# COMPREHENDING STRUCTURED PROOFS

**Evan Fuller**[1]
**Keith Weber**[2]
Rutgers University
**Juan Pablo Mejia-Ramos**[3]
Rutgers University
**Kathryn Rhoads**
Rutgers University
**Aron Samkoff**
Rutgers University

## ABSTRACT

In a widely cited paper, Leron (1983) proposed presenting proofs in a novel format that he called "structured proofs" and suggested that presenting proofs in this format improved students' comprehension. Our research investigates how structured proofs might aid or hinder students' comprehension. In a qualitative study, we presented structured proofs to students to examine how they read and perceived this type of proof presentation. Although some students valued the summaries contained in structured proofs, many complained that structured proofs "jumped around" and required them to scan different parts of the proof to coordinate information. In a larger quantitative study, we found that students who had read a structured proof were better at identifying a good summary of the proof than students who had read a linear proof, but performed somewhat (although usually not statistically significantly) 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; Undergraduate mathematics education.

[1] fuller.evan@gmail.com
[2] keith.weber@gse.rutgers.edu
[3] pablo.mejia@gse.rutgers.edu

# 1. INTRODUCTION[4]

## 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 lecture format where a significant portion of these lectures consists of the professor presenting proofs of theorems to his or her students. For instance, based on her observations of the lectures of three mathematics professors, Mills (2011) observed 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 will learn mathematics by reading and studying the proofs that their professors present. However, many question whether this assumption is justified. Both mathematicians and mathematics educators have remarked that students generally are confused by the content of formal proofs (e.g., Alcock, 2010; Davis & Hersh, 1981; Hersh, 1993; Leron & Dubinsky, 1995; Mamona-Downs & Downs, 2002; Rowland, 2001; Thurston, 1994; Weber, 2010; Weber & Mejia-Ramos, 2014).

Griffiths (2000) defined 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 formal syntax, and make little use of informal representations of the relevant mathematical concepts, such as diagrams and examples. Mathematicians and mathematics educators have argued that this type of presentation might hinder comprehension for several reasons. Some researchers claim that the linear nature of proof presentation can prevent students from seeing the structure of the proof or the overarching method being applied in the

[4] An earlier version of this paper appeared in the *Proceedings of the 14th Conference for Research in Undergraduate Mathematics Education* (Fuller et al, 2011), which is not copyrighted. This earlier manuscript differs from this paper in that the Interview Study Group A and Interview Study Group B are now treated as pilot data.

proof, making the ideas of the proof appear mysterious (Anderson, Boyle, & Yost, 1986; Davis & Hersh, 1981; Leron, 1983). The use of formal syntax and domain-specific jargon can be intimidating to students and mathematicians alike (Davis & Hersh, 1981; Hersh, 1993; Kline, 1973; Thurston, 1994). Several researchers have argued that the formal presentation of proofs masks the intuitive mathematical ideas and models that were used to produce (and are needed to comprehend) these proofs (Dreyfus, 1991; Hersh, 1993; Thurston, 1994). Consequently, some mathematics educators have proposed that proofs would be more easily understood if they were presented in a different manner.

### 1. 2. Alternative formats for presenting proofs

Several researchers have proposed alternative methods of proof presentation. Rowland (2001) suggested using "generic proofs": proofs that illustrate why a general theorem is true by showing how the theorem holds for a specific example, so that the reasoning used for this specific example could generalize to any other example. Rowland suggested that formal proofs can be preceded by generic proofs or, perhaps, generic proofs can be given in lieu of formal proofs. Rowland reported that his students generally preferred generic proofs to formal proofs. In a subsequent qualitative study, Malek and Movshovitz-Hadar (2011) performed a preliminary assessment of generic proofs, asking students to try to prove a theorem, then giving these students a generic or linear proof of the theorem, and finally giving students a test on how well they comprehended the proof. Malek and Movshovitz-Hadar found that for non-routine proofs, students who read the generic version of the proofs performed better on the comprehension test than those who read the linear version. However, only three to four students read the generic proofs; hence, the generalizability of these results is limited.

Alcock (2009) suggested using "e-proofs": proofs presented in a computer environment that highlighted the logical connections between different parts of the proof and allowed users to request explanations about these connections. Alcock designed this type of proof presentation with the intention of improving students' proof comprehension (Alcock, 2009, 2010; Alcock & Inglis, 2010). However, a subsequent study found that students who read a proof in real analysis using e-

proofs did not perform better on an immediate post-test and a delayed post-test than students who read the same proof as normal text; the e-proofs students performed significantly worse on both immediate and delayed post-tests when compared with students who saw the same proof given in a lecture (Roy, Alcock, & Inglis, 2010).

Finally, some researchers have suggested using less rigorous but more intuitive arguments to support a theorem, such as arguments based on graphs or diagrams (e.g., Hersh, 1993). To our knowledge, empirical studies have not yet been conducted to assess the efficacy or drawbacks of this recommendation.

There are three themes in this review of the literature that are worth noting. First, empirical assessments of alternative formats for proof presentation are rare and only beginning to emerge. Second, the few studies conducted have not found generalizable evidence that the suggested proof formats actually improve comprehension[5]. Third, there is considerable variety in the methodologies used to assess the efficacy of alternative proof formats, with these formats being assessed in a lecture format (Rowland, 2001), a reading format (Roy et al, 2010), and reading after attempting to prove the theorem oneself (Malek & Movloshitz-Hadar, 2011). Assessments ranged from paper-and-pencil tests (Roy et al, 2010) to self-reported preferences (Rowland, 2010).

