COMPREHENDING LERON'S STRUCTURED PROOFS

Evan Fuller
Montclair State University
fullere@mail.montclair.edu

Juan Pablo Mejia Ramos, Keith Weber, Aron Samkoff, Kathryn Rhoads,
Dhun Doongaji, Kristen Lew
Rutgers University

In a widely cited paper, Leron (1983) proposed presenting proofs in a novel format that he called "structured proofs" and suggested that this format improves comprehension. In a qualitative study, we found that participants had difficulty with this format for several reasons, including a lack of familiarity with the format and the requirement on the reader to jump around between different sections of the proof. In a larger quantitative study, we found that students reading a structured proof were better than students reading a linear proof at identifying a good summary of the proof, but performed slightly (but not statistically reliably) worse on questions concerning justifications within the proof, transferring the ideas from the proof to another setting, and illustrating the ideas of the proof using examples.

Key words: Proof; Proof comprehension; Structured Proofs

1. Introduction

1.1. Proof presentation and comprehension in advanced mathematics courses

Advanced mathematics courses—that is, tertiary proof-oriented mathematics courses for mathematics majors—are typically taught in a lecture format, a significant portion of which consists of the professor presenting proofs of theorems to his or her students. Based on her observations of the lectures of three mathematics professors, Mills (2011) reported that roughly half the lecture time was spent on proof presentation. Numerous mathematicians and mathematics educators have commented that mathematics is typically presented to students in a definition-theorem-proof format (e.g., Davis & Hersh, 1981; Dreyfus, 1991; Thurston, 1994; Weber, 2004).

Presumably an assumption behind this pedagogical practice is that students can learn mathematics by reading and studying the proofs that their professors present or that appear in textbooks and notes. Although there have been few systematic studies on undergraduates' comprehension of proofs (as noted by Mejia-Ramos & Inglis, 2009), there are many who question whether this assumption is accurate. Both mathematicians and mathematics educators have remarked that students are generally confused by the content of a formal proof (e.g., Alcock, 2010; Davis & Hersh, 1981; Hersh, 1993; Leron & Dubinsky, 1995; Rowland, 2001; Thurston, 1994), with students reporting that the proofs they see are pedantic or pointless (e.g., Porteous, 1986). Further, empirical studies demonstrate that undergraduates have difficulty distinguishing between valid proofs and invalid arguments (Selden & Selden, 2003; Weber, 2010). There are, of course, many potential reasons that undergraduates gain little from the proofs that they read. However many researchers suggest that some student difficulties with proof are due to factors inherent in the way formal proofs are written.

Griffiths (2000) defines a proof as "a formal and logical line of reasoning that begins with a set of axioms and moves through logical steps to a conclusion" (p. 3). These types of proofs usually proceed linearly, incorporate logical syntax, and make little use of informal representations of the relevant mathematical concepts, such as diagrams and examples. Mathematicians and mathematics educators argue that this type of presentation might hinder proof comprehension for several reasons. The linear nature of proof presentation can prevent students from seeing the structure of the proof or the overarching method being applied in the proof, making the ideas of the proof appear to be unmotivated and mysterious (Davis & Hersh, 1981; Leron, 1983). The use of logical syntax and domain-specific jargon can be intimidating to students and mathematicians alike (Davis & Hersh, 1981; Hersh, 1993; Thurston, 1994). Finally, the formal presentation of proofs masks the intuitive mathematical ideas and models that were used to produce the proof (Dreyfus, 1991; Thurston, 1994).

To address this situation, several researchers have suggested alternative methods of proof presentation. These include the use of generic proofs (Rowland, 2001), informal visual arguments in lieu of formal proofs (Hersh, 1993), e-proofs (Alcock, 2009), and structured proofs (Leron, 1983). Although each of these approaches is interesting and potentially valuable, there is to date little to no empirical evidence that any of these approaches improve proof comprehension. Indeed, Roy, Alcock, and Inglis (2010) found that students who studied a proof using Alcock's e-proofs (Alcock, 2009) performed significantly worse on a comprehension post-test than students who observed the same proof in a lecture. We are not aware of any other studies that examine the effects of these novel proof presentations on students' proof comprehension.

1. 2. Leron's structured proofs

In this study, we examine the ways in which Leron's (1983) structured proofs can improve students' comprehension of a proof. Leron proposed a novel way to present proofs in terms of levels. The highest level (Level 1) provides a summary of the main ideas of the proof without providing detail on how these main ideas will be carried out. The next level (Level 2) provides a summary of how each of the main ideas will be implemented. Successively lower levels fill in the details of the implementation of higher levels of the proof. An additional feature in some structured proofs is an "elevator" between levels that provides a rationale for why the proof is proceeding the way that it is. Leron illustrated the nature of structured proofs by comparing a linear and a structured proof of the following claim: "There are infinitely many triadic primes". We used these proofs in our studies (nearly verbatim from Leron's article) and they are presented in the Appendix of this paper. The reader can note that in this proof, Level 1 lays out a summary of the three main aims of the proof, while the "elevator" provides the motivation for introducing and defining the variable M in a particular way.

There are several theoretical benefits for using structured proofs in lieu of traditional linear proofs. The format provides the reader with a summary of the proof and enables the reader to grasp the main ideas of the proof without getting lost in its logical details. However, no information is lost as this format still enables the reader to study or verify these logical details if he or she desires to do so. Finally, the higher-level structure and the "elevator" commentary make clear the reasoning behind some of the choices made in the proof that might otherwise seem arbitrary.

Numerous mathematics educators approvingly cite Leron's structured proofs as a potential way to improve students' proof comprehension (e.g., Alibert & Thomas, 1991; CadwalladerOlsker, 2011; Hanna, 1990; Hersh, 1993; Selden & Selden, 2003, 2008). In a book chapter providing research-based pedagogical suggestions to mathematicians, Selden and Selden (2008) recommended Leron's structured proofs as a means of helping students understand the proofs that they read and learn about the process of writing proofs. Mamona-Downs and Downs (2002) suggested that Leron's structured proofs may be influencing the ways in which proofs are currently written in advanced mathematics textbooks. Melis (1994) wrote that "Uri Leron shows how proofs are better comprehensible by structuring them into different levels" (p. 2), implying that the claim that structured proofs improve comprehension is not a theoretical hypothesis but rather an established fact.

