these efforts, studies have suggested that students have significant difficulties with reading¹ the proofs that are presented to them. Mathematics majors are largely unable to determine if a proof is correct (Alcock & Weber, 2005; Ko & Knuth, 2013; Inglis & Alcock, 2012; Selden & Selden, 2003; Weber, 2010) and researchers have remarked that mathematics majors seem to have serious difficulties understanding proofs (Conradie & Firth, 2000; Cowen, 1991). The empirical research on proof reading has largely consisted of either measuring mathematics majors' success at determining if an argument is valid (e.g., Alcock & Weber, 2005; Ko & Knuth, 2013; Inglis & Alcock, 2012; Selden & Selden 2003; Weber, 2010) or asking students whether an argument is convincing and using these responses to provide insight into

¹ By "reading a proof", we are referring to looking at the proof with the intent of developing meaning from it. By "understanding a proof", we mean developing the facets of the Mejia-Ramos et al. (2012) model which we describe shortly. Reading is a behavioral activity while understanding is a cognitive state. Hence, students will read proofs with the aim of understanding them, and they may or may not be successful at achieving this goal. As we argue in this paper, the research literature suggests that mathematics majors frequently do not understand the proofs that they read.

students' standards of conviction (e.g., Martin & Harel, 1989; Healy & Hoyles, 2000; Segal, 2000; Weber, 2010). Mejia-Ramos and Inglis (2009) argued that there has been comparatively little research on students' *comprehension* of proofs. Because undergraduates encounter many proofs in their advanced mathematics courses that are supposed to increase their mathematical understanding, there is an urgent need for research on ways in which students can be successfully taught to improve their reading comprehension. Our goal in this paper is to contribute toward this gap in the literature.

In our previous work, we identified proof reading strategies (Weber, in press; Weber & Samkoff, 2011) that we hypothesized had the potential to improve mathematics majors' comprehension of the proofs that they read. We used this research as a starting point for an instructional experiment where we tried to teach students these strategies using a modeling-scaffolding-fading approach (Brown et al, 1989). By analyzing this intervention, we provide suggestive confirmatory evidence that some of these strategies can improve some students' proof comprehension. We also illustrate that sometimes students do not benefit from using these strategies, largely because students could not implement these strategies correctly. In these cases, we present an analysis of why students failed to correctly implement the strategies and suggest modifications to how the strategies were introduced to students or to the strategies themselves that might be beneficial for future rounds of the study.

Theoretical perspective

In this paper, we incorporate the model of Mejia-Ramos et al. (2012) for characterizing and assessing students' understanding of a proof in advanced mathematics. In this model, there are seven dimensions to understanding a proof: three "local" and four "holistic" dimensions. The

local dimensions deal with understanding that can be gleaned from carefully reading a small number of statements in the proof, and consist of the following:

- 1) Meaning of terms and statements: Understanding the meaning of terms and individual statements of the proof. This includes stating the definitions of terms used in the theorem statement and proof and identifying trivial implications of a given statement.
- 2) Justification of claims: Understanding why each claim made in the proof follows from previous ones, and being able to identify claims that follow from a given statement later in the proof.
- 3) Logical status of statements and proof framework: Understanding the logical relation between the assumptions and conclusions in a proof, identifying the proof technique being used, and conceptualizing the proof in terms of its proof framework (cf. Selden and Selden, 1995).

“Holistic” ways to understand a proof concern synthesizing the entire proof or entire parts of the proof as a coherent whole, and include:

- 1) Identifying the modular structure: Understanding how a proof can be broken into mathematically independent parts or sub-proofs, and how these parts logically relate to one another.
- 2) Illustrating with examples: Understanding how a sequence of inferences can be applied to verify that a general theorem is true for a specific example.
- 3) Summarizing via high-level ideas: Understanding the overarching logical structure of the proof and being able to summarize a proof in terms of these ideas.

- 4) Transferring the general ideas or methods to another context: Being able to use the ideas or methods in the proof to establish a different theorem.

This model was used to specify the learning goals of the instructional experiment; we wanted students to improve their understanding of proofs as judged by this model. The model was also used to generate assessment questions that were asked to students after each proof, which was done to provide students with feedback on their own understanding. This is further described in the methods section of the paper.

Strategies for understanding proofs

Weber (in press) observed two pairs of successful mathematics majors trying to understand six mathematical proofs and identified six potentially useful strategies that these students used along with a theoretical rationale for why these strategies might improve comprehension. Weber (in press) then surveyed 83 mathematicians about whether they desired that mathematics majors use these strategies. For five of these strategies, most mathematicians claimed that they desired their students use these strategies. When Weber and Mejia-Ramos (2013a) surveyed 175 mathematics majors about whether they used these strategies when reading proofs, there was no strategy that the majority of mathematics majors claimed that they used. Consequently, Weber and Mejia-Ramos (2013a) conjectured that mathematics majors' understanding could improve if they could be taught to apply these strategies. In the current paper, we explore this conjecture.

In addition to these five strategies, we also included some strategies that were suggested in the literature on proof. Below, we present each strategy that we used, as well as a description of what the successful implementation of each strategy would entail and why we anticipate the

strategy will be beneficial. We categorize these strategies as those one can use before reading a proof, while reading a proof, and after reading a proof.

Pre-reading strategies, or strategies to apply after one reads the statement of the theorem and before reading its proof:

1. *Make sure you understand the definitions of all the terms:* The student will try to state the definition of any term used in the theorem statement. This strategy was included in response to Conradie and Firth's (2000) observation that students frequently read proofs without being able to state the definitions of the involved concepts. We anticipate this strategy can help students with the Meanings of Terms and Statements aspect of the proof comprehension model. We note that being able to state the definition of a term is clearly not a sufficient condition for understanding the meaning of that term. Bills and Tall (1998) called a definition "formally operable" for a student when that student "is able to use it in creating or (meaningfully) reproducing a formal argument". Selden (2012) and Harel, Selden and Selden (2006) argued that undergraduate students need to come to make sense of definitions in this way to more fully understand key concepts in proofs. While we agree with this point, we doubt definitions could become operable for students if they are not first able to recall and consider these definitions.
2. *Illustrate the theorem with an example:* The student will try to instantiate an example in the theorem statement and check that the theorem holds true in this case. This was based on Weber's (in press) observation that their strong students would use examples to understand both the theorem and statements in the proof they found to be problematic and should help students with the Meanings of Terms and Statements aspect of the proof comprehension model.
3. *Describe how you would try to prove the theorem:* This was a strategy highlighted in Weber (in press). We found that when the students tried to prove the theorem, they were led to consider issues such as what proof methods might be used, what should be the assumptions and conclusions of the proof, and what parts of the proofs would cause them difficulty, thereby helping students with the Logical Status of Statements

and Proof Framework aspect of the proof comprehension model. We also found that this motivated students to read these proofs and focused their attention on conceptually interesting aspects of the proof rather than routine details. To avoid misinterpretation, we note that the intent here is *not* have student actually prove the theorem, but only to have students choose a method to see how it might proceed. In their proof validation study, Selden and Selden (2003) asked participants to prove the statement for which they would subsequently read supporting arguments. Not surprisingly, participants were generally unable to produce a proof.

While reading the proof:

4. *See if the proof can be broken into independent parts*: This was a strategy highlighted in Weber (in press). The student will try to identify any logically independent modules in the proof, as well as any sufficient conditions that are established to arrive at the conclusion. This will help students see the overall structure of the proof (in the sense of Leron, 1983), helping with the Modular Structure and Summary aspects of the proof comprehension model, as well as see how different assumptions in the proof were used.
5. *Identify what proof methods are being used*: This was a strategy highlighted in Weber (in press). The student will try to identify the assumptions and conclusions of the proof to determine what proof technique (e.g., direct proof, proof by contradiction, proof by contraposition, proof by induction, or proof by cases) is being used in the proof, which directly leads students to consider the Logical Status of Statements and Proof Framework aspect of the proof comprehension model.
6. *See why confusing statements are true with a particular example*. This was a strategy highlighted in Weber (in press). We also note that mathematicians frequently apply this strategy when they read proofs (e.g., Weber, 2008). This can help students with the Justification of Claims and Illustrate with Examples aspects of the proof comprehension model.

After reading the proof:

7. *Compare the approach you chose to the proof that you read:* This was a strategy highlighted in Weber (in press). The student will be able to note any differences in how they would approach the proof and how the proof proceeded. This will help students understand the overall procession of the proof as well as help them appreciate the strengths of these methods. This may help students with the Transfer the Methods of the Proof aspect of the proof comprehension model.

The goal of our instructional experiment was to teach students to apply these strategies when carefully reading a proof, as they were described above. By analyzing the implementation of this instruction and our participants' application of these strategies, we found suggestive evidence that some of these strategies appeared to increase participants' comprehension of proofs of the specific proofs that they read. We present illustrations that by applying the strategy, the participant developed some insight about the proof they were reading or later commented that the application of the strategy was helpful to them (although we did not verify that students performed better on a post-test, so the evidence here is merely suggestive). We also found other strategies were problematic for the participants to implement. When participants had difficulties implementing a strategy, we provide hypotheses for why these difficulties arose and suggest improvements for future iterations of our study.

The design research paradigm and “instructional experiments”

This study employs many of the principles of design research (Cobb et al., 2003) to develop an instructional intervention with the goal of having students to apply the strategies listed in this paper to improve their abilities to understand proofs. The goal of design experiments is to develop effective teaching methods while simultaneously developing theories for how students can learn the target domain. Following Cobb et al. (2003), we conceptualize design research as forming initial hypotheses on the following: (i) how students reason upon entering the classroom, (ii) what appropriate learning goals should be from the point of view of

the researcher (i.e., what it means to understand the concepts in question), (iii) how students might come to achieve these goals, and (iv) how instruction can be designed to help students achieve these goals. An instructional intervention is implemented and this implementation is recorded. Careful observations of this implementation, both between sessions during which this intervention is taking place and after the intervention is completed, allow the researcher to seek confirming or disconfirming evidence of the each of the initial hypotheses. Any observed disconfirming evidence can then be used to refine the hypotheses, and thereby develop better instruction, to achieve the study's instructional goals. This can also advance our understanding of how students learn and what it means to understand a concept. We describe each of our four initial sets of hypotheses below.

How students reason.

The research literature reveals several limitations in how mathematics majors typically read proofs. Weber and Mejia-Ramos (2014) highlighted several undesirable beliefs that mathematics majors hold about proof reading: Mathematics majors believe understanding a proof is tantamount to understanding each inference within a proof (as opposed to the holistic understandings described earlier in this paper). Mathematics majors do not believe they should have to justify statements when reading a proof; a well-written proof will include those justifications. Finally, the majority believe that most proofs they encounter in advanced mathematics can be understood in under 15 minutes. Weber and Mejia-Ramos (2014) also found that mathematicians find it undesirable that students hold such beliefs.

