An assessment model for proof comprehension in undergraduate mathematics.

Juan Pablo Mejia-Ramos [1]

Evan Fuller [2]

Keith Weber [1]

Kathryn Rhoads [1]

Aron Samkoff [1]

[1] Graduate School of Education, Rutgers University

[2] Department of Mathematical Sciences, Montclair State University

Abstract

Although proof comprehension is fundamental in advanced undergraduate mathematics courses, there has been limited research on what it means to understand a mathematical proof at this level and how such understanding can be assessed. In this paper, we address these issues by presenting a multi-dimensional model for assessing proof comprehension in undergraduate mathematics. Building on Yang and Lin's (2008) model of reading comprehension of proofs in school geometry, we contend that in undergraduate mathematics a proof is not only understood in terms of the meaning, logical status, and logical chaining of its statements, but also in terms of the proof's high-level ideas, its main components or modules, the methods it employs, and how it relates to specific examples. We illustrate how each of these types of understanding can be assessed in the context of a proof in number theory.

Keywords: Proof comprehension, proof reading, assessment, undergraduate mathematics education.

An assessment model for proof comprehension in undergraduate mathematics

1. Introduction

In advanced mathematics courses, students spend a substantial amount of time reading proofs. They read proofs in their mathematics textbooks and their professors' lecture notes and they read and listen to the proofs their professors present in class. Presumably a central reason that students are expected to read and study proofs is that they can come to understand the proofs and learn from them. However, the extent to which this pedagogical goal is realized is largely unknown and we contend that this is due, in part, to the lack of assessment instruments on proof comprehension.

It is widely accepted that the purpose of proof, both in the mathematics community and mathematics classrooms, is not merely to convince students that an assertion is true, but also to provide students with some form of mathematical insight (e.g., Hanna, 1990; Hersh, 1993; Thurston, 1994). However, exactly what this insight is, what it means for a proof to be understood, and how we can tell if students comprehend a given proof remain open questions in mathematics education. In a systematic study of the literature, Mejia-Ramos and Inglis (2009) found that in a sample of 131 articles related to the notions of proof and argumentation in mathematics, only three articles focused on students' comprehension of given proofs. This finding is consistent with calls from other researchers (e.g., Mamona-Downs & Downs, 2005; Selden & Selden, 2003) who have suggested that more research on proof reading is needed. Furthermore, Conradie and Frith (2000), Rowland (2001), Schoenfeld (1988), and Weber (in press) have argued that students' comprehension of a given proof is often measured by asking them to reproduce it or modify it slightly to prove an analogous theorem, even though these types of assessments offer only a superficial view of students' comprehension. These findings suggest that more sophisticated ways of assessing students' comprehension of a proof are needed. The objective of this paper is to present a model for assessing proof comprehension in advanced mathematics.

We argue that our assessment model for proof comprehension in advanced mathematics is particularly useful for researchers in undergraduate mathematics education, but could also be helpful for professors who teach advanced mathematics courses.

For researchers in mathematics education, a means to assess students' comprehension of proofs could be important for evaluating the effectiveness of mathematics instruction. For instance, researchers often cite structured and generic proofs (e.g., Leron, 1983; Rowland, 2001) as having the potential to improve how well students understand particular proofs (as well as their understanding of the enterprise of proof in general), but these broad claims have not been empirically tested. Our assessment model could be used to study a related, more specific question: how and to what extent do these novel ways of presenting proofs improve proof comprehension? More generally, our model could be used to examine how proof comprehension develops in students and to evaluate different means of improving it.

Employing assessment instruments based on this model could also inform teachers what *specific aspects* of a given proof students understand and what aspects they do not understand. As Conradie and Frith (2000) argued, appropriate comprehension tests can provide teachers with more insight into how effective their lectures were, and possibly how to improve them. Teachers of advanced mathematics might also employ this model to design assessment instruments that convey to students the type of understanding they are expected to develop. As Resnick and Resnick (1992) argued, it is usually unrealistic to expect students to learn a piece of mathematics if their understanding of this mathematics is not assessed. If comprehension tests only ask students to reproduce a proof by rote, students are likely to develop a superficial understanding of that proof and emphasize form over substance (as is illustrated by Schoenfeld, 1988). A useful assessment instrument could highlight to students what they are supposed to understand and direct their attention to appropriate aspects of the proof.

2. Related literature

In the last two decades, there has been a tremendous increase in research on proof. However, the substantial majority of the empirical studies on proof have focused on students' construction of proofs (Mejia-Ramos & Inglis, 2009). Further, the studies that have focused on the reading of proof have usually examined the ways in which students make judgments on mathematical arguments—such as whether a particular argument is convincing or would qualify as a proof—often with the aim of providing insights into students' perceptions of proof. Mejia-Ramos and Inglis (2009) note that there are few empirical studies on how well students are able to understand proofs.