Despite the theoretical promise of Leron's structured proofs and the esteem in which Leron's suggestion is held by some in the mathematical community, we are not aware of any empirical studies that suggest that structured proofs actually improve comprehension. Based on their observations of three students reading a structured proof using hyperlinks in a computer environment, Cairns and Gow (2003) noticed several difficulties that students had with this format. The novelty of the format did not map easily to the way that students were used to reading proofs (in a linear fashion), the students had difficulty evaluating their progress while reading the proof, and relevant information within the proof was often hidden or spaced far apart. However, given that Cairns and Gow investigated students' reading of proofs using hyperlinks (rather than as traditional text), and that the proofs were presented using an approach that was similar but not identical to Leron's (specifically, the approach of Lamport, 1995), the relevance of these findings to Leron's structured proofs should be viewed as suggestive at best.

1. 3. Research questions

The general aim of this paper is to examine the efficacy of Leron's structured proofs—to what extent do structured proofs improve comprehension, and what difficulties do students experience when reading them? However, we first emphasize the limitations and qualifications inherent in this type of research. Schoenfeld (2000) warns that questions such as "does group work improve mathematical learning?" are not well-formed. Before attempting to address this question, one would need to specify what type of group work was being implemented, what context was being studied, and what constituted learning. Even here, a single study or series of studies would not produce conclusive results as one would need to consider the quality of the implementation of the instruction, among other factors.

The pedagogical suggestion of structuring proofs can be implemented in a variety of ways. Students could be asked to read structured proofs as text in a short period of time, a student could be asked to study a structured proof overnight, or a professor could present a structured proof in lecture for his or her students. Structured proofs can even serve as an overarching paradigm for instruction, where many proof presentations involve structured proofs and students are involved in the structuring of some proofs themselves (indeed, a former student of Leron informed us that this was what he did in his own classroom). In this paper, we compare students' comprehension of a proof shortly after reading a linear or structured version of the same proof. However, any learning benefits that we did not observe in this study might occur if the structured proofs were used in a different way (and vice versa).

The two specific research questions we investigate are:

- What specific features of a structured proof help or hinder proof comprehension?
- To what extent do students who read a structured proof display a greater ability to (1) summarize the proof, (2) transfer the ideas of the proof to prove a different theorem, (3) illustrate the ideas of the proof with a specific example or diagram, or (4) see how particular assertions within the proof were justified, than students who read the same proof as a traditional linear proof?

Even here, as we only tested these ideas with two proofs, we do not view our answers as conclusive, but rather as the start of a conversation on the benefits and limitations of structured proofs.

2. Theoretical perspective

The assessment items that we use in this paper are based on our assessment model for proof comprehension (Mejia-Ramos et al, 2010). Based on interviews with mathematicians and our survey of the mathematics education literature, we developed seven different ways to assess students understanding of a proof. However, only four were utilized in the quantitative study that we focus on here:

- *Justification of claims*. Questions of this type assess whether students can justify how a new statement in a proof follows from previous assertions as well as how a particular statement is used to justify subsequent statements within the proof.
- *Summary*. Questions of this type assess whether students can recognize the overarching ideas used in the proof.
- *Transfer*. Questions of this type assess whether students can apply the ideas used in the proof in a new setting to prove a new theorem.
- *Illustration with examples*. Questions of this type assess examine whether students can relate the chain of argumentation in a proof to a particular example (i.e., transform the formal proof into a generic proof) or a relevant diagram.

The rationale for developing all seven categories, as well as a scheme for generating questions of each type, are described in detail by Mejia-Ramos et al (2010, submitted). Leron's structured proofs seem theoretically designed to help with the summary portion of this model (since Level 1 is essentially a summary of the proof) and transfer (since Level 1 and the between-levels elevator provide insight into the method being applied).

We note that our assessment model for proof comprehension is not designed to be hierarchical, but rather is a means to assess separate dimensions of proof comprehension. We can imagine one type of proof presentation helping students with some of the dimensions of this model while giving no benefit or hindering students with others.

3. Study 1: Qualitative interview study

3. 1. Rationale

We first conducted a set of interview studies in which students were videotaped while reading linear or structured proofs, answering questions about those proofs, and responding to open-ended questions about how they felt about the structured proof presentation. The purpose of these interviews was twofold: First, we wanted to examine how students read and reacted to structured proofs to gain insight into how this type of proof presentation might help or hinder students' comprehension of a proof. Second, we wanted to use students' behavior during our study to improve the quality of the proofs and assessment questions that we used during the quantitative study that we report in the next section.

3. 2. Methods

Participants. We recruited three groups of six participants to take part in this study. Participants were paid a fee for their participation. We recruited Groups A and B from math major courses at a large state university in the northeast United States in the beginning of the Fall 2010 semester.

After interviewing the students in Groups A and B, it became clear that some of them were deeply confused about the nature of structured proofs. We believed one cause of their difficulty was that they were not given a description about the nature and purposes of these structured proofs. At the end of the Fall 2010 semester, we recruited six more students and asked them to read two structured proofs, but only after giving them instructions about the nature of a structured proof (these instructions appear in the Appendix).

Materials. Two proofs were used for this study. We called the first proof the "Only Zero" proof, since it used a calculus-based argument to prove that 0 was the only solution to a given equation. The rationale for using the Only Zero proof is that we believed the content of the proof would be accessible to second- and third-year mathematics undergraduates and the proof could be effectively structured. The second proof, which we call the "Triadic Primes" proof, establishes that there are infinitely many triadic primes (i.e. primes congruent to 3 modulo 4). The linear and structured proofs we presented to participants in this study were taken nearly verbatim from Leron's (1983) original article on structured proofs. These proofs are both included in the Appendix of this paper.