Previous research supports the idea that mathematics majors are limited in the ways they attempt to read proofs: They tend to focus on calculations, rather than the overarching structure of a proof, when reading proofs (Inglis & Alcock, 2012; Selden & Selden, 2013). They do not

spend much effort inferring how new statements in a proof follow from previous ones (e.g., Alcock & Weber, 2005; Inglis & Alcock, 2012). They also are unlikely to generate examples or draw diagrams to help them understand a proof and spend little time reading the proof (Weber, 2009). Further, students are often unfamiliar with the formal definitions of concepts (Moore, 1994) and might read a proof without resolving their difficulties in this regard (e.g., Conradie & Firth, 2000). In summary, this research reveals that students do not attend to pertinent information when reading a proof and that this may be due, at least in part, to their beliefs about what it means to understand a proof and what their responsibility when reading a proof should be.

Learning goals

The goal of this study was for students to come to a more complete understanding of proofs that they read based on the model of understanding by Mejia-Ramos et al. (2012) described earlier in this paper. It is reasonable to expect that what students would gain from reading the proof would be affected by their personal goals for reading the proof. Therefore, we made our desired goals explicit to students at the beginning of each of the two iterations of our study, as described in the Procedure.

How students might come to achieve these goals

We hypothesized that students could become better readers of proofs if they enacted the strategies that experts (mathematicians or strong students) use to read proofs. These strategies were described in the previous section.

How instruction can be designed

Our first attempt to teach students to apply these strategies was through a “cognitive apprenticeship” (Brown et al., 1989). In this paradigm, a mentor models a skill, provides scaffolding to learners as they begin to learn the skill, and gradually “fades” as students become

proficient in the skill. This is all done in the context of authentic tasks; this allows students to perceive the contextual application of the skill (rather than asking students to apply the skill in decontextualized academic or laboratory settings and hoping that the student can later transfer this skill to more authentic tasks). The use of authentic tasks also allows students to appreciate the value of implementing the skill since they can see how it contributes to accomplishing a meaningful higher-order goal. While cognitive apprenticeship provides principles for designing instruction, this perspective does not directly prescribe a specific instructional intervention. To specifically teach the reading strategies described in this paper, we adapted Palinscar and Brown's (1984) "reciprocal teaching" method. Palinscar and Brown introduced "reciprocal teaching" in a classic article on reading comprehension where they reported a series of studies that showed this method could improve children's comprehension of non-mathematical expository texts. In these studies, students were taught to use four reading strategies: 1) summarize the main ideas, 2) ask questions related to the text, 3) clarify difficult parts of the text, and 4) make predictions. These strategies were first defined and modeled by a teacher-researcher. Students were given more responsibility in using the strategies during a guided practice phase. During this phase, the teacher-researcher provided positive feedback for good uses of a strategy, and provided scaffolding in the form of prompts or modifications when their application of these strategies was problematic, remodeling the strategy if necessary. Students took turns playing the "teacher" role by implementing each of the strategies. Gradually, the teacher-researcher's role was diminished as students gained proficiency in implementing these strategies. Palinscar and Brown (1984) reported large and robust learning gains when struggling readers were provided with this instruction. Given the similarity of our instructional goals (i.e., having students implement research-based strategies to improve comprehension of texts) and the

success of the reciprocal teaching studies in multiple environments (e.g., see Doolittle et al., 2006, for a summary of the success of using this method in different domains at the collegiate level), we judged reciprocal teaching to be a useful starting point for instructional design. We describe how this structure was adapted for our purposes in detail in the Procedure.

Instructional experiment

Although our study is based on the broad goals and philosophies of design research—specifically we use an iterative process to improve our understanding of how students can understand proofs, how they should learn, and how instruction should be designed—there are some key differences between our research methodology and that used in a design experiment. First, the instructional method that we used, reciprocal teaching, is not based on a theory of a process for *how* students' understandings transform over time. Rather, reciprocal teaching is what diSessa and Cobb (2004) referred to as a “framework for action”—or heuristic guidelines for designing effective instruction. Second, the bulk of the analysis that we report in this paper is based on retrospective analysis rather than the analysis that occurs as the instruction is progressing. To distinguish between what we did and a design research study, we refer to this as an “instructional experiment” which we view as the first stage to developing an instructional innovation.

Methods

Participants.

Our instructional experiment consisted of two iterations. All participants were recruited from the same large public university in northeastern United States. For the first iteration, students were recruited from an undergraduate real analysis course. Eight students agreed to participate in the study, but attendance varied by meeting. On average, each meeting was attended by six participants. Meetings were conducted once every one to two weeks during the

semester, and each meeting lasted approximately 2 hours. There were four meetings with this group of participants. For the second iteration, two students, Matt and Tony, were recruited from a class preparing mathematics high school teachers. Both students had just finished their fourth year in a five-year program, and had already taken real analysis. The second iteration also met for four sessions. The purpose of the second iteration was to examine in depth the hypotheses that were formed in the first iteration. The more personal format of a two-person case study facilitated this goal, as the focused discussion allowed us greater insight into the thinking of each individual participant.

Materials.

The original proof strategy sheet designed for this study can be found in Appendix A and was used for the first two sessions of Iteration 1. This strategy sheet was modified after the second session of Iteration 1 after new hypotheses emerged from our analysis. The modified strategy sheet is included in Modified proof reading strategies. Appendix B and was used for sessions 3 and 4 of Iteration 1 and for the entirety of Iteration 2. Additionally, eleven proofs for students to read were used for this study. The proofs came from two sources. First, some proofs were obtained from the textbook used for the real analysis course that the students were completing. Other proofs were designed by our research team. We strove to choose proofs with the following characteristics: First, the proofs were accessible to students with the particular background knowledge we had assumed (math majors who were enrolled in or had completed a real analysis course). We did not want students to spend a substantial amount of their efforts attending to unfamiliar concepts, and as such, proofs were chosen to rely on content from basic number theory, calculus, and real analysis (a course all students were completing in Iteration 1 or had completed in Iteration 2). Second, for the purposes of authenticity, proofs were intended to be similar to those they would see during the course of their programs of study. Finally, proofs

were chosen to be both moderately difficult, so that we could observe students encountering real difficulties in understanding the proofs they read, and complex, so that students had opportunities to implement the strategies provided for them.

All proofs were formatted in the same way: Each proof appeared on an individual page with the theorem at the top of the page, each assertion set on a new line, and lines numbered for facilitating discussion. These proofs are included in Appendix C, and the order in which the proofs were presented in each iteration can be found in Table 1. In the remainder of this paper, this paper, we refer to the n th proof in Appendix C as “Proof n ” and the corresponding theorem statement as “Theorem n ”.

	Iteration 1	Iteration 2
Session 1	Proof 1 Proof 2	Proof 1 Proof 2
Session 2	Proof 3 Proof 4 Proof 8	Proof 3 Proof 4 Proof 5
Session 3	Proof 6 Proof 10	Proof 6 Proof 7
Session 4	Proof 11	Proof 8 Proof 9*

Table 1. The order in which the proofs were presented to students in Iterations 1 and 2.

(* - Proof 11 was not used in Iteration 2 as the theorem was proven in a problem-solving class that the pair of students completed. The students in Iteration 1 had not taken this problem-solving course).

Procedure.

Palincsar and Brown’s (1984) “reciprocal teaching” was adapted for teaching these comprehension strategies. In the initial class meeting, the teacher-researcher introduced the study and explained that the students should read proofs for understanding. The teacher-researcher did not clarify what was meant by understanding. The teacher-researcher then modeled each of the strategies he or she wants students to use within the context of Proof 1, a

proof of the claim that a positive integer is square if and only if it has an odd number of divisors. In Iteration 1, each student was handed the strategy sheet in Appendix A. In Iteration 2, both students received the revised strategy sheet in Appendix B. The strategy sheets described each of the strategies that were discussed earlier in this paper. Next, students were given proofs and individual students were asked by the teacher-researcher to apply a specific strategy that the teacher-researcher had modeled. After the student had implemented the strategy, other students were encouraged to provide feedback on whether the strategy had been implemented successfully or whether they had done anything differently. After inviting student participation, the teacher-researcher provided scaffolding by giving feedback to the students on how well they had implemented the strategies, giving positive feedback on aspects of the implementation that had been done well and building off students' contributions if a strategy had been implemented incorrectly. In the latter cases, the teacher-researcher would provide corrective feedback, explaining why the students' approach was inadequate and again modeling the desired behavior. The teacher-researcher progressively reduced this support, allowing students to take increasing responsibility for implementing the strategies.

We provide a few illustrations of the teacher-researcher's feedback. First, in Iteration 1, Session 2, a student attempted to implement the strategy of trying to prove a theorem oneself before reading its proof with Theorem 3 by claiming that he could prove it "by contradiction", perhaps by using "the Squeeze Theorem". The instructor believed that the student did not go into sufficient detail to reap the benefits of this strategy, as important benefits of this strategy include having students actually produce the proof framework and see where their methods might be difficult to implement. Consequently, the instructor prompted the student to think more deeply about how he would prove the theorem by asking what difficulties he thought this

approach would encounter. In a separate instance, during Iteration 1, Session 3, a student noted that Proof 10 was a direct proof but did not say how he made this (correct) observation. The instructor then asked the student how he knew this, leading the student to observe that the last line of the proof ends with the conclusion of the theorem statement. The instructor brought the group's attention to this response and noted that this approach is a good one for determining if the proof is a direct proof. The support for a given strategy gradually decreased as students showed proficiency in adopting the strategy. To help the reader get a sense for how the instruction proceeded, we include a transcript of an in-class discussion in Appendix D.

After students had read and discussed a proof, the written proof was taken away and the instructor asked individual students comprehension questions. These questions were based on the theoretical model of Mejia-Ramos et al. (2012) discussed earlier in the paper and typically included asking a student to summarize the ideas of the proof, provide justification for particular steps in the proof, and determine if the method the proof used could be applied in a similar setting. These questions may have contributed to students' conceptions of what it meant to understand a proof and what types of understandings the teacher-researcher was expecting of the students. Students' responses were evaluated immediately by the instructor, and students were given corrective feedback as necessary. This methodology of giving students frequent assessments was also employed in the Palincsar and Brown (1984) studies. The purpose of these assessments was both to provide feedback to students on their understanding of texts and to provide feedback to the teacher-researcher on the extent to which students were understanding the text.

Throughout both iterations, students were asked for their feedback on the strategies, including which strategies they found helpful and which they found unhelpful, both during the

class meetings and on questionnaires distributed at the end of iteration session. From this feedback as well as their actual implementation of these strategies, hypotheses were made about the effectiveness of these strategies as well as the methods used to teach these strategies. All student work was videotaped so that a retrospective analysis could be conducted.

Analysis.

The sessions were analyzed both retrospectively and in real-time by the instructor. Real-time analysis consisted of noting occasions when there was evidence that students either benefitted or failed to benefit from the use of a particular strategy. Tentative hypotheses were formed to explain students' successes or failures, and corrective action was taken on the basis of these hypotheses if deemed appropriate. After each session, these hypotheses were recorded and were revisited during the retrospective analysis that occurred after each iteration was complete. In the retrospective analysis, each of the videotapes of student work was broken into segments. Each segment consisted of one or more students attempting to implement a strategy in a particular way. Instances in which students expressed confusion over a particular aspect of a proof or a strategy were also coded, as were students' responses to the assessment questions asked after each proof was discussed. In total, 119 segments were coded from the videotapes where the students attempted to implement a strategy that was appropriate for that proof, with the time length of each segment ranging from 30 seconds to several minutes.