## 1. 3. Leron's structured proofs

### *1. 3. 1. Description of Leron's structured proofs*

Leron (1983) proposed a novel way to present proofs in terms of levels, where each level is an independent module of the proof. 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

[5] This does not imply that these alternative presentation formats lack pedagogical value; perhaps they benefit students in ways that were not assessed or would benefit students more if these formats were introduced differently. This raises the possibility that students' failures to learn from proof are due to other factors, such as poor proof reading strategies, or that having students read theorems and then proofs, without other activities such as exploring the meaning of the theorem or trying to prove the theorem oneself, is not conducive to learning. Note that in this paper, we are neither advocating nor discouraging the use of the definition-theorem-proof format commonly used in advanced mathematics courses, but are simply assessing a suggestion for making this format more effective.

details of the implementation of higher levels of the proof. An additional feature in some structured proofs is a section called an “elevator,” which is located between levels and provides a rationale for why the proof is proceeding the way that it is. Leron illustrated the nature of a structured proof by comparing a linear and structured proof of the claim: “There are infinitely many triadic primes” (where a triadic prime is a prime congruent to 3 modulo 4). We used these proofs in our studies and they are presented in Appendix A of this paper, nearly verbatim from Leron’s article. In this proof, Level 1 lays out a summary of the three main aims of the proof while the “elevator” provides a motivation for defining the variable *M* as it is defined in the proof. Leron noted that not all proofs were amenable to a structured proof format; the ideal candidates for these proofs are longer, complicated proofs whose structures are not transparent in a linear presentation (p. 176).

### *1. 3. 2. Theoretical benefits of structured proofs*

Leron argued that structured proofs possess several desirable properties. 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. In addition, the high-level structure and the “elevator” commentary explicate the reasoning behind some of the choices made in the proof that might otherwise seem arbitrary. Weber and Mejia-Ramos (2011) found that mathematicians claim to understand proofs in terms of their high-level ideas and methods (see also Mejia-Ramos & Weber, 2014). Structured proofs make these constructs explicit for the reader. We note that Leron’s suggestion seeks to aid comprehension in a different manner than the other alternative proof formats mentioned previously. Whereas generic proofs (Rowland, 2001) and informal proofs seek to reduce the abstraction in a formal proof by framing it in less intimidating representation systems (e.g., with examples or diagrams) and Alcock’s (2009) e-proofs aim to make explicit the justifications and connections in a proof that are usually tacit, Leron advocates helping students manage the complexity of formal proofs by making the structure of the proof apparent and motivating the ideas in it.

*1. 3. 3. Structured proofs as a pedagogical suggestion or a research-based claim*

Leron (1983) published his discussion of structured proofs in the *American Mathematical Monthly*, an expository mathematics journal, to share his ideas of how proofs might be presented better. He was clear that he did not advance his suggestion as an established research-based claim, writing "I do not know of any way to prove (or disprove) my claims on the merits of the structured methods" (p. 176). However, his pedagogical suggestion was nonetheless influential in the mathematics education community, with many citing Leron's structured proofs as a potential way to improve students' proof comprehension (e.g., Alibert & Thomas, 1991; CadwalladerOsker, 2011; Hanna, 1990; Hersh, 1993; Mamona-Downs & Downs, 2002; Movshovitz-Hadar, 1988; Selden & Selden, 2003, 2008).

Some researchers have gone further, touting structured proofs as a pedagogical suggestion supported by research or claiming that structured proofs are appropriately influencing the way that advanced mathematics courses are taught. 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) claimed that structured proofs are influencing the ways that 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 hypothesis but rather an established fact. Despite these claims, we are aware of no empirical studies that suggest structured proofs actually improve mathematical instruction or proof comprehension.

### 1. 4. Research questions

The studies in this paper are based on the premise that an important mechanism for growth in mathematics education is to examine the effects of promising pedagogical suggestions, where we follow Schoenfeld (1994) in

operationalizing a “promising pedagogical suggestion” as one evaluated favorably by many mathematics educators. As structured proofs are regarded highly by many mathematics educators and presented by some as a research-based suggestion, we are trying to measure the ways in which structured proofs improve student comprehension and to explore any difficulties that students may experience when reading them.

Of course, one cannot address questions such as “do structured proofs improve comprehension?” in a single series of studies, as the answer to this question likely depends on what constitutes comprehension, which proofs were structured, and how structured proofs were introduced to students, among other factors (see, e.g., Schoenfeld, 2000). In particular, whether a researcher observes the utility of an alternative proof format depends on whether his or her method for measuring comprehension is consistent with the intended benefits of the format. 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, students could be asked to study a structured proof overnight, or a professor could present a structured proof in lecture for his or her students. Further, perhaps using linear and structured proofs in tandem (e.g. having students structure linear proofs they are given) might yield greater learning benefits than using either in isolation. In this paper, we compare students’ comprehension of a proof shortly after reading a linear or structured version of the same proof. We contend that this choice is appropriate for three reasons. First, Leron (1983) asked readers of his pedagogical suggestion to “pretend to be a student reading these proofs for the first time” (p. 176), implying that he felt his suggestion would benefit students reading proofs. Second, textbook presentation is a natural way to implement structured proofs in the undergraduate mathematics curricula (see Mamona-Downs & Downs, 2002). Third, previous studies on the benefits of alternative formats for proof presentation (e.g., Malek & Movshovitz-Hadar, 2011; Roy et al, 2010) have also measured students’ comprehension shortly after they read these proofs in traditional and innovative formats. Using a similar procedure facilitates comparisons between our findings and those in the literature. However, it is certainly plausible that any learning benefits that we did not observe in these studies might occur if the structured proofs were used in a different way, or if students were given more experience with this format.

Based on our model of proof comprehension assessment (Mejia-Ramos et al., 2012), our primary assessment items evaluated students' comprehension in four ways:

(a) Can students recognize a good summary of the proof that they read?

(b) Do students recognize how the ideas used in the proof can transfer to another setting?

(c) Do students recognize how various claims within the proof are justified? and

(d) Can students apply the ideas in the proof that they read to a specific example or diagram?

There are other potential benefits of structured proofs that we did not assess in these studies. For instance, perhaps by reading structured proofs, students might develop a more accurate understanding of the proof writing process and a deeper appreciation for the enterprise of proof (as suggested by Selden & Selden, 2008).