Procedure. During each semi-structured interview, participants were given a version of the Only Zero proof and asked to read this proof until they felt that they understood it. The participants were informed that after they had studied the proof, the proof would be taken away from them and they would have to answer questions about what they had just read.

When participants finished reading the proof, they were asked to say, on a scale of one through five, how well they felt they understood the proof. They were then asked the questions about the proof, one at a time, and were given a sheet of paper with the question written on it. In interviews with Groups A and B, participants were given a set of multiple choice questions following the open-ended questions, most of which were similar to the questions they had just answered. As we explain shortly, one goal of having students answer these multiple choice questions was so we could refine the questions for our quantitative study. If participants read a structured version of the proof, they were asked how they felt about this new format of the proof, whether they liked this new format, and what they thought the strengths and weaknesses of this format were. This entire process was then repeated with the Triadic Primes proof. Since we found that we had sufficient data from Groups A and B in order to refine the multiple choice questions, Group C was given only open-ended questions in the interviews.

Group A received a linear version of the Only Zero proof and a structured version of the Triadic Primes proof. Group B received a structured version of the Triadic Primes proof and a linear version of the Only Zero proof. Group C first read the instructions describing the nature of structured proofs that are presented in the Appendix and then read structured versions of both the Only Zero and Triadic Primes proofs. Interview lengths ranged from 60 to 150 minutes.

Analysis. Our analysis consisted of two phases. First we sought to identify attributes of structured proofs that may have aided or hindered the participants' comprehension of the proof. In an initial pass through the data, we noted three sources of difficulty that participants had with structured proofs: Participants expressed displeasure that the proof seemed to "jump around", they had misconceptions about the nature or purpose of structured proofs, and they failed to understand the high-level summary expressed in the top level of the structured proof.

In the second phase of analysis, we aimed to improve the readability of the structured proofs and the quality of the questions for the large-scale quantitative study that we conducted subsequently. To this end, we noted any common confusions that participants had about the proof that was not related to the format of the proof and sought to revise the proofs to avoid these errors. For instance, many participants had trouble inferring what "triadic prime" meant from the definition of triadic number. This was explicitly explained in the proof for the follow-up study. We also examined any instances in which participants answered the open-ended questions incorrectly but answered the multiple choice questions correctly, and vice versa, and formulated new multiple choice questions that more accurately measured participants' understandings of the proofs that they read.

3. 3. Results

We noticed three common difficulties that participants had when reading structured proofs that may have inhibited their comprehension of the proof.

Jumping around: Fourteen of the eighteen participants had difficulty reading the structured proofs because they felt the proof jumped around between the main ideas of the proof, including five of the six participants who received instructions on structured proofs. We provide three representative comments below, including two from students in group C (C1 and C3) who received instructions:

> B3: I'm not really sure the reason why it couldn't be just one flowing proof. It kind of bounced from parts, like from 2b, it would go to 3b, or whatever it is.
>
> I: Do you think this type of format increased your understanding of the proof, decreased your understanding, or neither?
>
> B3: Probably decreased it slightly, just because it was moving around with the statements a lot, jumping from the middle of the page to the bottom of the page. Things weren't flowing one after another, which made it a little more confusing[…] it just didn't allow for one thing to flow to the next in a way that was easy to read.
>
> I: In what ways if any did this format hinder your understanding of the proof?
>
> B3: The same way. It just bounced around too much.

> I: What did you think of the format?
>
> C1: I kind of like it because in a sense it's kind of how I would break down proofs sometimes. Because when I do a proof, if I think of something, I'll do a little scratch work over here and then more scratch work over here, and I had a sense that it was like

that. But since there was just so much stuff and you had to keep going back and drawing arrows, that was the only confusing part […] Just because someone says something is true doesn't necessarily mean it's true, so I kind of want to show for myself that it's true, so that's what it did. But just because the stuff was all over the place and there was like in three places like we support the claim in 2c, but by the time I got down here I forgot what 2c was, so I had to keep going back and drawing. [This participant drew arrows between the assertions in Level 1 and the sub-proofs in Level 2].

C3: But if you were to read it like that like I did the first time, you're jumping around everywhere. Because you're saying level two, but I don't really read level two because I see more text down here. So I figure let me read this first. [the rest of level 1]. But then it gets very confusing. Because for example I read 2c, and I go back to 3a but then they're talking about the claim in 2b. But I don't remember the claim in 2b, so I have to go back and read 2b properly. But then I have to remember how does this relate to what the goals were? And go back here, and I'm going all over the place.

The "jumping around" is, of course, a crucial feature of structured proofs, but the students comments reveal a limitation of this feature—specifically, participants have to keep several different ideas in their heads as they read these proofs, which likely puts a strain on their working memories. This difficulty that the participants experienced is related to Cairns and Gow's (2003) that this genre of proofs does not keep relevant information near each other.

Failure to understand the high-level summary. For the structured Triadic Primes proof, many participants could not understand the main ideas expressed in Level 1 of the proof. When asked to provide a summary of the structured Triadic Prime proof, six of the twelve participants who read this proof either provided no response or a response that was fundamentally flawed. For instance, one participant, C1, began her summary by defining $3p_2p_3\ldots p_n + 1$ as a triadic prime. Further, participants who did provide sensible summaries claimed that this was not due to Level 1, which they found confusing, but rather from what they gained from studying the proof itself. Consider A1's comments below:

A1: Reading the proof for the first time I got a little confused, because we're talking about M as if it had already been constructed but then I was starting to think 'Oh well what's M?', and I realized that it hadn't actually been constructed yet. You had to get down here (pointing to a lower level) and I was trying to think about it and I was like, 'Wait! What's M? Am I missing something?' And then when I got further down I understood.