For each segment, we noted the strategies to which students were attending, as well as the particular actions taken by the student (e.g., reading, listening to their classmates) and the instructor (e.g., clarifying a strategy, prompting students to use the next strategy). We also coded whether a strategy was relevant to the proof. For instance, for Theorem 3 (There is a real number whose fourth power is exactly one larger than itself) and Proof 3, the strategy of illustrating the

proof with an example would not be relevant (Proof 3 is not constructive), nor would the strategy of breaking the proof into independent sub-parts (Proof 3 is only 4 lines long). The segments that were not relevant were not included in the 119 segments for subsequent analysis. We then coded the quality of students' implementation of the strategy as (i) complete and correct, (ii) partially correct, or (iii) superficial or incorrect. As an example, the strategy of "partitioning a proof" would be appropriate when applied to a proof that consisted of proving two directions of an "if and only if" statement. If a student correctly partitioned the proof into the two separate sub-proofs, this would be coded as a "complete and correct" implementation of the strategy but if students claimed that the proof could not be partitioned or partitioned the proof incorrectly, this segment would be coded as a "superficial or incorrect" implementation. The "partially correct" code was used when there were correct aspects of a student's implementation but other aspects were either incomplete or incorrect. For instance, in identifying the proof methods used in a proof, a "partially correct" implementation could consist of correctly identifying the method in one part of the proof, but failing to identify the correct method in another part. Finally, it was noted whether there was any evidence of students' benefitting from their use of the strategy, as well as a possible explanation for why students benefited. Evidence consisted of excerpts from students' dialogues as they discussed the proof as well as student feedback on the helpfulness of the strategy sheet.

We were particularly interested in instances in which the strategy was coded as relevant but the implementation of the strategy was not coded as "complete and correct". We revisited each of these instances and developed hypotheses for why students failed to implement the strategy appropriately.

Results

The result of the analysis was a table of instances in which students attempted to implement strategies, their successes and difficulties in implementing the strategies, and hypothesized causes for these successes and difficulties. Table 2 presents the number of times each strategy implementation was coded as “complete and correct,” “partially correct,” and “superficial or incorrect”.

The results section is organized by strategy. For each strategy, we present the themes that emerged, illustrating these themes with excerpts of student work from Iterations 1 and 2. In cases where we identified students’ difficulties in implementing a strategy, we provide a hypothetical cause for these difficulties, and describe what steps, if any, we took to alleviate these difficulties, and the perceived effects of making these modifications. In instances where we did not alter instruction to improve students’ implementation of the strategy, we hypothesize actions that we could take in future iterations of this study. Data from Iterations 1 and 2 are considered together, and the results are organized by strategy.

Strategy	# Complete and correct implementations	# partially correct implementations	# superficial or incorrect implementations
Understand definition	8	3	12
Illustrate theorem with example	13	1	2
Try to prove before reading proof	6	8	10
Compare proof approach to one’s own approach	10	0	5
Partition proof	9	0	8
Identify proof methods	15	3	1

Instantiate lines of proof with example	3	2	0
---	---	---	---

Table 2. Counts of occurrences of different qualities of implementation for the strategies.

Strategy 1: Understand the definitions of terms in the theorem statement.

This strategy was coded as relevant for all theorems except Theorems 2 and 3, which did not contain advanced mathematical terms. For this strategy, attempts were coded as “complete and correct” when students correctly recalled the definition of a concept, “partially correct” for flawed accounts of definitions, and “superficial or incorrect” when the definition contained major errors or no definition was provided. . There were 8 “complete and correct”, 3 “partially correct”, and 12 “superficial or incorrect” implementations recorded. We first describe a case where this strategy appeared to be helpful for students. An instance where this was particularly salient was in Iteration 2, Session 3, where Matt and Tony were reading Proof 7, a proof of the claim that e is irrational. This proof relies on the power series definition of e and begins by assuming, for contradiction, that e is rational. After seeing the theorem statement, Matt and Tony considered applying this strategy:

Tony²: Do you want to write the definition of e ?

Matt: No. [laughs]

Tony: I don’t know if it will be used, I don’t think it will be.

[from Iteration 2, Session 3, Proof 7]

Matt agreed with Tony, suggesting that these students did not feel it was appropriate to implement this strategy because they did not feel the definition of e would be used in the proof. (This is consistent with the findings of Conradie and Firth (2000), in which these authors remarked that students would read a proof without knowing the definitions of the involved concepts as well as illustrating an interesting unproductive belief about students and the

² All names are pseudonyms. Quotes from students have been lightly edited for readability. At no point were words added and we do not believe we changed the meaning or spirit of the utterances.

importance of definitions). The instructor then urged the students to generate a definition for e , after which Tony wrote down the limit as n approaches infinity of $(1+1/n)^n$. The instructor asked if they remembered any other definitions, and they recalled there was a “summation” definition but were unable to remember it. The instructor then provided them with the power series definition of e , and the students proceeded to implement the next strategy of thinking how they would prove the theorem. Matt and Tony discussed how the proof might look, comparing it to a proof of the square root of 2 being irrational:

Tony: Would you do a contradiction, assume it is rational?

Matt: Assume it is rational.

Tony: Then it could be written as a over b . I remember the square root of 2 proof going like this—

Matt: --Yeah but, the thing is when you square the square root of 2, you get 2. I mean what are we going to...what is our contradiction going to come from?

Tony: Oh right, that's where that comes from.

Matt: Because like you're just going to get eb equals a . [laughing] What is that going to tell us?

Tony: Good point.

Matt: There's nothing we know how to work with, like the square root.

[from Iteration 2, Session 3, Proof 7]

This excerpt illustrates how students were able to generate one definition of the mathematical constant ' e '. After this excerpt, these students expressed that they were expecting a definition of ' e ' to be used somewhere in the proof, and expressed curiosity as to how the proof was going to use this definition. In the post-interview, both Matt and Tony identified this strategy as being particularly helpful for them in real analysis. In Tony's words, “The definitions aren't easy to remember, and a lot of times in proofs, just as long as you can remember the

definitions, you've pretty much got it." Nevertheless, Matt admitted to not always implementing this strategy when reading proofs:

Matt: I think I'm guilty of reading proofs and not really understanding the definitions fully. Like you just take it for granted that you think you know it.

[from Iteration 2, Session 3, post-interview]

These excerpts suggest that attending to the definitions of concepts can not only help students understand the meaning and terms of statements, but can be critical to students' understanding of the nature of proof in general.

In the first two sessions of Iteration 1, there were three instances of students merely identifying the key terms in the theorem statement without stating their definitions. Clearly this would not help students *understand* the terms, a primary reason for having students list the definitions. We believed they, like Matt, may have assumed that they knew the definitions of these terms. As a result of our observations, we modified the strategy sheet immediately after the second session of the first iteration, making it more explicit that students should actually state the definition, and if they realized that they were unable to correctly state the definition, to consult a teacher or textbook. The modified sheet is given in Appendix B. There were no recorded instances of students only identifying terms occurring after the second session of the first iteration, suggesting that our modifications alleviated this problem.

The previous excerpt also revealed to us a complication with this strategy. Although the students were able to recall one definition of e , their definition was not particularly useful for the proof we chose. The teacher-researcher found it necessary to encourage the students to seek another formulation. When concepts are defined or conceptualized in multiple ways (e.g., there are many formulations of normal subgroups in group theory), even a proper implementation of

this strategy does not guarantee that students will recall the relevant definition. This finding suggests that an important role that a teacher could play is stating which definition of a concept is relevant for the proof that is being written, acknowledging that there are multiple definitions that can be chosen, and explaining why the particular definition was suitable for this proof.

Strategy 2: Illustrate the theorem with an example.

This strategy was coded as relevant for all of the theorems except Theorem 7, which claimed that e is an irrational number. An implementation of this strategy was coded as “complete and correct” when a student instantiated a theorem statement with a particular example, and verified that the theorem held true for this example. “Partially correct” responses involved a student claiming that an object was an example, but failing to verify that the theorem held true for this example. “Superficial or incorrect” implementations occurred when a student was unable to think of a valid example instantiating the theorem statement. There were 13 instances coded as “complete and correct”, 1 coded as “partially correct”, and 2 coded as “superficial or incorrect”. We describe two instances from Iteration 2 in which students’ use of this strategy led them to investigate why theorem might be true in general. In Iteration 2, Session 4, Matt and Tony implemented this strategy by computing the first few terms of the sequence s_n in Theorem 8, defined by

$$s_1 = \sqrt{2}, s_{n+1} = \sqrt{2 + s_n} \text{ for } n \geq 1$$

Immediately after doing this, Tony remarked, “all right, well it looks like it’s getting closer to two. Let’s see if we can get past two.” After checking the next few terms of the sequence, Tony remarked, “what’s happening here? We’re adding two, but it’s never actually going to get past four. So we just keep on taking square roots, it should always be below two.” This logic is close to that used in the actual proof in showing s_n is bounded. Hence, this can help students appreciate

the method used in the proof and see the overarching logic used in the proof, which is important for providing a summary of the proof.

A similar instance occurred in Iteration 2, Session 2, where Matt and Tony implemented this strategy for Theorem 3, that there is a real number whose fourth power is exactly 1 larger than itself, by raising different numbers to the fourth power. After evaluating 1^4 and 2^4 , Tony quickly realized that 1 was “too small” and 2 was “too large”, and that there must be a number between 1 and 2 satisfying the claim. Again, this was similar to the logic employed in the actual proof, which formalized this idea by employing Rolle’s Theorem on the appropriate function.

There was also a case in which students revealed a key misconception about a mathematical concept in the course of implementing this strategy. Though this was not one of our original research goals—indeed, we deliberately chose proofs where we thought students’ concept understanding would be unproblematic—there were nonetheless instances when this strategy appeared to help students in this manner. In Iteration 2, Session 4, Matt and Tony attempted to implement this strategy for Theorem 9, a claim pertaining to subsequences. These students initially struggled to recall the definition of a subsequence, and even after this definition was provided to them, they continued to grapple with the concept by considering subsequences of the sequence $1/n$:

Tony: OK. $1/n$. Positive integer, so our first few terms are 1, $\frac{1}{2}$, [writing], $\frac{1}{3}$, $\frac{1}{4}$ and so on, all the way up to $1/n$, it’s an infinite series, so actually it should go further. Right? If we want to break it down into subsequences...

Matt: Mm-hmm. We can take uh...1 over 2—

Tony: --we could break it down into sub—like uh, really any number of those, we could just take any number of those sequences, right? So I could say like—

Matt: Wait, couldn't we like—

Tony: $1/1$, $1/2$, $1/3$ as my first one, and the rest of them? Or do they both have to be infinite?

[from Iteration 2, Session 4, Theorem 9]

In the above excerpt, we see that Tony, in attempting to implement this strategy, revealed a key misunderstanding about subsequences—that they include finite as well as infinite sequences. This gave the teacher-researcher the opportunity to intervene and address this misunderstanding before students saw the proof. The instructor pointed out that subsequences were not allowed to be finite, after which Matt suggested they look at subsequences of the form $1/2p$ and $1/(2p+1)$, and promptly verified that the theorem statement held in this example:

Matt: So why don't we take an even denominator and an odd denominator.