The two specific research questions we investigate are:

- From the students' perspective, 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 answer assessment items of the four above types than students who read the same proof as a traditional linear proof?

Given the general limitations of our study, as well as the difficulty of addressing these types of conceptual questions, we do not view the results of our study as conclusive (if conclusive answers to these types of questions are even possible to obtain). Rather, we view these studies as an important first step to examining the benefits and drawbacks of using structured proofs.

## 2. STUDY 1: QUALITATIVE INTERVIEW STUDY

### 2. 1. Rationale for the qualitative study

Study 1 was an interview study in which students were videotaped while reading 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 to examine how students read and reacted to structured proofs, and to gain insight into how this type of proof presentation might help or hinder students' comprehension of a proof.

## 2. 2. Methods

### *2. 2. 1. Participants*

Students were invited to participate in the study from a large state university in the northeast United States. We initially recruited a group of 12 students to participate in a pilot study in the beginning of the Fall 2010 semester. These students were recruited from a transition-to-proof course, an introductory real analysis course, and a mathematics education course for mathematics majors who were prospective secondary mathematics teachers. These courses were primarily composed of second and third-year mathematics majors. These 12 students were divided evenly into two groups, each group looking at one of two structured proofs and another linear proof (we refer to these groups as Pilot Group A and Pilot Group B). We discuss the materials and the treatment given to these groups in the following sections.

After interviewing the students in this pilot study, it became clear to us that some of the students were deeply confused about the nature of structured proofs. We believed one cause of their difficulty was that they were not given an adequate description of the nature and purposes of structured proofs. Furthermore, the linear proofs did not generate data of significance to our study. Consequently, we viewed data from these students as pilot data. At the end of the Fall 2010 semester, we recruited six more students from another section of the introductory real analysis course and conducted the study again (we refer to these students as the Interview Group). This time we asked each student to read both structured proofs, but only after giving them a description about the nature of a structured proof. Our analysis focuses primarily on the six students in the Interview Group, but we also discuss the

students from the Pilot Groups to illustrate how they corroborate the themes we saw from the Interview Group.

### *2. 2. 2. Materials*

Two proofs were used for this study. We call the first proof the Only Zero proof, which was a calculus-based proof of the claim that the only solution to the equation $x^3 + 5x = 3x^2 + \sin x$ is $x = 0$. The linear and structured Only Zero proofs are presented in Appendix B. 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 Triadic Primes proofs were taken nearly verbatim from Leron's (1983) original article on structured proofs[6] and are presented in Appendix A.

### *2. 2. 3. Procedure*

Each participant met individually with an author of this paper for a video-recorded semi-structured interview. Participants were given a version of one of the proofs 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 judge, on a scale of one through five, how well they felt they understood the proof. They were then asked a set of open-ended questions about the proof, assessing the four aspects of proof comprehension mentioned earlier. Participants were given a sheet of paper with each question written on it so they would not have to keep in mind the question while trying to remember the relevant details of the proof. After participants answered these questions, they were given a set of multiple-choice questions, which were similar to the questions they had just answered. After answering all the

[6] We chose the Triadic Primes proof because this would clearly be consistent with Leron's (1983) intention. However, many who have heard presentations of this research and two reviewers of this manuscript note that the high-level structure of the Triadic Primes proof seemed needlessly complex. Consequently, any difficulties understanding this structured proof could plausibly be attributed to how it was written.

questions, the interviewer returned the proof to the participants and invited them to change any of their answers as they saw fit. Participants usually did not change any of their answers. If participants read the structured version of the proof, they were asked how they felt about this new format, whether they liked it, and what they thought the strengths and weaknesses of this format were. This entire process was then repeated with the other proof.

Pilot Group A received a linear version of the Only Zero proof and a structured version of the Triadic Primes proof. Pilot Group B received a structured version of the Only Zero proof and a linear version of the Triadic Primes proof. The Interview Group first read a description describing the nature of structured proofs (this is presented in Appendix C), and then read structured versions of both the Only Zero and Triadic Primes proofs.

#### *2. 2. 4. Analysis*

In our analysis, we sought to identify attributes of structured proofs that may have aided or hindered the participants' comprehension of the proof. We made an initial pass through the data noting: (a) any positive comments that participants made about the structured proofs that they read, (b) any moments of insights that participants expressed while reading the structured proofs, (c) any negative comments that participants made about the structured proofs that they read, and (d) any confusion expressed by the participants while reading the structured proofs or answering questions about the proofs. During this initial pass through the data, we noted three themes that emerged from our data: (1) some participants valued the summaries provided in Level 1 of the structured proofs, while others had difficulty understanding these, (2) some participants valued that structured proofs provided insight into the thought processes used to create the proof, and (3) most participants complained that structured proofs were difficult to follow because they "jumped around". We then systematically re-analyzed the data, coding for any instances of these phenomena.

### 2. 3. Results

A summary of the results is presented in Table 1.

Table 1: Number of participants in the qualitative study expressing each positive or negative opinion about the structured proof.

| | GROUPS | | |
|---|---|---|---|
| Category | Interview Group (N=6) | Pilot Group A (N=6) | Pilot Group B (N=6) |
| Summaries in Level I | | | |
| Appreciated the summaries | 3 | 0 | 2 |
| Did not like the summaries | 2 | 2 | 0 |
| Explaining reasoning behind the proof | | | |
| Appreciated that this was done | 2 | 2 | 1 |
| Did not like that this was done | 2 | 1 | 0 |
| Thought proof jumped around too much | 5 | 4 | 5 |

### *2. 3. 1. Summaries*

Level 1 of a structured proof is essentially a summary of the main ideas of the proof. Three participants in the Interview Group explicitly mentioned the value of providing this summary, as it provided them with a framework to interpret the rest of the proof. For instance, consider the comment below:

> *IG 6 [7]: I actually like how this proof is written out. It's really step-by-step. Because first they said in level one, this is what we're going to attempt to do. So he said first we're going to prove that zero is a solution, and that's kind of clear, you can just plug it in. And then they said now we want to show that the function is increasing for all x, and they do that here (pointing to 2b on the structured Only Zero proof).*

Later, this student commented that this made the proof easier to follow:

> *IG 6: I like how he took the different steps and made it really easy to follow. I don't think you really need to be a math major to kind of follow that.*

Two students in Pilot Group B, who read the structured version of the Only Zero proof, also complimented the summary in Level 1. One student described this as, "Okay, there's three goals in this proof… it's like stating objectives. It clearly

[7] IG 6 stands for the sixth interviewed student in the Interview Group.

states at the beginning, we're going to prove this using these three goals. Bang, bang, bang". It is interesting to note that no student in Pilot Group A, who read a structured version of the more challenging Triadic Primes proof, spoke favorably of the summary in Level 1 of this proof.

Two participants in the Interview Group did not appreciate the summary in Level 1 because they did not understand it. Consider the comment by an Interview Group participant:

> *IG 1: [Referring to Level 1 of the Triadic Primes Proof] This by itself doesn't make a lot of sense to me. I couldn't figure out what this even meant or why they were doing it, where M came from, none of it. So when I read this [pointing to the rest of the proof], it made more sense […] If you put Level 2 in it, it makes so much more sense.*

One intended benefit of structured proofs is that the high-level summary of Level 1 enables the reader to more fully appreciate and comprehend the logical justifications that appear in the lower levels of the proof. However, for this participant, the levels had the opposite effect. He needed to read Level 2 to make sense of the summary in Level 1. Two participants from Pilot Group A, who read a structured version of the Triadic Primes proof, made similar complaints.

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 struggle to remember the ideas contained in Level 1, leading to a strain on their working memory and a dissatisfaction that the proof was jumping around (discussed further in the next sub-section).

We propose the following hypothesis to account to account for this data: If participants are familiar with the ideas and methods outlined in Level 1 of a structured proof, reading Level 1 may provide them with a useful framework to understand the logical details that follow. However, if the ideas and methods outlined in Level 1 are novel to students, the summary here is more likely to confuse students and fail to provide a useful framework for integrating information. This could account for the fact that the summaries for Level 1 were more valued by participants reading the structured version of the Only Zero proof, where the calculus-based techniques

were presumably familiar to students, than to those reading the structured version of the Triadic Primes proof, where the techniques were presumably less familiar to students. The following excerpt below provides evidence that the familiarity with the technique of the proof was key to understanding Level 1:

> *IG 2: [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.*

This participant 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.

### *2. 3. 2. The reasoning behind the proofs*

Two participants from the Interview Group valued structured proofs because they made clear the reasoning behind the proofs. Both participants mentioned the "between levels" step in the Triadic Proof. For example:

> *IG 6: The between steps really kind of breaks down what the thought process is when defining that M. Like why we chose it, what we would normally choose, why that isn't a good one, and how we lead up to what that number is going to be.*

Three participants from the Pilot Groups expressed a similar appreciation. However, we note that the "between levels" was not universally appreciated by all participants. One participant in the Interview Group was confused about why the "natural" choice for $M$ that was *not* chosen in the Triadic Proof was natural, saying, "It's not really clear because he says, 'It is natural try $M = 4p_1p_2...p_n + 1$'. I don't

understand why it's natural to choose that". Another participant from the Interview Group did not seem to appreciate the distinction between the motivation contained in the "between levels" and the justifications present in the lower levels of the proof, saying, "if you want to make levels, then why are you having this between levels thing? Can't you just put it in a level? Isn't that the whole point?" One participant from Pilot Group A made a similar criticism, claiming that the structured Triadic Primes proof "included things it just didn't need", citing the between levels because these make the argument "sound like a story, not a proof". This suggests that this participant (and perhaps others) do not view providing motivation for where the proof idea came from as appropriate to include in a proof, suggesting that these types of students may need to refine their epistemological beliefs about the purposes of proof if they are to appreciate this aspect of structured proofs.

### *2. 3. 3. Jumping around*

Five students in the Interview Group complained that the structured proofs "jumped around" too much. That is, that the proof was difficult to follow because ideas that were closely related in the proof appeared spatially far apart from one another:

> *Interviewer: What did you think of the format?*
>
> *IG 1: 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].*
>
> *IG 3: 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 a characteristic feature of structured proofs, but students' comments reveal a drawback 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. Comments of this type were made by students in the Pilot Groups as well; nine of the 12 Pilot Group students complained that the structured proofs jumped around.

### 2. 4. Summary and discussion

Three participants in the Interview Group appreciated the summary given in Level 1 of the structured proofs. However, our data suggest that these high-level summaries are most effective in cases where the participants are familiar with the proof techniques being employed. If the Level 1 summary presents ideas with which the students are not familiar, students might not gain insight from reading it. Indeed, some participants in our study indicated that they only understood Level 1 of the Triadic Proofs after reading Level 2 and Level 3.

Two participants in the Interview Group liked how structured proofs provided the motivation behind the proof, but some participants in this study did not value this, in some cases because they did not understand the motivation upon reading it and in one case because they thought such motivation did not belong in a proof. While structured proofs arguably have the potential to transform students' beliefs about the enterprise of proof (as suggested by Selden and Selden, 2008), students' pre-existing beliefs about what proofs should contain may initially inhibit their comprehension of these proofs.

Most participants in the Interview Group, as well as both Pilot Groups, expressed frustration that structured proofs jumped around. It is not clear to us how structured proofs could be altered to address this issue, although perhaps this feature of the proof would be less prevalent in a lecture or after students gained experience reading structured proofs. In general, these findings should all be viewed tentatively given the small number of participants in the study and the fact that students had so little experience with reading structured proofs.