From our perspective, a substantial benefit of including Level 1 in a structured proof is to provide a framework or schema to integrate the chain of deductions that follow by providing them with purpose. However, if participants fail to adequately understand Level 1, they will either be trying to integrate the ideas from Level 2 into an inappropriate framework, which could lead to confusion, or they might forget the ideas contained in Level 1, leading to a strain on their working memory and a dissatisfaction that the proof was jumping around. This concern is conveyed nicely in the comments from C2.

C2: [Structured proofs] were a disaster, to be blunt. That proof using Rolle's theorem and using derivatives and stuff. I took analysis so we did proofs like that for an entire semester. So when I was faced with that proof, I already knew where it was going, I had seen things like it before, it was very familiar to me. So even though the style of the proof

> wasn't exactly my cup of tea I was still able to absorb a lot of the information because I was familiar with saying ok you know we're using a derivative argument, we're using a Rolle's theorem argument and what not. But for this proof [the Triadic Primes Proof], like I said the only proof I've ever seen in that form was the proof that there were infinitely many primes and if you asked me to write that down right now, I probably wouldn't be able to, to be honest. Although I think it's very elegant, it's not a proof skeleton or a proof form that I use very often, so although I knew what the proof was getting at, I wasn't able to absorb it as well simply because it's not something I see very often.

C2 argues that he was able to understand the structured proof of the Only Zero theorem in spite of the structured proof format because he had familiarity with the type of argumentation used in the proof. He lacked similar experience in number theory and consequently could not use the proof skeleton in Level 1 to "absorb" the details of the proof. Cairns and Gow (2003) note that structured proofs might be ineffective for students who do not possess a strong enough understanding of the material to see why the proof is structured the way that it is.

Failure to understand the nature of structured proofs. Participants in our study often held significant misconceptions about what a structured proof was and how a structured proof works. Three participants believed the structured proofs that they read were proofs by cases and one participant thought the structured proof was three different independent proofs of the same claim. Five participants expressed confusion regarding how the concluding sentence for why the theorem was proved occurred in Level 1, whereas the final sentence of the proof that appeared in Level 3 did not establish the theorem. For instance, when reading the structured triadic primes proof, one participant, C4, reached the end of the proof and said, "Is that it?", and turned the paper that the proof was written on over to look for more text. When told that was the proof in its entirety, C4 remarked, "I feel like we didn't prove anything." Later, when asked why he didn't understand the proof, C4 remarked:

> C4: I feel like we didn't prove anything […] I just didn't understand the ending at all […] At the end, showing that when you multiply each together you're going to get that kind of number [the end of the proof was a proof of the assertion that the product of monadic numbers was monadic], I'm going to say I'm really confused.

A main motivation behind structured proofs is that the proofs of the lemmas used in the proofs, which are less significant to the main line of argumentation, appear near the end of the proof to reflect their relative lack of importance. However, the participants who read these proofs did not appreciate this and expected the last line of the proof to conclude with the theorem being proven.

Most likely one contributing cause to this difficulty was the participants' lack of familiarity with this way of writing proofs and this would likely be mitigated if participants were given more experience with structured proofs. Indeed, eight participants remarked on the novelty of this format with some suggesting that their lack of familiarity with structured proofs may have limited comprehension. For example, in describing the structured proofs, B5 said, "It was kind of all over the place. I've never seen anything like it. [laughing] It was just bizarre." Cairns and Gow (2003) also noted that students' lack of familiarity with these types of proofs might pose a barrier to comprehension.

4. Study 2: Quantitative internet study

4. 1. Rationale for this study

In the previous section, we highlighted several potential attributes of structured proofs that might inhibit students' comprehension of a structured proof. However these data are limited for two reasons. First, the fact that students found some aspects of structured proofs to be problematic does not mean that structured proofs might not be beneficial for comprehension. It is possible that the other theoretical benefits of structured proofs more than compensate for the reported drawbacks. It is also plausible that the difficulties that students experience are not drawbacks at all; for instance, even though "jumping" between different sections of the proof was unpopular with our participants because it made the process of reading the proof longer and more difficult, perhaps this activity helped students make connections between ideas in the proof that they would have missed if they had read the proof in a linear format. Second, the sample of eighteen students reading structured proofs was relatively small, and only six students read structured proofs after being given instruction about the nature of structured proofs. Hence the findings from the qualitative interview study might not be generalizable.

The goal of the larger quantitative internet study was to search for systematic evidence with a larger number of students that structured proofs might help students summarize proofs, transfer the ideas of the proof to other settings, understand how the proof relates to a specific example or diagram, and see how particular statements within the proof were justified.

4. 2. Research methods

To maximize our sample size, we conducted the quantitative study recruiting mathematics majors using the internet. The validity and reliability of this type of studies have been extensively discussed in the research methods literature (e.g. Gosling et al., 2004; Reips, 2000). To deal with the common threats of validity, we employed the methodology of Inglis and Mejia-Ramos (2009) to discard participants who showed evidence from participating in the experiment twice (e.g., multiple submissions from the same IP address) or who had revisited earlier pages in the study.

Participants. We recruited 300 mathematics undergraduate students from 50 top-ranked mathematics departments in US universities. These students participated without payment and were recruited via an email from their departmental secretary. The email explained the purpose of the experiment and asked third and fourth year mathematics major/minor students to visit the experimental website if they wished to participate.

Materials. All materials used for this study are presented in the Appendix. For the internet study we used the two proofs employed in the qualitative interview study (i.e. the Only Zero proof and the Triadic Primes proof) with minor modifications. As mentioned earlier, we made these minor modifications in order to avoid confusion unrelated to the format of the proof. For instance, for the "triadic primes" proof, in addition to defining *triadic (monadic) numbers* right before the theorem and its proof, we added a statement clarifying the meaning of *triadic (monadic) primes*.

For each proof, we designed a proof comprehension test consisting of four (for the Triadic Primes proof) or five (for the Only Zero proof) items. As mentioned earlier, these assessment items were selected and improved from the questions used in the interview study.