Tony: OK, that would both be infinite. Good.

Matt: So for our first one, $1/2p$...[both writing]

Tony: That...see that helps, because [inaudible] it is kind of...now I'm remembering a little bit.

Matt: So if we show...

Tony: Do these sequences converge?

Tony: For sequences, I believe this does, right? Because as n approaches infinity—

Matt: --so it's convergent is...zero?

Tony: --it should get to zero.

Matt: Right.

Tony: For a series—a sequence, it is. And so for both of these, since this one must reach infinity, and that one as well, they would both also converge.

Matt: To zero, yeah [...]

Tony: ...yeah. That was a good example, too. Now I have a better idea of subsequences as well. So that helped, I think.

[from Iteration 2, Session 4, Theorem 9]

Hence, exploring this example helped reveal a misconception that students had, which in turn enabled the students to develop a better understanding of the term subsequence used in the proof.

Despite the above successes in using this strategy, there were a few instances in which the students did not implement this strategy correctly. Instances of implementations that were not coded as “complete and correct” included being unable to generate an appropriate example or generating an example but being unable to articulate how it related to the theorem. Overall, however, this strategy was easily implemented by students, and the previous excerpts indicated several potential benefits of employing it.

Strategy 3: Describe how you would try to prove the theorem|
Strategy 7. Compare the approach used in the proof to the one you
described.

We treated these as “sister strategies” since we conjectured that the implementation of the first is related to the implementation of the second. Both strategies were coded as being relevant for every proof. For Strategy 3, attempting to prove the theorem, a “complete and correct” implementation was one in which the student considered the proof methods that might be used, what they should assume and prove, and any difficulties they might encounter. The “partially correct” code was assigned when a student indicated an appropriate proof approach but did not investigate this approach in considerable detail. The “superficial or incorrect” code was applied when a student’s response was very short, or the suggested proof approach was inappropriate. Students were not necessarily coded as not implementing a strategy in a “complete and correct” manner if the proof approach they chose differed from the one used in the proof, as long as we determined that it could be adapted into a workable proof. There were 6 “complete and correct” implementations, 8 “partially correct” implementations, and 10 “superficial or incorrect” implementations in all. No student was observed creating a complete proof.

Of the 18 attempts not coded as “complete and correct”, 9 occurred in the first two sessions of Iteration 1. In these first sessions, it became clear to us that students were only

superficially attempting to prove the theorem. For example, in Iteration 1, Session 3, after asking students how they would prove the theorem statement for Proof 3, one student replied by saying “induction” without further elaboration. This behavior was in contrast to the students in Weber (in press), who considered the proof methods that might be used, what they should assume and prove, and any difficulties they might have encountered. We therefore included these goals as sub-steps for this strategy (see Appendix B). Modifications were also made to the instructional intervention by providing more opportunities for students to discuss their implementations of the strategies with one another (which we discuss later in this paper). There was some evidence of improvement in subsequent sessions. As indicated in Table 3, “complete and correct” responses occurred with a higher frequency after making the changes described above. While the implementation was still sometimes problematic, there were several instances where students showed evidence of potential learning, especially when this strategy was coupled with comparing the predicted proof approach to the one used in the proof.

	Complete and correct	Partially correct	Superficial or incorrect
Before modifying Strategy 3	1	4	5
After modifying Strategy 3	5	4	5

Table 3. Frequency of occurrence of different qualities of attempting to prove the theorem before and after modifying the strategy.

For the strategy of comparing proof approaches, an attempt was coded as “superficial or incorrect” when a student made a superficial comparison between the approach used in the proof and the one generated before reading the proof. For instance, some students merely compared the proof technique employed in each case (e.g., whether or not the proof was direct, used induction, contraposition). Conversely, when a student made significant links between the proof

approaches or stated how the proof approach could be used in other mathematical situations, the attempt was coded as “complete and correct”. There were 10 “complete and correct” attempts, no “partially complete” attempts, and 5 “superficial and incorrect” attempts.

Several students mentioned this strategy as being particularly helpful for them, both during the teaching sessions and in the feedback forms. This is illustrated in the following excerpt from Iteration 1, Session 1, in which a student discusses Proof 2, of the claim that the equation $x^3+5x = 3x^2 + \sin(x)$ has only one root. The proof for this theorem relies on establishing that the function $f(x) = x^3+5x - 3x^2 - \sin(x)$ is increasing by showing that $f'(x)$ is positive for all real-valued x , and hence has only one root by Rolle’s Theorem. When students tried to implement the strategy themselves, they attempted to solve the equation by manipulating it algebraically; we do not think this approach could be successful. After reading the proof, some students commented on advantages the calculus approach had over their algebraic approach.

Rahim: “*When I first looked at the [theorem], it seemed to me it would just be arithmetic. I guess just at first glance, it seemed like a pretty basic proof where you could pretty much just do it [another student says ‘plug and chug’] yeah you could just plug and chug, and you could just put steps together. Now I know what to do next time, I guess, to just think a little bit more outside the box. Because even though it may use more difficult concepts, it’ll become a simpler proof altogether because I guess with plugging and chugging with this would be much longer than a nine-step proof, and calculus made it a lot easier, even though you wouldn’t automatically think to use Rolle’s Theorem.*” (italics were our emphasis)

[from Iteration 1, Session 1, Proof 2]

In this excerpt, Rahim claimed to have gained an appreciation for the method used in the proof after comparing it with his own method. He implied that this was helpful because he could now transfer this method into another setting (although confirmatory evidence of Rahim actually doing this would be needed to claim that he fully developed this knowledge). Another student

cited the strategy of trying to prove a theorem before reading its proof as being helpful before seeing the proof:

Matt: Describing how you would prove the theorem is useful, for when you would go to read the proof, you know what you're expecting.

[from Iteration 2, Session 4, post-interview]

We were surprised that the strategy of comparing proof approaches appeared to be effective even when students' attempts at implementing Strategy 3 were coded as not "complete and correct". For instance, in Iteration 1, Session 2, students were reading Theorem 4, that monotone, bounded sequences are convergent. Aneta's initial prediction of the proof approach consisted only of the fact that it would "use epsilons". The group agreed that this was a reasonable approach and did not have anything further to add. This resulted in a coding of "partially correct" for attempting to prove the claim. After reading the proof, however, Rahim realized that the proof had assumed the sequence was increasing, a crucial step for making progress in the proof, which he had not anticipated. The instructor then highlighted this, emphasizing that whenever a statement is encountered about monotonic sequences, it may be useful to assume, without loss of generality, the sequence is monotonic increasing.

In sum, describing how one would prove the claim and comparing this approach to the one used in the proof were generally effective for students, especially after the first of these was modified. However, students had to be informed that they needed to go beyond stating a superficial proof method and actually think about how their strategy would be implemented. Comparing the proof approach to the one that the student generated created the opportunity to learn about when different proof approaches are useful in general. We hypothesize that this is an important way in which this strategy can help students learn proof-based mathematics.

Strategy 4: See if the proof can be broken into independent parts.

Students were encouraged to try to break proofs into independent parts, based on any sub-goals or lemmas established in the proof. Proofs 3, 5, 7, and 10 were considered to not have any sub-goals or lemmas, while all other proofs were judged to be able to be partitioned. An implementation of this strategy was coded as “complete and correct” when the student appropriately partitioned a proof according to the sub-goals achieved in the proof. When a student inappropriately partitioned a proof or failed to realize that a proof could be partitioned, the implementation was coded as “superficial or incorrect”. There were no cases of a “partially correct” implementation of this strategy. In total, there were 9 “complete and correct” instances and 8 “superficial or incorrect” instances recorded.

A common problem with students’ implementation of this strategy was their tendency to partition the proof according to the quality of the prose, rather than the overarching mathematical goals. For instance, in Iteration 1, Session 3, one student suggested that Proof 6 could be partitioned into an “introduction,” “main proof,” and “conclusion”, and another student approved of this implementation. The researchers’ intention was the first part of the proof was to establish that $xy = 2^n - 1$ and the second was to establish that $1 < x < 2^n - 1$ and hence $2^n - 1$ had a factor other than 1 and itself. The researcher was surprised by this response and illustrated the intended partitioning.

To avoid the differing interpretations of this strategy by the students and the teacher-researcher, we attempted to clarify the intended purpose of this strategy by pointing out that the parts had to be mathematically independent from one another. The strategy sheet was modified after the second session of the first iteration, and included ways in which proofs could be broken into parts: by identifying sub-goals, lemmas, or cases, among others (see Appendix B for the

modified strategy sheet). In Iteration 2, we tried having students implement this strategy at different points during proof reading, both before reading the proof line-by-line and after. Despite these modifications and repeated feedback, students' confusion with this strategy persisted throughout the study. One student in Iteration 2 expressed his frustration with this strategy, saying "I don't know if I saw the benefit of breaking the proof up into pieces... Because it seems like... whenever we wanted to, it didn't meet your criteria."

In addition to students' different interpretations of what was meant by a "part", students sometimes had difficulties recognizing the structure of a proof before reading the proof line-by-line. In Iteration 2, Session 4, after reading Theorem 9, that a sequence converges if and only if all of its subsequences converge, Matt and Tony predicted that the proof would be comprised of two parts, one for each direction of the bi-conditional claim. However, after receiving the proof, the students were silent for half a minute, after which the following exchange occurred:

Tony: Do you see that it can be broken into parts somewhere?

Matt: I was looking for it, but I don't see the two distinctions that we made...

Tony: That's exactly what I was looking for.

Matt: ...so he [the proof author] is doing something.

[from Iteration 2, Session 4, Theorem 9]

Matt and Tony appeared confused by the apparent lack of distinct parts in the proof, and the teacher-researcher intervened, asking why they were confused. This prompted Tony to attempt to explain that they were searching for two directions corresponding to the bi-conditional claim. In the course of explaining this, he re-read the first two lines of the proof and realized that this constituted one part, proving the trivial direction that if every subsequence converges, then the original sequence converges since it is itself a subsequence. One explanation for these

students' difficulties is that the students overlooked the first part of the proof due to its short length. It is also apparent that these students initially skimmed the proof, without reading each line carefully. It might be that in certain cases, it can be difficult for students to identify the high-level structure of a proof without reading in detail at a line-by-line level. This hypothesis is supported by comments made by the Iteration 2 students during the post-interview:

Matt: Is it fair to say that we can only break it up by your definition if we've read it line by line, do you know what I mean? When I first look at it...I can't necessarily see if there's a sub-goal without reading the whole proof. So I don't see how I can [partition the proof] first.

Interviewer: Do you think it would be more helpful if it were after reading [the proof] line by line?

Matt: I think so, yeah.

Tony: I mean really it kind of goes, during, really. Because as you're going through--because you can't like read it, and then be like, 'oh wait, was there multiple parts here?' It's really *as you're reading it, line by line, you find parts in it.*

[from Iteration 2, Session 4, post-interview]

To summarize, in both iterations, students were able to correctly partition proofs in over half the recorded instances, which may have helped them to understand the high-level structure of the proof. However, students had difficulties understanding what should constitute a part, even after this strategy was moved earlier among the proof-reading strategies. Having students apply this strategy before reading the proof also may have contributed to students' difficulties identifying the high-level structure of a proof in one instance.