Our assessment model is based on two pioneering articles in this area—the work of Yang and Lin (2008) and that of Conradie and Frith (2000). Yang and Lin (2008) made an important first step toward understanding proof comprehension by introducing what they called a *model of reading comprehension of geometry proof* (RCGP). Yang and Lin's model consists of four levels and five facets of proof comprehension.

At the first level, termed *surface*, students acquire basic knowledge regarding the meaning of statements and symbols in the proof. At the second level, which Yang and Lin called *recognizing the elements,* students identify the logical status of the statements that are used either explicitly or implicitly in the proof. At the third level, termed *chaining the elements*, students comprehend the way in which these different statements are connected in the proof by identifying the logical relations between them. Finally, at the fourth level, which Yang and Lin referred to as *encapsulation,* students interiorize the proof as a whole by reflecting on how one may apply the proof to other contexts.

Yang and Lin (2008) focused predominantly on the first three levels of their model (which seem to be crucial in the comprehension of high-school geometry), while leaving the fourth level noticeably unspecified. In particular, Yang and Lin indicated that their instrument for measuring students' proof comprehension was not aimed at diagnosing if a student had reached this top level (p.71).

The assessment model we propose seeks to adapt Yang and Lin's RCGP model to contexts in advanced mathematics. Critically, this involves expanding upon their encapsulation level of understanding (which we believe is of crucial importance in the comprehension of proofs in advanced mathematics courses) and restructuring their facets to make them more relevant to proof comprehension in undergraduate mathematics.

Conradie and Frith (2000) raised the issue of comprehension tests in advanced mathematics. In addition to stressing their importance, these researchers provided illustrations of comprehension tests for two different proofs. In this paper, we aim to go further by providing a more systematic way of generating this type of tests.

3. Method for generating the assessment model

In this paper, we propose seven types of questions one could employ to assess students' understanding of a proof in advanced mathematics. We believe each of these types measures a different facet of proof comprehension—for instance, we can imagine a student (or mathematician) who grasps the big picture of a proof without understanding the technical details of the proof, and vice versa. Thus we do not view these different types of assessment as part of a hierarchy, but rather as measuring students' proof comprehension along different dimensions.

The components of our model can be separated into two groups. The first group focuses on assessing one's understanding of *local* aspects of the proof. By this we mean understanding that can be discerned either by studying a specific statement in the proof or how that statement relates to a small number of other statements within the proof. Critically, questions assessing the comprehension of local aspects of the proof could be answered by ignoring most of the statements in the proof. Our work in generating this part of our assessment model consists of adapting the first three levels and facets of Yang and Lin's (2008) model of reading comprehension of geometry proofs to the more complex proofs that undergraduates encounter in their advanced mathematics courses.

The second group focuses on assessing one's *holistic* understanding of the proof, which cannot be gleaned by examining a small number of statements in the proof, but rather must be ascertained by inferring the ideas or methods that motivate a major part of the proof, or the proof in its entirety. This second group is an elaboration of Yang and Lin's (2008) notion of encapsulation, and it was generated by studying the types of understanding valued by the mathematics education research community and mathematicians who teach advanced mathematics courses. We took the following steps to generate this second group of components of our assessment model.

First, we reviewed the mathematics education research literature on the purposes of proof (e.g., de Villiers, 1990; Hanna, 1990; Hersh, 1993; Weber, 2002; Hanna &

Barbeau, 2008). We reasoned that one way we can evaluate whether a student understood a proof is by measuring the extent to which the proof achieves these purposes for that particular student. To illustrate, some mathematics educators (e.g. Hanna & Barbeau, 2008) have argued that a primary purpose of proof is to illustrate new methods for proving and problem solving. Hence, we deduced that one way to assess whether a student understood a given proof is by evaluating the extent to which that student could apply the proof's method in other situations.

Second, we reviewed the recommendations for alternative methods of presenting proofs (e.g., Leron, 1983; Hersh, 1993; Rowland, 2001; Alcock, 2009). The authors who proposed these formats all contended that these formats had the potential to improve students' understanding of a proof, although none defined what understanding a proof actually entailed. We inferred what *specific* understandings the authors felt that students might gain from these alternative presentations. For instance, Leron's (1983) structured proofs were organized in terms of the proof's components to help students see how these components supported the main idea of the proof. Accordingly, we inferred that being able to both summarize a proof and identify the relationship between the components of the proof constitute two dimensions of understanding it.

Finally, we investigated what types of proof comprehension were valued by the mathematicians who teach advanced mathematics courses. To investigate this, we conducted semi-structured interviews with nine such mathematicians about why they read proofs, what they thought it meant to understand a proof, and their goals for presenting proofs to their students. The methodology and results of this study are discussed in detail by Weber and Mejia-Ramos (2011) and Weber (2010).