Procedure. Depending upon the school they were at, participants were invited to read the Only Zero proof or the Triadic Primes proof. Once participants had loaded the experimental website, they were randomly assigned into one of three conditions: LIN where participants were presented with the linear version of the proof; S(no dir) where participants were presented with the structured version of the proof but with no directions; and S(w dir) where participants were presented with the structured proof after reading the instructions describing the nature and purpose of structured proofs (in the Appendix).

After reading the proof, participants were asked, "How well do you feel you understand this proof?" A five-point Likert scale was used to record each participant's reported level of understanding. Next, participants took the corresponding comprehension test. Each of the questions appeared in a new screen and the order in which they were displayed was randomized for each participant. Upon completion of the test, participants in a structured condition were asked to use a three-point Likert scale to report the extent to which they liked the format in which the proof had been presented.

4. 3. Results

We found no statistically significant differences between the performance on the items between the Structured-No Directions and Structured-Directions condition and, surprisingly, the Structured-No Directions did better than the Structured-Directions conditions on most items. Consequently, for the purposes of comparing the efficacy of structured proofs, we collapsed the Structured-No Directions and Structured-Directions groups into a single Structured group in what we report here. Participants' performance in the comprehension tests is presented in Table 1.

Table 1: Performance in linear versus structured conditions.

Only Zero proof

Condition	Summary Diagram	Transfer	Justification	Justification Question 1	Application to Question 2
Lin (N=67)	63%	61%	85%	85%	69%
Struct (N=135)	76%	53%	88%	80%	78%

Triadic Primes proof

Condition	Summary	Transfer		Application to Justification	Examples
Lin (N=33)	67%	45%	55%	91%	
Struct (N=65)	74%	38%	40%	63%	

Summary. The Structured Proof group did better on the summary questions than the Linear Proof group. For the Only Zero proof, the Structured Proof group answered the summary question correctly 76% of the time as compared to 63% for the Linear Group, a statistically reliable difference (c(1, 200)=4.095, p<0.05). For the Triadic Primes proof, the Structured Proof also did better on the summary question, although this difference was not statistically reliable (c(1, 96)<1). In structured proofs, Level 1 essentially provides a summary of the proof and, while some students may find it difficult to understand (see section 3), others may grasp it. The data demonstrates that some of the participants who read a structured proof were better able to retain this information and recognize a high-quality summary.

Transfer. We did not observe evidence that the Structured Proof group outperformed the Linear group on the transfer questions. On the contrary, although inconclusive, the evidence suggests a trend in the opposite direction. For both the Only Zero and Triadic Primes proofs, the Linear Group had a higher percentage of students who answered the transfer question correct, although neither of these results were statistically reliable (c(1, 200)=4.095, p=0.24, c(1, 96)<1, p=0.43). Consequently, the data provide no evidence that structured proofs improve students' ability to answer these types of transfer questions and suggestive evidence that they hinder students' abilities in this regard.

Justification. We did not observe evidence that the Structured group outperformed the Linear group on the justification questions. We found essentially no difference between participants' performance on the Justification questions for the Only Zero proof. For the Triadic Primes proof, the Linear group performed somewhat better than the Structured group (55% vs. 40%), but this difference was not statistically reliable (c(1, 96)=1.872, p=0.17). Hence, like the transfer questions, our results are inconclusive, but to the extent that they suggest anything, they favor the Linear group.

Illustration with examples. This was the only category in which we observed different findings on the Only Zero and Triadic Primes proofs. On the Triadic Primes proof, the Structured Group performed better than the Linear Group in recognizing that the ideas in the argument would not be applicable to a graphed function (78% vs. 69%), although this difference was not statistically reliable ((c(1, 200)=1.974, p=0.16). However, for the Triadic Primes proof, the Linear Group did substantially better on the question illustrating the ideas of the proof with a specific example (91% vs. 63%; c(1, 96)=9.252, p<0.01). In hindsight, it appears that evaluating whether the ideas of a proof *cannot* be applied to a specific diagram and actually applying the ideas of the proof to a specific example may be two very different skills.

In analyzing the conflicting results for illustration with examples, we suggest two factors. First, once *M* is chosen in the Triadic Primes proof, the linear version of the proof essentially offers a step-by-step recipe for reaching a contradiction with this *M*. For the structured proof, this information is dispersed throughout the proof and more emphasis is given to overarching methods and the motivation for choosing *M*. Reading a linear proof may help students with these types of questions because the step-by-step procedure is made explicit. Second, our post-hoc analysis reveals that when Leron (1983) originally posed the structured proof for the Triadic Primes Theorem, he did not just change the presentation of the linear proof, but he also changed the *content* of the linear proof as well. The linear version of the Triadic Primes proof reaches a contradiction by showing that *M* must be monadic and triadic while the structured version of the proof arrives at a contradiction by showing that *M* must have a triadic prime factor not included in the finite list of triadic primes. As a result, the assessment question that we asked the participants may have unintentionally favored the linear group, leading to the observed differences. Hence, despite the statistically reliable effect, we also view this result as inconclusive.

Participants' subjective evaluation of structured proofs. Detailed results on this topic will be presented in a forthcoming paper. For the purposes of brevity, we provide a summary here. There was remarkably little difference between the participants' condition and how well they understood the proof that they read. However, we found that participants in the structured proof with directions were about 50% more likely to view the structured proof format positively and 50% less likely to view the proof negatively. Hence, although providing directions did not improve students' professed understanding of a structured proof or their performance on the comprehension of the test, it did lead them to view structured proofs in a more positive light.

Finally we note that the correlations between participants' reported understandings of the proof and their cumulative comprehensive scores were negative, both for the Only Zero proof (-0.54) and the Triadic Primes proof (-0.185), suggesting that participants' judgments of their understandings should not be used as a primary means to assess the efficacy of new proof formats as some researchers have done (e.g., Rowland, 2001).