Strategy 5: Identify what proof methods are being used.

This strategy was relevant for all of the proofs used in this study. A “complete and correct” implementation of this strategy consisted of correctly identifying the proving technique

used in the proof (e.g., contradiction, contraposition, direct proof). When a student was unable to identify the proving technique used in the proof, the implementation was coded as “superficial or incorrect”. Students in both iterations were largely successful in implementing this strategy, with 15 “complete and correct” attempts at implementing the strategy and 1 “superficial or incorrect” attempt. There were also 3 instances of students making a “partially correct” implementation of this strategy. This occurred when a student correctly identified the direct proof method used in Proof 4, but overlooked the use of cases. The other instances coded as “partially correct” were of students identifying the correct proof technique, but then were unable to give a correct justification for how they knew these methods were used, as described next.

In general, although students were able to determine the method used in a proof, they were not always able to immediately cite how they knew these methods were used. In one instance in Iteration 1, Session 3, students claimed that Proof 10 was a direct proof, but were initially unable to say why. One student claimed “they head straight into it,” another stated “they don’t use negations, ‘not convergent’”. A third student pointed out that the final assertion of the proof ends with the conclusion of the theorem, a justification that we saw as superior to the previous two. The other students quickly agreed with this reply, suggesting that they saw the value of this justification. Exchanges like these highlight a benefit of the intervention used in this study. The processes used to implement proof-reading strategies are likely to be rapid and tacit. By having students reflect on these processes and make them explicit, others can deepen their own understanding of how to better read proofs. These data suggest that students are generally able to identify the proof method used in the proof, and that learning from peers’ responses may be valuable for understanding how to know these methods are being used.

Strategy 6: Instantiate lines you do not understand with examples.

This strategy encouraged students to use examples to see why assertions were true when it was not immediately obvious why this was so. There were only five recorded cases of students attempting to implement this strategy. This is fewer than the other strategies, likely because the decision of whether to implement this strategy was left up to the students' discretion. That is, students were free to use this strategy if they did not understand a specific line within the proof but were not required to do so. An attempt was coded as "complete and correct" when an example was used to correctly instantiate an assertion and there was evidence that the students better understood this assertion as a result. Three attempts were coded as "complete and correct". There were also two instances that earned a "partially complete" code, which are described below. We first describe two of the "complete and correct" instances here. In the first instance, a student, Tony, in Iteration 2, Session 4 had difficulty understanding the assertion " $n_k \geq k$ for all k since a_{n_k} is a subsequence of a_n " in Proof 9. After spending some time attempting to understand this assertion, he referred back to an example that he had considered before reading the proof, of the sequence $1/p$ and two subsequences, $1/(2n)$ and $1/(2n+1)$:

Tony: How is that true? ... We're looking at n_k , n_k by that definition, is n_1, n_2, n_3 , and so on... Right. And then k is going to be starting at 1, and going on. But your subsequence... is not necessarily 1 and going on... it could be though, which is why it says 'or equal to'. So, yeah, like in our example. This is actually what I thought back to [points to sheet where he wrote the example] Because these were the... I should say like a_1, a_3, a_5 , and so on [labels these on his subsequence $1/(2n+1)$] This one was a_2, a_4, a_6 , right? *So the n th term is going to be—in any sequence it's going to start at 1 or whatever number then just keep going up.* And it doesn't have to go up by ones, it could go up by whatever you define the subsequence as. *But it could never be less than k , so greater than or equal to has to be true.* (italics were our emphasis)

[from Iteration 2, Session 4, Proof 9]

In this excerpt, we see that this strategy prompted Tony to refer to his previously constructed example. This example helped him gain insight into why the statement was true. In another instance, Tony was confused about a claim in the proof of the fact that e is irrational (Iteration 2, Session 3, Proof 7). In this proof, the quantity

$$p((q-1)!) - \sum_{n=0}^q \frac{q!}{n!}$$

is claimed to be an integer. After Tony expresses initial confusion about this claim, the instructor recommended that he try an example. Tony chose $q = 3$, $p = 5$, and, focusing on the summation term of the expression, factored out a “3 factorial” and uttered:

"Well since it's factorial, that would make sense. All these would have to cancel out because we're talking about 3 times 2 times 1 and so it's always going to have something to cancel out the denominator with. So all these fractions will go away, and it has to be an integer. That makes sense. *So I think in this case that specific example actually really helped, because I could actually see the different factors that cancelled out.*" (italics were our emphasis)

[from Iteration 2, Session 3, Proof 7]

In both excerpts, we see Tony developing a better understanding of what each statement was asserting as well as developing a justification for why these assertions are true. We further note that in each previous excerpt, Tony correctly followed the logic of the proof with a particular example, while viewing this example *generically* (Rowland, 2001). That is, the student mirrored the calculations in the proof with the particular example, and moreover, was able to refer to the original proof and compare the work he had done to the assertions made in the proof. This allowed the student to see why a particular statement was true. In contrast, this strategy was not as clearly beneficial when the student did not appear to view the example generically. For example, in the proof that $2^n - 1$ is prime implies that n is prime (Iteration 1,

Session 3, Proof 6), students attempted to see why the quantities $x = 2^a - 1$ and $y = 1 + 2^a + 2^{2a} + \dots + 2^{(b-1)a}$ have the product $2^{ab} - 1$. To do so, Aneta set $a = 2$, and $b = 3$, and immediately computed that $x = 3$, and $y = 21$ so that $xy = 63$. However, in doing this calculation, Aneta skipped the intermediate steps of the calculation followed in the proof. Had she done so, she would have been able to check that the claim was true, and might have gained a better understanding of the initial computational steps in this proof. However, we believe this student missed an opportunity to see *why* the claim was true through cancellation of the various terms.

From these limited cases, we observe that asking students to apply this strategy might not always benefit students. It appears critical that students view their example verifications generically. One way to address this problem might be for the teacher-researcher to highlight when a student fails to use an example generically, and demonstrate what would be a generic use of example in this case. For instance, in the last excerpt, with $a = 2$ and $b = 3$, the teacher-researcher could have kept x and y in the fully expanded form, and showed exactly how the intermediate terms cancel out to arrive at $2^6 - 1$. We also noted that students did not often spontaneously use this strategy, a finding consistent with the survey data of Weber and Mejia-Ramos (2013a) and also with the behavior of the 28 students observed in the qualitative study of Weber (2009). Making students aware of the strategy and having the instructor model the strategy was not sufficient motivation for the students in this study to implement this strategy. In subsequent iterations of implementing this instruction, giving students cases where they were required to use this strategy for steps in proofs that were known to be problematic might provide students with the opportunity to see the benefits of applying this strategy, making it more likely for them to use the strategy in the future.

Other observations about the strategies and instruction

A complaint that occurred repeatedly throughout the study was that students sometimes realized that certain lines in a proof were true, but felt unsatisfied because they felt they would not have thought to make such an inference if they were writing the proof. Leron (1983) referred to this as proofs “which pull a rabbit out of a hat” in which some new quantity is “magically” produced to satisfy some desired properties. This was especially apparent with in Proof 6, of the fact that $2^n - 1$ is prime implies that n is prime. This statement is proved using contraposition and assumes that $n = ab$ with a, b , both greater than 1. The proof then defines two numbers, $x = 2^a - 1$ and $y = 1 + 2^a + 2^{2a} + \dots + 2^{(b-1)a}$, multiplies these numbers, and checks that their product equals $2^n - 1$. Several students complained that although they were convinced the calculation was correct, they did not see how the author of the proof thought to define these quantities:

Rahim: I see how $[xy = 2^n - 1]$ is a true statement, but I still don't see how you derive x and y .
[Iteration 1, Session 3, Proof 6]

Matt: The whole idea of bringing up this y , it does seem to come out of nowhere. It's one of those things that's like they make it so that it works...I get it logically, I understand what he's trying to do, but how they came up with it...*I would like to see his scratch work, his idea.* (italics were our emphasis) [Iteration 2, Session 3, Proof 6]

In a post-interview with the two students from Iteration 2, they elaborated on this aspect of proofs as typical to their experience with proof-based mathematics:

Tony: Whenever you're reading proofs, and stuff like that, whenever they just say that line, like let *this* be something really specific, and it's just like, OK... (italics were Tony's emphasis)

Matt: ...all right. [laughing]...when they present the non-obvious so obviously, it's just like [laughing, shaking his head and shrugging] it makes you feel *less* about yourself mathematically.

[from Iteration 2, Session 4, post-interview]

In the previous excerpts, students expressed confusion and frustration by what they perceived to be a mathematical sleight of hand. This phenomenon is related to our goals for how students can understand proofs, from the proof comprehension model of Mejia-Ramos et al. Specifically, it represents an inability to see how the method of the proof could transfer to another context. This proof comprehension model highlighted applying a holistic and presumably transparent method to another context, but this model does not include applying specific heuristics in other contexts. Our strategies did not seem to help students overcome their perceived shortcomings in this regard. This may be an area where an instructor may need to help students cope by explaining the reasoning used to generate the difficult step.

As described in the Procedure, our instructional method was based on Palincsar and Brown's (1984) reciprocal teaching studies. An integral part of this method, as we interpreted it, is for the instructor to provide scaffolding and regular feedback, particularly as students are initially learning the strategies. However, this appeared to have a negative effect on student engagement. As the instructor assumed the responsibility for evaluating students' responses, students took a passive role. In fact, in the first two sessions of Iteration 1, there was not a single instance of students providing any feedback other than passive confirmation to a students' use of these strategies. Instead, students would simply repeat what others had already said or say nothing meaningful at all (e.g., "sounds good"). We therefore conjectured that students had assumed that the instructor would continue to evaluate their use of strategies, and so did not feel the need to critically evaluate their peers' responses. We decided to adopt an instructional design principle observed from a different teaching study that we conducted (Weber et al., 2010): we hypothesized that this situation could be ameliorated by having students form smaller groups (two to three students each) in which they would be responsible for discussing how they would

implement each of the strategies. Students were told that they needed to reach agreement in their groups before presenting their ideas to the class.

After these changes to the protocol were made, there were instances of students evaluating one another's strategy use. For instance, in Iteration 1, Session 3, in thinking how to prove that $2^n - 1$ is prime implies that n is prime in Proof 6, Aneta suggested using the Fundamental Theorem of Arithmetic. Another student suggested that induction could be used instead. Aneta argued against this saying, "I don't think induction will work because if you have n , then $n+1$ wouldn't be prime. Unless it's two."

Discussion: Lessons learned from these two iterations

Lessons learned about the proof reading strategies

In this section, we discuss each of the strategies and what we learned from the two iterations of the study.