## 3. STUDY 2: QUANTITATIVE INTERNET STUDY

### 3. 1. Research methods

To obtain a large sample size, we conducted the quantitative study with mathematics majors using an internet-based instrument. The validity of this type of method has been discussed extensively in the research methods literature, with the conclusion that internet-based research is as valid as laboratory research—provided that appropriate safeguards are taken (e.g. Gosling et al., 2004; Reips, 2000). To deal with the common validity threats for this type of study, we employed the methodology of Inglis and Mejia-Ramos (2009) and discarded participants who showed evidence of having participated in the experiment more than once (we controlled for multiple submissions by recording participants' IP addresses) or who revisited earlier pages in the study (we recorded the order in which each participant visited the pages in the experimental website).

*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. Consequently, the student population for this study were advanced mathematics majors and minors from elite universities in the United States.

*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)[8]. We note that, in the qualitative study, we found evidence that the Only Zero proof contained routine methods for students while the methods in the Triadic Primes proof were novel for undergraduate mathematics majors. 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. These items were refinements of those used in the qualitative study.

[8] Minor modifications were added to these proofs to clarify difficulties in interpretation that participants from the qualitative study had that were unrelated to the format of the proof.

*Procedure.* Depending upon their university, participants were either 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: *L*, where participants were presented with the linear version of the proof; *S(no desc),* where participants were presented with the structured version of the proof, but with no description about the nature of structured proofs; and *S(w desc)*, where participants were presented with the structured proof after reading a description of the nature and purpose of structured proofs. These descriptions are the same as those used in the qualitative interview study and are provided in Appendix C.

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 answered the corresponding set of comprehension questions. Each question appeared on a new screen and the order in which they were displayed was randomized for each participant. Participants were asked not to hit the "Back" button to review the proof or change their answers to any of the questions, and informed that if they did their data would not be considered for analysis. Before analyzing the data, we removed any participant that revisited pages while completing the study. Upon completion of the test, participants in either of the two structured conditions 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.

### 3. 2. Results

#### *3. 2. 1. Comprehension test*

Participants' performance in the comprehension tests is presented in Table 2. We note that we found no statistically significant differences between the performance on the items between the *S(no desc)* and *S(w desc)* conditions and, surprisingly, the *S(no desc)* did better than the *S(w desc)* conditions on all but one of the items. Consequently, for the purposes of comparing the efficacy of structured proofs, we collapsed the *S(no desc)* and *S(w desc)* groups into a single Structured group. We note that the lack of difference between *S(no desc)* and *S(w desc)* is inconsistent with the pilot data from Study 1, where we found students in the Pilot

Groups had difficulty understanding structured proofs because they were not given a description of them. It is possible that our assumption that this was a cause of difficulty for the Pilot Group was unfounded.

Table 2: Percentage of students from each group of the quantitative study who correctly answer each assessment item.

*Only Zero proof*

| Condition | Summary | Transfer | Justification Question 1 | Justification Question 2 | Application to Diagram |
|---|---|---|---|---|---|
| Lin (N=67) | 63% | 61% | 85% | 85% | 69% |
| S(no Desc) (N=69) | 70% | 54% | 90% | 87% | 78% |
| S(w Desc) (N= 66) | 83% | 51% | 86% | 73% | 77% |

*Triadic Primes proof*

| Condition | Summary | Transfer | Justification | Application to Examples |
|---|---|---|---|---|
| Lin (N=33) | 67% | 45% | 54% | 91% |
| S(no Desc) (N=32) | 72% | 44% | 47% | 65% |
| S(w Desc) (N= 33) | 75% | 33% | 33% | 60% |

Collapsing across both structured conditions

*Only Zero proof*

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

*Triadic Primes proof*

| Condition | Summary | Transfer | Justification | Application to Examples |
|---|---|---|---|---|
| Lin (N=33) | 67% | 45% | 54% | 91%* |
| Struct (N=65) | 74% | 38% | 40% | 63%* |

*- Indicates a statistically significant difference between groups with $p < .05$

The data indicate that the Structured Group performed better on the summary questions than the Linear Group did. For the Only Zero proof, the Structured group answered the summary question correctly 76% of the time as compared to 63% for the Linear group, a statistically significant difference ($\chi(1, 200)=4.095$, $p=.043$)[9]. For the Triadic Primes proof, the Structured group also did better on the summary question, although this difference (74% for the Structured group compared to 67% for the Linear group) was not statistically significant ($\chi$ (1, 96)<1, p=.457). This data is consistent with the qualitative data from Study 1. Participants appeared to benefit from reading Level 1 of the structured Only Zero proof. However, participants showed

[9] For all statistical tests, the null hypothesis is that there is no difference between how the Structured and Linear Groups would perform on that specific assessment item.

a more modest benefit from reading Level 1 of the structured Triadic Primes proof. This is consistent with our hypothesis that students might gain understanding from reading a summary primarily when they are familiar with the ideas and techniques in this summary.

The Structured Group did not perform better than the Linear Group for the other assessment items. In fact, for five of these seven assessment items, the Linear Group outperformed the Structured Group, although these differences were generally not statistically significant. The Linear group had a higher percentage of students who answered the transfer questions correctly for both the Only Zero (61% vs 53%) and Triadic Primes (45% vs 38%) proofs. However, neither of these results were statistically significant ($\chi$ (1, 200)=1.341, p=0.247, $\chi$ (1, 96)<1, p=0.506). For the justification questions, we found effectively no difference between participants' performance for the Only Zero proof. Averaging across the two justification questions, the Linear and Structured groups performed essentially the same (85% and 84% correct, respectively). For the Triadic Primes proof, the Linear group performed somewhat better than the Structured group on the justification question (54% vs. 40%), but this difference was not statistically significant ($\chi$ (1, 96)=1.872, p=0.173). Hence our results for these assessment items are inconclusive, but at least provide no evidence that structuring a proof improves students' comprehension in these regards.