5. Discussion

5. 1. Summary of main results and caveats

Our findings suggest that the summary provided in Level 1 of a structured proof can be a double-edged sword. On the one hand, if readers are able to understand that summary (which may depend on their familiarity with the subject and the difficulty of the proof) they will have a better chance of grasping the higher level ideas of the proof. Indeed, in the quantitative internet study, we found that participants who read structured proofs performed better on questions pertaining to a proof summary than participants who read an analogous linear proof of the same theorem, suggesting that some students were able to grasp the summary given in Level 1. However, as suggested by our qualitative study, this might not be the case, especially if the reader is not familiar with the ideas in the proof.

Further, we failed to find evidence that structured proofs improve participants' abilities to answer questions pertaining to transfer, justification, or relating the ideas of the proof to specific examples. In fact, while the evidence for each of these three categories was inconclusive, in each case the data suggested an effect in the opposite direction. The results from our qualitative interview study suggest an account for these findings. Students reading a structured proof may have difficulty comprehending the details of the proof, both because mastering the details of the proof requires the participants to integrate information that is far apart in the proof and because some students may have failed to understand the summary of the proof that they had read. These difficulties may have been exacerbated by their lack of familiarity with and understanding of the nature of structured proofs.

Any conclusions reached from this paper should be viewed as tentative for the reasons mentioned when we posed our research questions. We view our data as a contribution to the still open question of whether and how structured proofs improve proof comprehension. While our data are limited, we are not aware of studies that show that structured proofs improve students' appreciation for proof or that delivering structured proofs in a lecture format leads to more learning benefits than presenting these proofs as text. As we argue below, we strongly encourage such research. We hope the results from our qualitative interview study offer guidance on how such studies might be conducted.

5. 2. Significance of results

We discuss how our findings relate to three hypotheses about the pedagogical value of structured proofs: (a) Implementing structured proofs in instruction in a straightforward manner will lead to large and clearly observable learning gains, (b) with careful instruction and preparation, structured proofs have the potential lead to observable learning gains, and (c) structured proofs do not have the potential to lead to significant learning gains.

We believe our data provide evidence against hypothesis (a). Participants who read structured proofs were better able to identify an accurate summary of the proofs than their counterparts who read a linear proof. However, this is most likely because a summary was explicitly provided for them in Level 1 of the structured proof. It is possible that the same benefit could be observed in a linear proof if the linear proof began by providing a summary. We did not observe benefits for other assessment items, meaning that presenting structured proofs as text did not lead to large or observable learning gains. It is possible that using different assessment items or delivering structured proofs as a lecture or in another format would have yielded large and observable learning gains, but as of now, there is no empirical evidence that this is the case.

Some mathematics educators familiar with the challenges of designing instruction might object that it is naïve to expect large learning gains from a relatively simple instructional intervention, arguing that merely tinkering with the format of a proof would not be sufficient to improve student comprehension. However, if we accept this objection, then the suggestions in the literature that structured proofs can improve comprehension should be tempered. We should not recommend that mathematics professors implement structured proofs in their classrooms without also providing them with further recommendations on how they can be implemented effectively. Claims that structured proofs improve comprehension (e.g., Melis, 1994) should be qualified appropriately.

Our data do not distinguish between hypotheses (b) and (c), as we believe a more careful and nuanced implementation of Leron's structured proofs may lead to substantial learning gains. Hence we view it as a worthwhile direction for future research to design and assess instruction based on Leron's structured proofs that demonstrates its effectiveness. We feel that this research is valuable for three reasons. First, if we are going to make research-based suggestions to mathematicians on how to improve their pedagogical practice (as Selden and Selden (2008) do for structured proofs), it is important that we have empirical support for these suggestions. There are instances when theoretically-based pedagogical suggestions lead to minimal, or negative, learning gains, even in the domain of proof presentation (Roy, Alcock, & Inglis, 2010). Second, a positive demonstration of the efficacy of Leron's structured proofs would not only provide evidence *that* this improves learning, but would also provide insight into *what actions a teacher can take to make them work*.

Finally, as Alcock (2010) notes, implementing innovative pedagogy in an advanced mathematics classroom often involves making trade-offs. For instance, it might well be the case that structured proofs lead to increased proof comprehension; however, this only occurs with extended training on how these proofs should be read and with a lengthy, detailed classroom presentation. Ultimately, choosing the time to devote to different activities in advanced mathematics courses involves value judgments by the professor. However, without knowing the types and magnitudes of the learning gains of structured proofs or what a professor would need to do for these to be effective, we are not yet in a position to make these judgments.

REFERENCES

Alcock, L. (2009). e-Proofs: Students experience of online resources to aid understanding of mathematical proofs. In *Proceedings of the 12th Conference for Research in Undergraduate Mathematics Education*. Available for download at: http://sigmaa.maa.org/rume/crume2009/proceedings.html. Last downloaded April 10, 2010.

Alcock, L. (2010). Interactions between teaching and research: Developing pedagogical content knowledge for real analysis. In R. Leikin & R. Zazkis (Eds.) *Learning through teaching mathematics*. Springer: Dordrecht.

Alibert, D. and Thomas, M. (1991). Research on mathematical proof. In D. Tall (Ed.) *Advanced Mathematical Thinking*. Kluwer: The Netherlands.

CallwalladerOlsker, T. (2011). What do we mean by mathematical proof? *Journal of Humanistic Mathematics*, 1, 33-60.

Cairnes, P. & Gow, J. (2003). A theoretical analysis of hierarchical proofs. In A. Asperti, B. Buchberger & J.H. Davenport (Eds.) *Mathematical knowledge management. Proceedings of MKM 2003*. LNCS 2594. Springer.

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

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

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

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

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

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

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

Lamport, L. (1995). How to write a proof. *American Mathematical Monthly, 102,* 600-608.

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

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

Mamona-Downs, J. & Downs, M. (2002). Advanced mathematical thinking with a special reference to mathematical structure. In English, L. (Ed.) *Handbook of international research in mathematics education*. Lawrence Erlbaum Associates: Mahwah, NJ.

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. Weber, K., Fuller, E., Samkoff, A., Search, R., & Rhoads, K. (2010). Modeling the comprehension of proof in undergraduate mathematics. In *Proceedings of the 13th Conference on Research in Undergraduate Mathematics Education*. Raleigh, NC.