Strategy 1: We found that asking students to "know the definition of the terms of the theorem" was not helpful for students at the start of Iteration 1; the students simply would identify the terms but not state their definitions, perhaps because they felt they already knew the definitions. Asking students to specifically *state* the definition and to look up definitions they may have forgotten led students to use the strategy more productively. In Iteration 2, we saw that this not only helped students understand some proofs, but it also called their attention about an important epistemological aspect of proof: proofs about concepts in mathematics should be based on the definitions of those concepts. In Iteration 2, we discovered that one limitation of this strategy occurs in the case where there are multiple common characterizations of the concept, as this strategy would not prescribe to students which characterization that they should consider.

Strategy 2: The data in both iterations support the claim that by instantiating a theorem statement with a particular example can help students understand the proofs that they read. As predicted, this helped students understand why the theorem statement was true and how the proof might proceed. An unanticipated benefit that we observed in Iteration 2 was that when a student instantiated a theorem with an example, a classmate, the teacher-researcher, or the student herself sometimes realized that they had a misconception about what the theorem was asserting. This realization provided the opportunity to take actions to remedy this situation.

Strategy 3: In Iteration 1, we found that simply asking students how to prove the theorem led to superficial responses, such as “use epsilons”, that would not aid their understanding of the proof. We presented students with more detailed instructions, such as stating how the idea would be carried out and what difficulties may arise. These instructions led students to avoid this shallow implementation of the strategy. When students implemented the strategy successfully, some cited that it was beneficial to them because it helped them know what to expect when reading the proof.

Strategy 4: We did not find that asking students to break the proof into independent parts was helpful for students to build an understanding of the proof. Even after being precise about how this strategy could be implemented, students still complained that they had difficulty implementing it and that it was not helpful. When students were asked to implement this strategy *before* reading it at a line-by-line level, they found the task difficult, saying it was tough to know when one section of a proof ended and another began without reading it closely first. It seemed as if students could only implement this strategy *as a result* of understanding the proof, making it problematic to use it as a strategy *as a means* to understand the proof.

We were surprised by this finding, given that we observed successful students using this strategy effectively in our earlier research (Weber, in press). Further, mathematicians claim both to use this strategy (Mejia-Ramos & Weber, 2014; Weber & Mejia-Ramos, 2011) and to desire that their students use this strategy as well (Weber, in press). One possibility is that being able to break the proof into independent parts represents a deep understanding of the proof; that is, this is more an indicator of understanding rather than the means to achieve it. Inglis and Alcock (2012, 2013) contended that mathematicians may allege to skim proofs for comprehension, but their actual behavior when reading a proof is not consistent with this claim. (Weber and Mejia-Ramos (2013b) objected to the interpretation of their data). The data in this study are consistent with the claim of Inglis and Alcock.

Another possibility is that our atypical proof presentation, where the proof did not appear in paragraph form but as a sequence numbered lines, made the use of this strategy difficult. Our proofs did not contain any advanced organizers, paragraph breaks, or indentations. Konior (1993) suggested that such “delimiters” in proofs may be helpful for readers’ understanding.

Strategy 5: Students were generally successful at identifying the proof methods used, although in Iteration 1, students sometimes did not articulate how they made this identification. We found that helping students know that the proof method was determined by the assumptions and conclusions of the argument led to more articulate descriptions of how the strategy was implemented.

Strategy 6: We encouraged students to instantiate a line of a proof that they did not understand with a specific example. This was an optional strategy and we found that students usually did not use the strategy. It was only used five times across the two iterations of the study.

From studying the implementation of this strategy in both iterations, when the strategy was implemented, it was successful for students if they treated the example generically but not when they focused on calculation-based aspect of the example. This is consistent with a broader point that examples may have more value for mathematicians than for students since mathematicians use the examples in more sophisticated and productive ways (e.g., Iannone et al, 2011).

Strategy 7: The data from both iterations confirmed that when students compared the proof that they read to their own approach proving the theorem, they felt that this helped them see situations that it might be applicable elsewhere. We were surprised to find that students could reap benefits from the strategy even if they failed to articulate how they would try to prove the theorem before reading the proof.

Themes across the strategies

Specific prescriptive guidance helped students implement them more effectively. This finding is consistent with the teaching of heuristics in mathematical problem solving. In summarizing the research literature at the time, Schoenfeld (1985) observed that describing and modeling heuristics for students is not sufficient for students to use these heuristics effectively. One needed to be clearer at a prescriptive level as to how the heuristics were implemented. We found that when we offered this prescriptive guidance for several of the strategies, such as trying to prove a theorem before reading its proof and stating the definition of the terms in the theorem statement, students' rate of success in implementing the strategy successfully improved.

These strategies were beneficial to students. In the previous section, we illustrated how, with the exception of Strategy 4, there were instances where implementing the strategies appeared to help students improve their understanding of the proofs that they read. For instance,

illustrating a theorem with an example before reading its proof helped students understand what the theorem was asserting and how the proof might proceed. Comparing the methods of a proof with one's own approach helped students expand their arsenal of proving techniques. When students stated the definitions of terms before reading its proof, some reported having a greater appreciation of the role that definitions play in the proving process and this provided the opportunity to identify gaps in students' knowledge. At this stage, this evidence is only suggestive as students were not systematically given post-tests on their understanding of proof. Further research can corroborate these findings by demonstrating that students who are taught these strategies subsequently perform better on proof comprehension tests.

There were impediments to proof comprehension that could not be addressed by these strategies. We designed these strategies so that students could increase their comprehension of a proof that they read independently. Our goal was to provide students with the means to autonomously mine the insights out of proofs. There were two aspects of the proof that students were not able to understand using our strategies—choosing the right definition to base a proof upon when multiple definitions were available and understanding the rationale for unmotivated steps in the proof. We wonder if there was anything students could have done in these situations. It might be the case that the presenter of the proof might have to explain these aspects to the student-readers.

Asking students to implement and discuss the strategy in small groups led to more student involvement in the evaluation of the strategy. At the start of Iteration 1, although we asked students to evaluate others' implementation of the strategies, the students tended to be passive as they did so. They simply agreed that the implementation was fine. To rectify this, we had students work in small groups, requiring that they reach a consensus before sharing their

implementation in the class. The responsibility of forming a consensus led students to be much more attentive and evaluative of their classmates' work.

Limitations of our study and future research

In this study, we asked students to apply our proof reading strategies to proofs from calculus, real analysis, and number theory. Our primary reason for this decision was we felt students would have an adequate knowledge base to understand these proofs; the number theory proofs only relied on basic number theoretic facts that was covered in the transition-to-proof course that these students had completed and the calculus and real analysis concepts were covered in the calculus sequence and real analysis course these students had completed. It is possible that the proof reading strategies that we described worked well for proofs in *these specific domains*, but might not be beneficial in other domains such as topology, set theory, or group theory. For instance, the example generation strategies might be difficult to apply for a beginning group theory student who may not yet have a rich collection of groups that he or she feels comfortable exploring.

We chose our proofs to pertain to content with which students are familiar. Of course, when students read proofs in their advanced mathematics courses, this will often not be the case. Students will have to simultaneously contend with understanding a proof as well as understanding the concepts within that proof. Indeed, a professor may hope that students come to understand a concept by reading a proof about that concept. As our study deliberately avoided such proofs, we cannot claim that our strategies are helpful in these situations. It is possible that our strategies might not be helpful or a different set of strategies is needed.

Our evidence for a strategy's effectiveness consisted of qualitative illustrations of students gaining an insight by applying a strategy and their personal reflections on using a

strategy. However, we acknowledge this is merely suggestive evidence that learning has occurred. For instance, it is possible that the gains that the participant made were fleeting or that they would not be able to leverage these gains to answer questions on a proof comprehension test. Further studies, where an actual post-test was given, would be needed to address this limitation in the study.

Relationship to other literature

We have discussed the results primarily in relation to the papers we have written regarding the strategies. There have also been other studies examining the strategies that students use when attempting to interact with mathematical texts. In the remainder of the paper, we compare our strategies and results with these other studies.

Shepherd, Selden and Selden (2012) asked 11 students to read excerpts from pre-calculus and calculus textbooks, and counted students' use of fifteen "constructively responsive reading" strategies. These strategies were adopted from the reading comprehension literature and most were more general in nature than ours, including "use strategies to remember text," and "relate text content to prior knowledge". Interestingly, Shepherd et al.'s study suggested that students were aware of, but not bothered by, their confusions or inadequate prior knowledge when reading the text. This is consistent with our findings that suggested that students were aware of their imperfect understanding of concepts (such as the mathematical constant e) but avoided addressing them because they did not think it was important to do so. On the other hand, several students in our study also voiced concerns when they were able to validate a series of steps, but unable to extract a deeper understanding of the method behind these steps. The dissatisfaction that these students expressed about their own understanding suggests a level of sophistication and curiosity in thinking that seems runs contrary to Shepherd et al.'s students. One explanation

might be that this level of sophistication and curiosity was absent in Shepherd et al.'s study because students were first-year mathematics students in pre-calculus and calculus. Perhaps experience with higher-level mathematics, as well as students' motivation and goals for continuing to enroll in mathematics courses, are relevant factors in students' curiosity and self-reflection.

Several of our strategies, such as identifying proof frameworks, were specific to proof reading and all of our strategies were geared toward reading *mathematical arguments*. In this sense, it is not surprising that our strategies were more specifically geared toward mathematics than the general comprehension strategies explored by Shepherd et al.

Our strategies can also be compared with those suggested in the literature for proof validation. We first note that it would not be surprising if the strategies for validation differed from those of comprehension. As Rav (1999) noted, how one reads a proof depends on what one hopes to get out of it; we have observed that mathematicians use different strategies for comprehension and refereeing (Mejia-Ramos & Weber, 2014). Inferring warrants has been suggested to be an important part of validating proofs (Weber & Alcock, 2005). While we did not directly instruct students to do this, students were told to instantiate steps of the proof they did not understand with specific examples (Strategy 6). As discussed above, we believe this can be helpful for students to understand how steps in the proof follow from one another, but only when those steps are followed through completely with the example, as far as the reader finds necessary. Selden and Selden (1995) recommended that students attend to the proof framework of a proof while attempting to validate proofs. There were two strategies used in our study that we believe could be beneficial to students' comprehensions of proof frameworks: attempting to prove the theorem (Strategy 3) and identifying the proof technique (Strategy 5). In these ways,

our strategies might be suitable for improving students' ability to validate proofs. Furthermore, Inglis and Alcock (2012) and Selden and Selden (2003) argued that students focus too much on local aspects of the proof, and failed to attend to global aspects when validating proofs. Some of our strategies, such as partitioning the proof (Strategy 4) and comparing the proof approach to the reader's own approach (Strategy 7) could be helpful for encouraging students to understand the global structure of the proof.