On the Only Zero proof, the Structured group performed better than the Linear group in recognizing that the ideas in the argument would not be applicable to the function shown in the diagram (78% vs. 69%), although this difference was not statistically significant ($\chi$ (1, 200)=1.974, p=.160). However, for the Triadic Primes proof, the Linear group did much better on the question applying the ideas of the proof to a specific example (91% vs. 63%; Fisher exact test, p=.034). It is plausible that this result is partly due to the fact that once *M* is chosen in the Triadic Primes proof, the linear version of the proof essentially offers a step-by-step prescription 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 the *M*. Reading a linear proof may help students with

these types of questions because the step-by-step procedure is made explicit. Of course, we would need to replicate this result with more proofs to test this conjecture.

*3. 2. 2. Participants' subjective evaluation of structured proofs*

After completing the assessment items, participants who read a structured proof were asked if they had a favorable or unfavorable view of the proof format. Participants who did not receive the structured proof description tended to have an unfavorable opinion of the proofs (20 viewed the proof favorably, 37 unfavorably, 44 neutrally), while those who did receive the description viewed the proof more positively (33 favorably, 21 unfavorably, 45 neutrally).

Participants were also asked to rate how well they understood the proof on a scale of 1 through 5. Participants' self-reported levels of understanding were similar across groups. For the Only Zero proof, the Linear Group, Structured with Descriptions, and Structured without Descriptions, participants said they understood the proof fully (i.e., provided a rating of 5) 72%, 75%, and 74% of the time respectively. For the Triadic Primes proofs, these percentages were 52%, 52%, and 48%.

## 4. DISCUSSION

### 4. 1. Implications for structured proofs

As Schoenfeld (2000) cautioned, questions such as "are structured proofs a useful pedagogical technique?" are impossible to answer in a single study. Therefore, it is important to situate these results carefully and to avoid drawing inappropriate conclusions. The central result from Study 2 was that structured proofs helped students identify a high-level summary of the proof, but there was no evidence that they improved students' performance on assessment items related to justification, transfer, or applying the ideas of the proof to a specific example.

There are several essential caveats that must be kept in mind when interpreting this result. First, we assessed specific aspects of proof comprehension

after students read a proof in a short period of time. There are numerous ways we could have changed the design of our studies; for instance, we could have presented the structured proofs in lectures instead of written text, or we could have assessed students' appreciation of proof or their abilities to write proofs rather than their comprehension of proofs. We could also have looked at using linear and structured proofs in conjunction. Had we made any of these changes, the outcome for structured proofs may have been more favorable. Second, our study looked for *immediate and short-term learning gains* for structured proofs. This is perhaps not a realistic expectation since undergraduates have extensive experience with reading linear proofs but essentially no experience with reading structured proofs. It is possible that the benefits of structured proofs would be more apparent if students had more experience with proofs in this format or training on how these proofs should be read and constructed. Third, we did not differentiate between different types of advanced mathematics majors for the purposes of this study. It is possible that structured proofs may have been beneficial for particular groups of students (e.g., new math majors who have not grown accustomed to linear proofs). This would be an interesting topic for future research.

These findings do suggest that simply changing the format of proof from the traditional linear proofs to structured proofs is not likely to substantially improve students' comprehension of these proofs, at least not in the short term. We believe this has two implications. First, if students are to benefit from structured proofs, they will probably require different pedagogical and mathematical experiences than the relatively shallow description of structured proofs that we provided students in this study. In hindsight, this perhaps is obvious. However, we note that Leron's (1983) article and the many articles endorsing structured proofs (e.g., Mamona-Downs & Downs, 2002; Selden & Selden, 2008) simply recommend the use of structured proofs as a way to improve instruction, but do not suggest specific instructional interventions that mathematics professors might use for structured proofs to be effective. The data for Study 1 suggests several such measures. One finding from Study 1 is that participants appreciated structured proofs more when they were familiar with the overarching method of the proof, as this allowed Level 1 of the proof to be accessible and serve as an organizing guide. This finding suggests that the potential benefit of structured proofs depends not only on the student or on the

content of the proof, but on the relationship between the two. Therefore we recommend that professors take care when presenting a summary of a proof (i.e. Level 1 of a structured proof), especially when the ideas in this summary are less familiar to students. When introducing novel and difficult methods, it perhaps would be better to first present these proofs linearly and only afterwards summarize the new methods. Second, it might be worthwhile for professors to discuss the epistemological role of proof with students in their lectures: proofs are not only meant to give conviction, but also to provide explanation, insight, and ways of reasoning. Doing so might help students better appreciate how the summary and between-levels sections of structured proofs illustrate and motivate the ideas of the proof.

These results also underscore the lack of empirical evidence supporting the claim that structured proofs actually improve students' comprehension, or benefit students in any other discernable way. We emphasize that our results certainly do *not* imply that structured proofs cannot be beneficial; indeed, perhaps assessing students' proof comprehension of a structured proof with such limited experience with the format is an unfair test. However, the central results from Study 1 suggest that structured proofs may have some inherent weaknesses for student comprehension: primarily, students perceive these proofs as "jumping around", making it difficult to follow the sub-arguments. Therefore, we would recommend that the field not take it as *a priori* that structured proofs improve comprehension (as some, such as Melis, 1994, have done). Further, other studies have found that other promising alternative proof formats had a negative effect on student learning (e.g., Roy et al, 2009). It might be the case that students' difficulties are not due to the presentation format, but rather, as Hanna (1990) suggests, to the content of the proof. In this case, proofs that employ explanatory arguments might be more comprehensible to students. Hence we recommend more research on the effectiveness of structured proofs before as advancing it as an effective pedagogical suggestion.

### 4. 2. Research on alternative proof formats

This research study failed to verify the effectiveness of a popular pedagogical suggestion in mathematics education. Consequently, we are not surprised that some view these findings as provocative. This article should not be read as a critique of

Leron's (1983) original contribution, which made no research claims and successfully opened a conversation about proof presentation in mathematics teaching. Rather, it should be viewed as an initial attempt to address the questions of when and in what ways structured proofs might benefit students.