Available for download at: http://sigmaa.maa.org/rume/crume2010/Abstracts2010.htm. Last downloaded March 20, 2011.
Mejia-Ramos, J.P. Weber, K., Fuller, E., Samkoff, A., Search, R., & Rhoads, K. (submitted). An assessment model for proof comprehension in undergraduate mathematics.
Melis, E. (1994). How mathematicians prove theorems. In *Proceedings of the 16th Annual Cognitive Science Conference*. Atlanta, GA.
Mills, M. (2011). Mathematicians' pedagogical thoughts and practices in proof presentation. Presentation at the *14th Conference for Research in Undergraduate Mathematics Education.* Abstract downloaded from: http://sigmaa.maa.org/rume/crume2011/Preliminary_Reports.html.
Porteous, K. (1986). Children's appreciation of the significance of proof. *Proceedings of the Tenth International Conference of the Psychology of Mathematics Education* (pp. 392-397). London, England.
Reips, U.-D. (2000). The web experiment method: Advantages, disadvantages, and solutions. In M. H. Birnbaum (Ed.), *Psychological experiments on the internet* (pp. 89–117). San Diego: Academic Press.
Rowland, T. (2001). 'Generic proofs in number theory.' In S. Campbell and R. Zazkis (Eds.) *Learning and teaching number theory: Research in cognition and instruction*. (pp. 157-184). Westport, CT: Ablex Publishing.
Roy, S., Alcock, L., & Inglis, M. (2010). Supporting proof comprehension: A comparison study of three forms of comprehension. In *Proceedings of the 13th Conference for Research in Undergraduate Mathematics Education*. Available for download from: http://sigmaa.maa.org/rume/crume2010/Abstracts2010.htm
Schoenfeld, A. (2000). Purposes and methods of research in mathematics education. *Notices of the AMS*, 47*,* 641-649.
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. (2008). Overcoming students' difficulties with learning to understand and construct proofs. In M. P. Carlson and C. Rasmussen (Eds.) *Making the Connection: Research and Teaching in Undergraduate Mathematics Education,* MAA Notes: Washington, D.C.
Thurston, W.P. (1994). On proof and progress in mathematics, *Bulletin of the American Mathematical Society*, 30, 161-177.
Weber, K. (2004). Traditional instruction in advanced mathematics classrooms: A case study of one professor's lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior*, 23(2), 115-133.
Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning*, 12, 306-336.

APPENDIX- Materials used in the quantitative internet study.
Directions for participants:

Recently a mathematician proposed a new way of presenting proofs to make them easier to understand. However this way of presenting proofs has not been tested with students. The purpose of this study is to see how students understand proofs written in this format.

This format presents proofs in levels. Level 1, the top level, gives in very general terms a description of how the proof will proceed. Level 2 carries out the arguments described in Level 1. If there are some logical details or computations for some of the ideas in Level 2, these details may be pushed down until Level 3.

The motivation behind this presentation is that the reader can gain a general sense of what the proof will do before seeing all the detailed arguments.

In reading this type of proof, you may encounter a "between levels" section. In this section, ideas that will be used in later levels are introduced. This way, statements in the proof will be motivated and not appear to come out of nowhere.

Proofs used in the studies:

Structured Linear version

Claim. The equation $x^3 + 5x = 3x^2 + \sin x$ has no nonzero solutions.

Proof:

Level 1. We define $f(x) = x^3 - 3x^2 + 5x - \sin x$. Solutions of $f(x$ precisely correspond to solutions of $x^3 + 5x = 3x^2 + \sin x$. Assume the cl false: then $f(x) = 0$ has a nonzero solution. We show (in Level 2):

a. $f'(x) > 0$ for all x

b. If $f(x) = 0$ has a nonzero solution, then there is a number c for $f'(c) = 0$.

Together, these conclusions clearly produce a contradiction, so the cl proved.

Level 2a. $f'(x) = 3x^2 - 6x + 5 - \cos x$. Using algebra (in Level 3a show this expression is always positive.

Level 2b. Suppose $f(x) = 0$ has a nonzero solution, that is, there $s \neq 0$ and $f(s) = 0$. Since $f(0) = 0$ (Level 3b), this implies there is a that $f'(c) = 0$, contradicting the fact that $f'(x) > 0$. for all x, which w tablished in Level 2a. (The details are given in Level 3c).

Level 3a. We support the claim in Level 2a as follows:

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

Since $3(x-1)^2 \geq 0$ and $2 - \cos x > 0$ for all real numbers x, $f'(x) > 0$ real numbers x.

Level 3b. $f(0) = 0^3 - 3(0)^2 + 5(0) - \sin 0 = 0$

Level 3c. We support the claim in Level 2b as follows: Rolle's theorem that if a differentiable function f has the property that $f(a) = f(b)$ for then there is a c such that $a < c < b$ and $f'(c) = 0$. In our case, we $f(0) = f(s) = 0$. Hence there is a c between 0 and s such that $f'(c) = 0$.

We define a number as *monadic* if it can be represented as $4j + 1$ for integer j, and *triadic* if it can be represented as $4k + 3$ for some intege triadic (monadic) prime refers to a number that is both triadic (monadi prime. Note that every odd prime is either a monadic prime or a triadic

Claim. There exist infinitely many triadic primes.

Proof:

Level 1. Suppose the theorem is false and let $p_1, p_2, \ldots, p_n$ be all the primes. We construct (in Level 2) a number M having the following prop

a. M as well as all its factors are different from $p_1, p_2, \ldots, p_n$;

b. M has a triadic prime factor.

These two properties clearly produce a contradiction, as we get a triadic which is not one of $p_1, p_2, \ldots, p_n$. Thus, the theorem is proved.