Acknowledgments

This material is based upon the work supported by the National Science Foundation under grants DUE-1245626 and DRL-0643734. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

- Alcock, L. & Weber, K. (2005). Proof validation in real analysis: Inferring and evaluating warrants. *Journal of Mathematical Behavior*, 24(2), 125-134.
- Bills, L., & Tall, D. (1998). Operable definitions in advanced mathematics: The case of the least upper bound. In A. Olivier & K. Newstead (Eds.) *Proceedings of the Conference for the International Group of the Psychology of Mathematics Education*, 2, (pp. 104-111). Stellenbosch, South Africa: PME.
- Brown, J., Collins, A., & Newman, S. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. Resnick's (Ed.) *Knowing, learning, and instruction: Essays in honor of Robert Glaser*. (pp. 32-42). Mahwah, NJ: Lawrence Erlbaum Associates.
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32(1), 9-13.
- Conradie, J., & Frith, J. (2000). Comprehension tests in mathematics. *Educational Studies in Mathematics*, 42, 225-235.
- Cowen, C. (1991). Teaching and testing mathematics reading. *American Mathematical Monthly*, 98(1), 50-53.
- diSessa, A. A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *The Journal of the Learning Sciences*, 13(1), 77-103.
- Doolittle, P. E., Hicks, D., Triplett, C. F., Nichols, W. D., & Young, C. A. (2006). Reciprocal Teaching for reading comprehension in higher education: A strategy for fostering the deeper understanding of texts. *International Journal of Teaching and Learning in Higher Education*, 17(2), 106-118.
- Harel, G., Selden, A., & Selden, J. (2006). Advanced mathematical thinking. In A. Gutierrez & P. Boero (Eds.) *Handbook of research on the psychology of mathematics education: past, present and future* (pp. 147-172). Rotterdam: Sense Publishers.

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

Iannone, P., Inglis, M., Mejia-Ramos, J.P., Simpson, A., & Weber, K. (2011). Does generating examples aid proof production? *Educational Studies in Mathematics*, 77, 1-14.

Inglis, M. & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, 43, 358-390.

Inglis, M. & Alcock, L. (2013). Skimming: A response to Weber and Mejia-Ramos. *Journal for Research in Mathematics Education*, 44, 471-474.

Ko, Y. Y., & Knuth, E. J. (2013). Validating proofs and counterexamples across content domains: Practices of importance for mathematics majors. *The Journal of Mathematical Behavior*, 32(1), 20-35.

Konior, J. (1993). Research into the construction of mathematical texts. *Educational Studies in Mathematics*, 24(3), 251-256.

Lai, Y. & Weber, K. (2014). Factors mathematicians profess to consider when presenting pedagogical proofs. *Educational Studies in Mathematics*, 85, 93-108.

Lai, Y., Weber, K., & Mejía-Ramos, J. P. (2012). Mathematicians' perspectives on features of a good pedagogical proof. *Cognition and Instruction*, 30(2), 146-169.

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

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

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

Mejia-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, 79(1), 3-18.

Mejia-Ramos, J.P. & Weber, K. (2014). How and why mathematicians read proofs: Further evidence from a survey study. *Educational Studies in Mathematics*, 85, 161-173.

Mills, M. (2011). Mathematicians' pedagogical thoughts and practices in proof presentation. In *Proceedings of the 14th conference for research in undergraduate mathematics education* (Vol. 2, pp. 283-297).

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics*, 27(3), 249-266.

Palincsar, A.S., & Brown, A.L. (1984). Reciprocal teaching of comprehension fostering and monitoring activities. *Cognition and Instruction*, 1, 117-175.

Rav, Y. (1999). Why do we prove theorems? *Philosophia Mathematica*, 7(3), 5-41.

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

Schoenfeld, A. (1985). *Mathematical problem solving*. Orlando, FL: Academic Press.

Segal, J. (2000). Learning about mathematical proof: Conviction and validity. *Journal of Mathematical Behavior*, 18, 191-210.

Selden, A. (2012). Transitions and proof and proving at tertiary level. In G. Hanna & M. de Villiers (Eds.) *Proof and Proving in Mathematics Education* (pp. 391-420). Springer Netherlands.

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, A., & Selden, J. (2013). The Genre of Proof. (No. 2013-1). Cookeville, TN: Tennessee Technological University.

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

Shepherd, M. D., Selden, A., & Selden, J. (2012). University students' reading of their first-year mathematics textbooks. *Mathematical Thinking and Learning*, 14(3), 226-256.

Vinner, S. (1983). Concept definition, concept image and the notion of function. *International Journal of Mathematical Education in Science and Technology*, 14(3), 293-305.

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

Weber, K. (2009). Mathematics majors evaluation of mathematical arguments and their conception of proof. In *Proceedings of the 12th Conference for Research in Undergraduate Mathematics Education*. Available for download at: <http://rume.org/crume2009/proceedings.html>

Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning*, 12(4), 306-336.

Weber, K. (2012). Mathematicians' perspectives on their pedagogical practice with respect to proof. *International Journal of Mathematics Education in Science and Technology*, 43, 463-482.

Weber, K. (in press). Effective proof reading strategies to foster comprehension of mathematical proofs. To appear in *International Journal for Research in Undergraduate Mathematics Education*.

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

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

Weber, K. & Mejia-Ramos, J.P. (2013a). Effective but underused strategies for proof comprehension. In M. Martinez & A. Superfine (Eds.) *Proceedings of the 35th Annual Meeting of the North American Chapter of the Psychology of Mathematics Education*. (pp. 260-267). Chicago, IL.

Weber, K. & Mejia-Ramos, J.P. (2013b). On mathematicians' proof skimming. *Journal for Research in Mathematics Education*, 44, 464-471.

Weber, K., & Mejia-Ramos, J. P. (2014). Mathematics majors' beliefs about proof reading. *International Journal of Mathematical Education in Science and Technology*, 45(1), 89-103.

Weber, K., Radu, I., Mueller, M., Powell, A. & Maher, C. (2010). Expanding participation in problem solving in a diverse middle school mathematics classroom. *Mathematics Education Research Journal*, 22, 91-118.

Weber, K. & Samkoff, A. (2011). Effective strategies that undergraduates use to read and comprehend proofs. In *Proceedings of the 14th Annual Conference on Research in Undergraduate Mathematics Education*. Available for download from: <http://sigmaa.maa.org/rume/Site/Proceedings.html> Last downloaded 7 May 2012.

Appendix A. Original proof reading strategies

What do we do before we read the proof?

- Know the definition of the terms of the theorem
- Work through some examples to make sure we understand the statement
- Try to rephrase the theorem in our own words
- How would you try to prove the theorem?

What do we do while we are reading the proof?

- Does the proof have a framework that we can recognize (such as direct proof, proof by cases, proof by contradiction, proof by contrapositive, mathematical induction, constructive proof, uniqueness proof)?
- Do we understand each step in the proof? Could we illustrate the step with an example?
- How does each step follow as a consequence of previous steps?

What do we do after we read a proof?

- Do we see how each of the hypotheses was needed and used in the proof?
- How can the proof be broken into parts?
- How did the approach shown in the proof differ from the way you would have chosen?
- Can we explain the proof to someone else?

Appendix B. Modified proof reading strategies.

Before reading the proof:

1. Make sure you understand the definitions of all the terms
 - a. What mathematical concepts are involved in the theorem statement?
 - b. What is the definition of each concept?
 - c. If you do not know a definition, use a resource to find it. In this study, ask your classmates or the instructor. In other settings, you can consult a teacher, use your textbook, or search the internet.
2. Illustrate the theorem with an example
 - a. Choose an appropriate example. The example should satisfy the hypothesis, not be too complicated, but not be too trivial.
 - b. Check to see if the theorem is true for this example.
 - c. If not, consult a resource (a classmate, a teacher) to see if you really understand what the theorem is saying.
3. Rephrase the theorem in your own words
4. Describe how you would try to prove the theorem
 - a. Choose a proof approach that you think might work
 - b. Describe what you think it might look like. What would you assume and what would you try to show?
 - c. What difficulties would you expect if you tried your approach?

While reading the proof:

1. See if the proof can be broken into parts.
 - a. See if the proof has lemmas or sub-goals. If it does, treat the proofs of the lemmas or sub-goals as separate proofs.
 - b. See if the proof is a proof by cases. If it is, you can treat the proof of each case as a separate proof. (But be sure the proof by cases is done correctly, see step 2).

- c. See if the proof requires you to prove two (or more) separate things. An if-and-only-if statement requires proving things in both directions. Proving s is a least upper bound often involves proving s is an upper bound and proving that there are no upper bounds smaller than s .
 - d. *After reading the proof:* What did each section of the proof accomplish?
 2. Identify what proof methods—direct proof, proof by contradiction, proof by contraposition, proof by induction, or proof by cases—is being used.
 - a. What assumptions does the proof (or sub-proof) start with?
 - b. How do these assumptions relate to a proof method? (If the proof begins by assuming the conclusion is false, you are using a proof by contradiction or a proof by contraposition).
 - c. Check the conclusion of the proof (or sub-proof). Use these to confirm the proof method.
 3. Do you understand each step in the proof?
 - a. Are there any terms in the proof that you do not understand? If so, consult a source (classmate, textbook, teacher) to find out what these terms mean.
 - b. How does the new statement follow from previous statements?
 - c. If you are confused by a statement, try to see why it is true for a particular example.

After reading the proof:

1. How did the approach you chose differ from the proof that you read?
 - a. Were there any differences that you observed?
 - b. What benefit did you think the author's approach to their proof had that yours did not? Were there any difficulties the author avoided that your approach did not?
 - c. For what other types of proofs might the method of this proof be applicable?

Appendix C. Proofs used in Study 2

Proof 1

Theorem: A positive integer, N , is a square if and only if the number of divisors of N is an odd number.

Proof:

1. Suppose that there are an odd number of divisors of N .
2. For each divisor, d , of N , N/d is also a divisor of N .
3. Thus, there are an even number of divisors of N unless there is some d with $d = N/d$.
4. In this case $d^2 = N$ so N is a perfect square.
5. Now suppose N is a perfect square: $N = r^2$.
6. If d is a divisor of N , then N/d is also a divisor of N .
7. Moreover, if $d < r$, then $N/d > r$.
8. Also, if $d > r$, then $N/d < r$.
9. So the divisors of N are $\{1 < d_1 < d_2 < \dots < r < \dots < N/d_2 < N/d_1 < N\}$
10. Hence, N has an odd number of divisors.

Proof 2

Theorem: The only solution to the equation $x^3 + 5x = 3x^2 + \sin x$ is $x = 0$.

Proof:

1. Check that $x = 0$ is a solution to the equation: $(0)^3 + 5(0) = 3(0)^2 + \sin(0)$.
2. We need to show that there are no other solutions.
3. Let $f(x) = x^3 - 3x^2 + 5x - \sin x$.
4. Roots of $f(x) = 0$ precisely correspond to solutions of $x^3 + 5x = 3x^2 + \sin x$.
5. Suppose $f(x) = 0$ has a nonzero root; that is $s \neq 0$ and $f(s) = 0$.
6. Then

$$f'(x) = 3x^2 - 6x + 5 - \cos x = 3(x^2 - 2x + 1) + 2 - \cos x = 3(x - 1)^2 + 2 - \cos x.$$

7. Since $3(x-1)^2 \geq 0$ and $2 - \cos x > 0$ for all real numbers x , $f'(x) > 0$ for all real numbers x .
8. Since $f(0) = f(s) = 0$ and, $s \neq 0$ by Rolle's theorem, there exists c between 0 and s such that $f'(c) = 0$.
9. However, this is a contradiction because $f'(x) > 0$ for all x .

Proof 3

Theorem: There is a real number whose fourth power is exactly one larger than itself.