We note that our presentations of these findings have generated a wide range of reactions from the mathematics education community. We believe that this is because the claim "structured proofs improve student learning in advanced mathematics" is not a simple assertion, but actually a wide range of claims. Mathematics educators have very different interpretations of what this claim means, including how structured proofs are used in the classroom and the goals of presenting proofs to students (e.g., proof comprehension vs. ability to construct proofs). As noted in the introduction, assessments for proof comprehension and research on proof alternatives are beginning to emerge (e.g., Malek & Movshovitz-Hadar, 2011; Mejia-Ramos et al, 2012; Roy et al, 2009). An important first step to building a systematic research base in this area is clarifying exactly what it means for an alternative proof format to be effective. With this clarification, comparisons between alternative proof formats and traditional proofs, as well as comparisons between different alternative proof formats, can yield a research base that is both cumulative and generalizable.

## 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* (pp. 227-245). Springer: Dordrecht.

Alcock, L. & Inglis, M. (2010). Visual considerations in the presentation of mathematical proofs. *Seminar.net- International Journal of Media, Technology, and Lifelong Learning*, *6*, 43-59.

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

Anderson, J., Boyle, C., & Yost, G. (1986). The geometry tutor. *Journal of Mathematical Behavior*, *5*, 5-20.

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

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.

Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics, 61*, 103–131

Fuller, E., Mejia-Ramos, J.P., Weber, K., Samkoff, A., Rhoads, K., Doongaji, D., & Lew, K. (2011). In S. Brown, S. Larsen, K. Marrongelle, & M. Oerterman (Eds.)Comprehending Leron's structured proofs. In *Proceedings of the 14th Conference for Research in Undergraduate Mathematics Education*, 1, 84-102. Portland, OR.

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

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

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

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

Harel, G. & Sowder L. (1998). Students' proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput, and E. Dubinsky (Eds.), *Research in Collegiate Mathematics Education III* (pp. 234-283). Providence, RI: American Mathematical Society.

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.

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

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

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

Malek, A. & Movshovitz-Hadar, N. (2011). The effect of using Transparent Pseudo-Proofs in linear algebra. *Research in Mathematics Education*, *13*, 33-58.

Mamona-Downs, J., & Downs, M. (2002). Advanced mathematical thinking with a special reference to mathematical structure. In L. English (Ed.), *Handbook of international research in mathematics education* (pp. 165-196). Lawrence Erlbaum Associates: Mahwah, NJ.

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). Why and how mathematicians read proofs: Further evidence from a survey study. *Educational Studies in Mathematics*, 85, 161-173.

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.

Movshovitz-Hadar, N. (1988). Stimulating presentation of theorems followed by responsive proofs. *For the Learning of Mathematics*, *8*(2), 12-19.

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. (1994). What do we know about mathematics curricula? *Journal of Mathematical Behavior*, 13, 55-80.

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.

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

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

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

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

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. (2014). Mathematics majors' beliefs about proof reading. *International Journal of Mathematics Education in Science and Technology*, 45, 89-103.

Submetido: Março 2014
Aceito: Maio 2014

Appendix A: Triadic Primes proof and questions (final version from quantitative internet study)

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

**Claim.** There exist infinitely many triadic primes.
*Linear Proof:*

1. Consider a product of two monadic numbers:
   $(4j+1)(4k+1) = 4j\,4k + 4j + 4k + 1 = 4(4jk + j + k) + 1,$
   which is again monadic.
2. Similarly, the product of any number of monadic numbers is monadic.
3. Now, assume the theorem is false, so there are only finitely many triadic primes, say $p_1, p_2, ..., p_n$.
4. Let $M = 4p_2 \cdots p_n + 3$, where $p_1 = 3$.
5. $p_2, p_3 ..., p_n$ do not divide *M* as they leave a remainder of 3, and 3 does not divide *M* as it does not divide $4p_2 \cdots p_n$.
6. We conclude that no triadic prime divides *M*.
7. Also, 2 does not divide *M* since *M* is odd.
8. Thus all of *M*'s prime factors are monadic, hence *M* itself must be monadic.
9. But *M* is clearly triadic, a contradiction.

*Structured Proof:*

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

(a) *M* as well as all its factors are different from $p_1, p_2, ..., p_n$;

(b) *M* has a triadic prime factor.

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

**Between Levels:** How shall we define *M*? It is natural to try $M = 4p_1p_2 \cdots p_n + 1$, which meets requirement (a) but not (b). In fact, since *M* itself may turn out to be prime, it must be triadic to meet requirement (b). A natural second guess is $M = 4p_1p_2 \cdots p_n + 3$. However, since $p_1 = 3$, *M* is divisible by 3, in violation of (a). This 'bug' is easy to fix: simply eliminate 3 (i.e. $p_1$) from the product in our definition of *M*.

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

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

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

**Level 3.** *Lemma:* The product of monadic numbers is again a monadic number.

Consider a product of two monadic numbers:
$(4j+1)(4k+1) = 4j4k + 4j + 4k + 1 = 4(4jk + j + k) + 1,$
which is again monadic. Similarly, the product of any number of monadic numbers is monadic.

*Questions:*

**Summary**
Which of the following is a better summary of this proof, A or B?

A. It lists all triadic primes, then uses this finite list of triadic primes to obtain a contradiction by constructing a number that is triadic but has no triadic prime factors.
B. It lets $M = 4p_2 \cdots p_n + 3$, where $p_i \neq 3$. 2 does not divide *M* because *M* is odd. $p_i$ does not divide *M* because it leaves a remainder of 3. This produces a contradiction.
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 primes of the form $6j+5$, what would be an appropriate definition for *M*? Assume first that there are finitely many primes of this form, say $p_1, p_2, ..., p_n$, with $p_1 < p_2 < ... < p_n$.