Between levels: How will we define M? It is natural to try $M = p_1 p_2 \cdots$ which meets requirement (a) but not (b). In fact, since M itself may tu to be prime, it must be triadic to meet requirement (b). A natural seconc is $M = 4p_1 p_2 \cdots p_n + 3$. However, since $p_1 = 3$, M is divisible by 3, in vi of (a). This 'bug' is easy to fix: simply eliminate 3 (i.e. p_1) from the p in our definition of M.

Level 2. Let $M = 4p_2 \cdots p_n + 3$. M is clearly triadic. We show tl satisfies the two requirements from Level 1.

Level 2a. Requirement (a) means that no p_i should divide M. I $p_2, p_3, \ldots, p_n$ do not divide M as they leave a remainder of 3, and 3 dc divide M as it does not divide $4p_2 \cdots p_n$.

Level 2b. As for requirement (b), suppose on the contrary that all prime factors were monadic. Then M, as a product of monadic numbers, itself be monadic (Lemma, Level 3), which is a contradiction.

Level 3. *Lemma:* The product of monadic numbers is again a monadic nu *Proof of lemma*: Consider a product of two monadic numbers:

$$(4j+1)(4k+1) = 4k \cdot 4j + 4k + 4j + 1 = 4(4jk + j + k) + 1$$

which is again monadic. Similarly, the product of any number of monadi bers is monadic.

Questions used in the quantitative study:

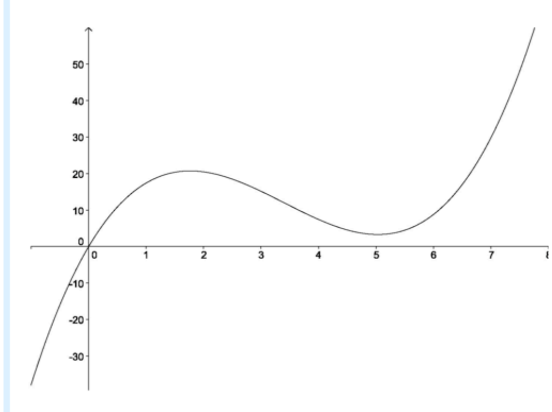
	Triadic Primes proof	Only Zero proof
Summary	Which of the following is a better summary of this proof, A or B? A. It lists all triadic primes and then uses this finite list of triadic prim obtain a contradiction by constructing a number that is triadic bu no triadic prime factors. B. It lets $M = 4p_2 \cdots p_n + 3$, where $p_i \neq 3$. 2 does not divide M bec M is odd. p_i does not divide M because it leaves a remainder of 3. produces a contradiction. C. I don't know which summary would be better.	Which of the following is a better summary of this proof, A or B? A. $f(x) = x^3 - 3x^2 + 5x - \sin x$. If there were non-zero s for which f then there would be a point where $f'(x) = 0$. However $f'(x)$ i positive. Hence 0 is the only solution. B. Since $f(x) = x^3 - 3x^2 + 5x - \sin x$, then $f'(x) = 3x^2 - 6x + 5 -$ $3(x-1)^2 + 2 - \cos x$. If s is a solution, $f(s) = 0$. A non-zero would mean that there is a c such that $f'(c) = 0$, which is a contra C. I don't know which summary would be better.
Transfer	If one were to adapt the proof to show that there are infinitely many p of the form $6j+5$, what would be an appropriate definition for M? A first that there are finitely many primes of this form, say $p_1, p_2, \ldots, p_n$ $p_1 < p_2 < \ldots < p_n$. Please type your answer in the space below. For example, if you th that the definition of M should be $2p_1 + 2p_2 + \cdots + 2p_n$, you could e this as "2p1+2p2+...+2pn."	Why is it impossible to use the ideas from the proof to show that the $x = \sin x$ has no nonzero solutions? A. To use Rolle's theorem, the function $f(x) = x - \sin x$ would ne differentiable. B. There is a nonzero solution to this equation. C. $f(x) = x - \sin x$ has critical points. D. None of the above. E. I don't know.
Justification	How was the conclusion that "the product of any number of monadic nur is monadic" used in the proof? A. It was used to show that no triadic prime can divide M. B. It was used to show that M must be monadic. C. None of the above. D. I don't know.	Where was the fact that $f(0) = 0$ used in this proof? I. It was used to show that $f'(x) > 0$ for all real numbers x. II. It was used to show that $f(s) = 0$, where $s \neq 0$, would imply a cc tion. A. It was used to show I but not II. B. It was used to show II but not I. C. It was used for both I and II. D. It was used for neither I nor II. E. I don't know.
		How was the fact that $f'(x) > 0$ for all real numbers x used in this prc A. It was used to show that $f(s) = 0$. B. It was used to show that solutions of $f(x) = 0$ precisely corres solutions of $x^3 + 5x = 3x^2 + \sin x$. C. It was not used to show either A or B. D. I don't know.

Application to examples/diagrams

Supposing 3, 7, 11, and 19 were the only triadic primes, which of the foll
illustrates the main steps of the proof?

- ○ A. Let $M = 4 \cdot 7 \cdot 11 \cdot 19 + 3$. 7, 11, and 19 do not divide M since leave a remainder of 3, and 3 does not divide M since it does not $4 \cdot 7 \cdot 11 \cdot 19$. Thus, M has only monadic factors and is monadic. How M is triadic because it is of the form $4k + 3$. This is a contradictio
- ○ B. Let $M = 4 \cdot 7 \cdot 11 \cdot 19 + 3$. We show that M is divisible by a triadic prime than 3, 7, 11, and 19, which is a contradiction. $M = 4 \cdot 7 \cdot 11 \cdot 19 + 3 =$ The prime factors of M are 5 and 1171. However, $1171 = 4 \cdot 292 +$ 1171 is triadic. This is a contradiction.
- ○ C. Neither of the above illustrates the main steps of the proof.
- ○ D. I don't know.

The graph of $g(x)$ is given below. Is it possible that the ideas of the pr
be used to show that the equation $g(x) = 0$ has no nonzero solutions?



- ○ A. Yes.
- ○ B. No.
- ○ C. I don't know.