Proof:

1. Let $f(x) = x^4 - x - 1$.
2. $f(1) = -1$ and $f(2) = 13$.
3. By the Intermediate Value Theorem, there must be some number c such that $1 < c < 2$ and $f(c) = 0$.
4. c has the desired property.

Proof 4

Theorem: A bounded monotone sequence is convergent.

Proof:

1. Let $\{a_n\}_{n=1}^{\infty}$ be a bounded monotone sequence.
2. Assume the sequence is increasing.
3. Let $s = \sup\{a_n : n = 1, 2, \dots\}$
4. It will be shown that $\{a_n\}_{n=1}^{\infty}$ converges to s .
5. Let $\varepsilon > 0$. Since s is the least upper bound of $\{a_n\}_{n=1}^{\infty}$, $s - \varepsilon$ is not an upper bound.
6. Hence there is an n_0 such that $s - \varepsilon < a_{n_0}$.
7. For all $n > n_0$, $s - \varepsilon < a_{n_0} \leq a_n \leq s$.
8. Hence $\{a_n\}_{n=1}^{\infty}$ converges to s .

9. For the case in which $\{a_n\}_{n=1}^{\infty}$ is decreasing, let $s = \inf\{a_n: n = 1, 2, \dots\}$
10. The proof in this case is similar to the above.

Proof 5

Theorem: If n is an integer and $n \geq 2$, then $n^3 - n$ is divisible by 6.

Proof:

1. Let n be an integer such that $n \geq 2$.
2. Then $n^3 - n = n(n^2 - 1) = n(n + 1)(n - 1)$
3. $n^3 - n$ is therefore the product of three consecutive positive integers.
4. Therefore, 3 must divide n , $n + 1$, or $n - 1$.
5. Also, 2 must divide n , $n + 1$, or $n - 1$.
6. Hence 2 divides a factor of $n^3 - n$ and 3 divides a factor of $n^3 - n$.
7. Therefore, 6 divides $n^3 - n$.

Proof 6

Theorem: If $n \geq 2$ and $2^n - 1$ is a prime number, then n is prime.

Proof:

1. We prove the contrapositive.
2. Suppose $n \geq 2$ is not prime.
3. Then $n = ab$ where a and b are both greater than 1.
4. Let $x = 2^a - 1$ and $y = (1 + 2^a + 2^{2a} + \dots + 2^{(b-1)a})$.
5. Then $xy = (2^a - 1)(1 + 2^a + 2^{2a} + \dots + 2^{(b-1)a})$
6. $= 2^a(1 + 2^a + 2^{2a} + \dots + 2^{(b-1)a}) - (1 + 2^a + 2^{2a} + \dots + 2^{(b-1)a})$
7. $= (2^a + 2^{2a} + 2^{3a} + \dots + 2^{ba}) - (1 + 2^a + 2^{2a} + \dots + 2^{(b-1)a}) = 2^{ab} - 1 = 2^n - 1$.
8. Since $a > 1$, $x = 2^a - 1 > 1$.
9. Since $b > 1$, $a < n$.

10. Hence $x = 2^a - 1 < 2^n - 1$.

11. Therefore $2^n - 1$ can be written as the product of two positive integers, both greater than one.

Proof 7

Theorem: e is irrational.

Proof:

1. Suppose that e is a rational number: $e = \frac{p}{q}$.
2. We know that $e = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \dots > 2$.
3. We also know that $e = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \dots < 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots = 3$
4. Since e is not a whole number, $q > 1$.
5. Then $\frac{p}{q} = e = \sum_{n=0}^{\infty} \frac{1}{n!} = \sum_{n=0}^q \frac{1}{n!} + \sum_{n=q+1}^{\infty} \frac{1}{n!}$
6. Rearranging and multiplying by $q!$, we obtain:

$$p(q-1)! - \sum_{n=0}^q \frac{q!}{n!} = \sum_{n=q+1}^{\infty} \frac{q!}{n!}$$

7. The left-hand side of the equation above is an integer.
8. For the right hand side of the equation,

$$\sum_{n=q+1}^{\infty} \frac{q!}{n!} = \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \frac{1}{(q+1)(q+2)(q+3)} + \dots$$

$$< \frac{1}{q+1} + \frac{1}{(q+1)^2} + \frac{1}{(q+1)^3} + \dots = \frac{1/(q+1)}{1 - 1/(q+1)} = \frac{1}{q} < 1$$
9. Thus the right-hand side of the equation is a positive integer less than one, a contradiction.
10. Hence, e is irrational.

Proof 8

Theorem: Consider the sequence defined as follows:

$$s_1 = \sqrt{2}, \quad s_{n+1} = \sqrt{2 + s_n}, \quad \text{for } n = 1, 2, 3, \dots$$

Show that the sequence $\{s_n\}_{n=1}^{\infty}$ is convergent.

Proof:

1. We will show that $\{s_n\}$ is a bounded, increasing sequence. Hence $\{s_n\}$ is convergent.
2. We will use proof by induction to show that $\{s_n\}$ is an increasing sequence.

3. $s_2 = \sqrt{2 + \sqrt{2}} > \sqrt{2}$ so $s_1 < s_2$
4. $s_{n+2} = \sqrt{2 + s_{n+1}} > \sqrt{2 + s_n} = s_{n+1}$ so $\{s_n\}$ is an increasing sequence.
5. We will use proof by induction to show that $s_n < 2$.
6. $s_1 < 2$.
7. $s_{n+1} = \sqrt{2 + s_n} < \sqrt{2 + 2} = 2$ hence $\{s_n\}$ is bounded.
8. Since a bounded, increasing sequence is convergent, $\{s_n\}$ is convergent.

Proof 9

Theorem: A sequence $\{a_n\}_{n=1}^{\infty}$ converges iff each of its subsequences converges.

Proof:

1. Suppose that every subsequence of $\{a_n\}_{n=1}^{\infty}$ converges.
2. Since $\{a_n\}_{n=1}^{\infty}$ is a subsequence of itself, $\{a_n\}_{n=1}^{\infty}$ converges.
3. Suppose that $\{a_n\}_{n=1}^{\infty}$ converges to L .
4. Let $\{a_{n_k}\}_{k=1}^{\infty}$ be a subsequence.
5. Let $\varepsilon > 0$.
6. Since $\{a_n\}_{n=1}^{\infty}$ converges to L , there exists a positive integer N such that, for $n \geq N$, $|a_n - L| < \varepsilon$.
7. Since $\{a_{n_k}\}_{k=1}^{\infty}$ is a subsequence of the given sequence, $n_k \geq k$ for all positive integers k .
8. Thus if $k \geq N$, $n_k \geq N$.
9. Hence, $|a_{n_k} - L| < \varepsilon$.
10. So $\{a_{n_k}\}_{k=1}^{\infty}$ is convergent.

Proof 10

Theorem: If $\{a_n\}_{n=1}^{\infty}$ converges to 0 and $\{b_n\}_{n=1}^{\infty}$ is bounded, then $\{a_n b_n\}_{n=1}^{\infty}$ converges to 0.

Proof:

1. Let M be a positive number such that $|b_n| \leq M$ for all n .
2. Choose $\varepsilon > 0$.
3. Then let $\varepsilon' = \frac{\varepsilon}{M} > 0$.
4. There is a positive integer N such that $n \geq N$ implies $|a_n| = |a_n - 0| < \varepsilon'$.
5. Then $|a_n b_n - 0| = |a_n b_n| = |a_n| |b_n| \leq |a_n| M < \varepsilon' M = \varepsilon$.
6. Thus $\{a_n b_n\}_{n=1}^{\infty}$ converges to 0.

Proof 11

Definition: A Primitive Pythagorean triple is a triple of positive integers (a, b, c) such that:

1. $a^2 + b^2 = c^2$ and
2. a , b , and c are co-prime. (That is, there is no common factor of a , b , and c other than 1.)

Theorem: If (a, b, c) is a primitive Pythagorean triple, then c is odd.

Lemma 1: For any natural number x , $x^2 \equiv 1 \pmod{4}$ if x is odd and $x^2 \equiv 0 \pmod{4}$ if x is even.

Proof of Lemma 1:

1. Suppose x is odd. Then $x = 2k + 1$ for some integer k . Then

$$x^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 4(k^2 + k) + 1$$
2. So x^2 leaves a remainder of 1 when divided by 4.
3. Suppose x is even. Then $x = 2k$ for some integer k . Then

$$x^2 = (2k)^2 = 4k^2$$
4. So x^2 is divisible by 4 and $x^2 \equiv 0 \pmod{4}$.

Lemma 2: For any natural number x , If x^2 is even, then x is even.

Proof of Lemma 2:

1. We prove the contrapositive.
2. Suppose that x is odd.
3. Then by Lemma 1, x^2 leaves remainder 1 mod 4.
4. Hence x^2 is odd, as desired.

Proof of Theorem:

1. Let (a,b,c) be a Primitive Pythagorean triple.
2. There are three cases. Either a and b are both even, or a and b are both odd, or a is even and b is odd (or vice versa).
3. Case 1: Assume a and b are both even. Since a and b are even, a^2 and b^2 are even. Hence $a^2 + b^2 = c^2$ is even.
4. By Lemma 2, since c^2 is even, c must be even.
5. Hence a , b , and c are all even and have a common factor of 2.
6. This contradicts the fact that (a, b, c) is a Primitive Pythagorean triple.
7. Thus a and b cannot both be even.
8. Case 2: Assume a and b are both odd. By Lemma 1, since a and b are odd, $a^2 \equiv 1 \pmod{4}$ and $b^2 \equiv 1 \pmod{4}$. So $a^2 + b^2 = c^2 \equiv 2 \pmod{4}$.
9. By Lemma 1, we know that c^2 cannot be congruent to 2 (mod 4).
10. Thus a and b cannot both be odd.
11. Case 3: Assume, without loss of generality, that a is odd and b is even.
12. Then a^2 is odd and b^2 is even. Hence $a^2 + b^2 = c^2$ is odd.
13. Since c^2 is odd, c is odd.

Appendix D. Excerpt of teacher-researcher intervention: Iteration 1, Session 2, Proof 4

[Students are handed a statement of Theorem 4, that a bounded, monotone sequence is convergent]

Interviewer: Bob, can you identify the definitions and key terms of the theorem?

Bob: Uh... bounded, monotone sequence, and convergent?

Interviewer: Does anyone know what it means for a sequence to be bounded?

Aneta: Uh, there is a small m and big M such that for all x_n in the sequence, $m \leq x_n \leq M$.

Interviewer: OK, does anyone have anything to add to that? [silence] Yes, that's right! All right, Rahim, can you tell us what monotone means?

Rahim: ...No.

Interviewer: OK, anybody?

Kirk: That means it's only entirely increasing [gestures with his hands to show the graph of an increasing function] or decreasing.

Interviewer: OK, does everybody agree? [students are nodding heads] Yes, I would agree with that, but we have to be a little careful. Monotone means, if it's monotone increasing, it's either staying the same or going up [gestures with his hands to show a function which is constant on an interval and then increasing]. So that's a little tweak about "strictly" monotone increasing. But once it starts going up, you know it's going to be going up and never going down.