**Justification**
How was the conclusion that "the product of any number of monadic numbers 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.

**Illustration with examples/diagrams**
Supposing 3, 7, 11, and 19 are the only triadic primes, which of the following 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 they leave a remainder of 3, and 3 does not divide *M* since it does not divide $4 \cdot 7 \cdot 11 \cdot 19$. Thus, *M* has only monadic factors and is monadic.
B. Let $M = 4 \cdot 7 \cdot 11 \cdot 19 + 3 = 5855$. Then the prime factors of *M* are 5 and 1171. 5 is monadic and 1171 is triadic. Thus, *M* is triadic.
C. Neither of the above illustrates the main steps of the proof.
D. I don't know.

Appendix B: Only Zero proofs and questions (final version from quantitative internet study)

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

*Linear Proof:*

1. Let $f(x) = x^3 - 3x^2 + 5x - \sin x$. Solutions of $f(x) = 0$ precisely correspond to solutions of $x^3 + 5x = 3x^2 + \sin x$.
2. Suppose the claim is false. Then $f(x) = 0$ has a nonzero solution *s*; that is, $s \neq 0$ and $f(s) = 0$.
3. $f'(x) = 3x^2 - 6x + 5 - \cos x = 3(x^2 - 2x + 1) + 2 - \cos x = 3(x-1)^2 + 2 - \cos x$.
4. 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*.
5. Clearly $f(0) = 0^3 - 3(0)^2 + 5(0) - \sin 0 = 0$, so $x = 0$ is a solution of $f(x) = 0$.
6. Since $f(0) = f(s)$ and $s \neq 0$, by Rolle's theorem, there is a *c* between 0 and *s* such that $f'(c) = 0$.
7. However, this is a contradiction because $f'(x) > 0$ for all *x*.

**Note:** Rolle's theorem states that if a differentiable function *f* has the property that $f(a) = f(b)$, then there is a *c* such that $a < c < b$ and $f'(c) = 0$.

*Structured Proof:*

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

- $f'(x) > 0$ for all *x*.
- $f(s) = 0$ for $s \neq 0$ and $f(0) = 0$ imply that there is a number *c* for which $f'(c) = 0$.

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

**Level 2.**
**2a.** $f'(x) = 3x^2 - 6x + 5 - \cos x$. Using algebra (in level 3a), we show this expression is always positive.
**2b.** Suppose $s \neq 0$ and $f(s) = 0$. Since $f(0) = 0$ (level 3b), this implies there is a *c* such that $f'(c) = 0$, contradicting the fact that $f'(x) > 0$ for all *x*, which was established in level 2a. (The details are given in level 3c).

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

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

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*.

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

**3c.** We support the claim in level 2b as follows:

Rolle's theorem states that if a differentiable function *f* has the property that $f(a) = f(b)$, then there is a *c* such that $a < c < b$ and $f'(c) = 0$. In our case, we have $f(0) = f(s) = 0$. Hence there is a *c* between 0 and *s* such that $f'(c) = 0$.

*Questions*

**Summary**

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 a non-zero *s* for which $f(s) = 0$, then there would be a point where $f'(x) = 0$. However $f'(x)$ is always positive. Hence 0 is the only solution.

B. Since $f(x) = x^3 - 3x^2 + 5x - \sin x$, then $f'(x) = 3x^2 - 6x + 5 - \cos x = 3(x-1)^2 + 2 - \cos x$. If *s* is a solution, $f(s) = 0$. A non-zero solution would mean there is a *c* such that $f'(c) = 0$, which is a contradiction.

C. I don't know which summary would be better.

**Transfer**

Why is it not possible to use the ideas from the proof to show that the equation $x = \sin x$ has no nonzero solutions?

A. To use Rolle's theorem, the function $f(x) = x - \sin x$ would need to be 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 1**

Where was the fact that $f(0) = 0$ used in this proof?

I. It was used to show that 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 contradiction.

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 and II.

E. I don't know.

**Justification 2**

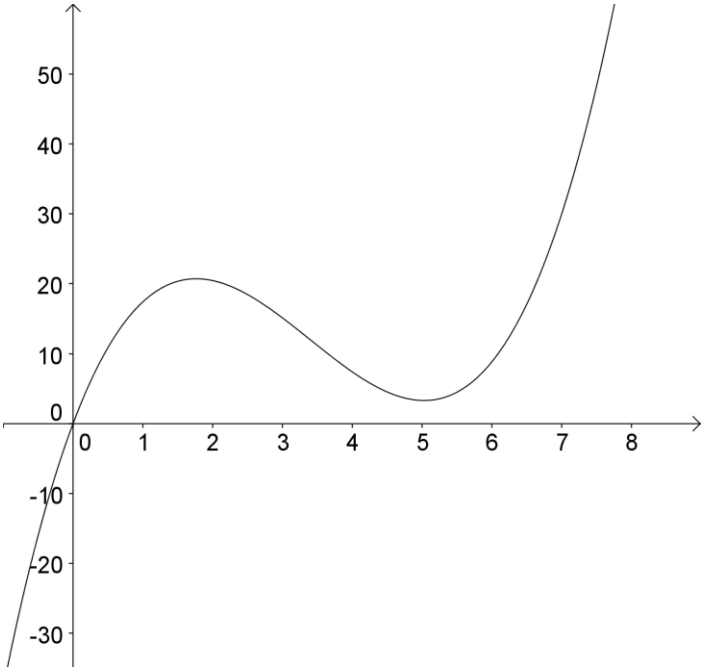
How was the fact that $f'(x) > 0$ for all real numbers *x* used in this proof?

A. It was used to show that $f(s) = 0$.

B. It was used to show that solutions of $f(x)=0$ precisely correspond to 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.

**Illustration with examples/diagrams**

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




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

Appendix C: Description for reading structured proofs

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.