If a particular facet of understanding was mentioned by at least two mathematicians *and* was discussed in the mathematics education literature, we incorporated it into our model of assessing the comprehension of holistic aspects of the proof. We found four such facets: summarizing the main idea of the proof, understanding the proof in terms of its components or modules, applying the method of the proof in other contexts, and illustrating the proof with examples or diagrams.[1]

4. A model for assessing proof comprehension

In this section, we describe ways to assess students' understanding of seven different aspects of a proof in advanced mathematics. For each type of question, we first describe the type of understanding it is designed to evaluate. Next, we describe how this type of assessment is an adaptation of a component in Yang and Lin's (2008) proof comprehension model, or how it was discussed by the interviewed mathematicians and in the mathematics education literature. After this, we present templates for how these types of assessments can be generated. Finally, we give specific examples of questions that can evaluate one's understanding of the following proof:

We say that a number is monadic if it can be represented as $4j+1$, and triadic if it can be represented as $4k+3$, for some integers j and k.

Theorem. There exist infinitely many triadic primes.

1. Consider a product of two monadic numbers:
   $(4j+1)(4k+1) = 4j \bullet 4k + 4j + 4k + 1 = 4(4jk + j + k) + 1,$
   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 $p_1, p_2, \ldots, p_n$.

---

[1] We note that there were two related types of proof comprehension mentioned by mathematicians that were *not* discussed in the mathematics education research literature—recognizing where a proof becomes non-routine or difficult, and recognizing why the obvious approach to proving a particular theorem fails.

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.

Before reading the following section, we suggest that the reader consider how one could assess students' understanding of this proof at the undergraduate level.

4. 1. Assessing the local comprehension of a proof

The first three levels of Yang and Lin's (2008) model address proof comprehension at the level of specific terms and statements (what they mean, what their logical status is, and how they connect to preceding and succeeding statements), as opposed to the fourth level, which addresses the comprehension of the proof as a whole (its generality and application to other contexts). In this sense, the first three levels of Yang and Lin's model address students' understanding of *local* aspects of the proof. In this section we focus on types of question to assess this type of understanding; i.e. questions that address only one, or a small number, of statements within the proof.

4. 1. 1 Meaning of terms and statements

One of the most fundamental ways to understand any type of text is to understand the meaning of individual words and sentences. In the case of proof, one can assess a reader's comprehension of this aspect by asking him or her to identify the definition of a key term in the proof, or to specify what is meant by some of its statements. Yang and Lin (2008) featured this aspect of proof prominently at their *surface level* of proof comprehension. Research suggests that students often fail to understand the meaning of key terms when reading a proof (Conradie & Frith, 2000), hindering their ability to comprehend other aspects of the proof, and that less successful students sometimes do not try to understand the meaning of key terms and statements (Weber, Brophy, & Lin, 2008).

Assessing the extent to which readers understand the meaning of specific terms appearing in the proof may involve asking them to:

1. *State the definition of a given term in the proof* (e.g., "define the given term in your own words", "which of the following statements defines the given term?"). For instance, in the proof above, one could ask, "what does it mean for a number to be triadic?"[2] or "what does it mean for a number to be prime?"
2. *Identify examples that illustrate a given term in the proof* (e.g., "give a specific example that illustrates the given term", "which of the following cases exemplifies the given term?"). For instance, for the proof above, one could provide a list of natural numbers and ask, "which of the following are triadic primes?"

Further, in order to assess a reader's comprehension of individual statements in the proof (including the proven statement) one may ask them to:

1. *State a given statement in a different but equivalent manner* (e.g., "write the given statement in your own words", "which of the following statements are equivalent to the proven theorem?").
2. *Identify trivial implications of a given statement* (e.g., "which of the following statements are true based on the given statement?", "which of the following are immediate consequences of the given statement?").

---

[2] The proof used in this paper introduces new terminology such as "triadic", but as the following question illustrates, one need not only ask about new terminology.

3. *Identify examples that illustrate a given statement* (e.g., “which of the following cases verify the statement for a particular example?”, “which of the following cases does the given statement rule out?”). In our example, one could ask: “according to the statements in lines 1-2, is it definitely true, possible, or impossible that 5^{20} is a monadic number?”

A reader could conceivably answer these types of questions without ever having read the proof itself (although reading the proof may help a learner develop a better understanding of key terms and statements). However, this is not always the case, as some proofs introduce new terminology.[3] We also note that, even though these questions may be simple to pose—because they pertain only to a single statement within the proof—they nonetheless can be designed to be quite challenging.

4. 1. 2. Logical status of statements and proof framework

Yang and Lin (2008) note that in high school geometry proofs, the assertions within the proof can have different statuses, such as being an axiom or postulate, an established fact or theorem, a hypothesis of the theorem to be proven, or a statement deduced from previous assertions. Understanding the status of the different assertions in the proof is necessary to understanding the logic of the proof. We agree with this analysis and argue that the situation is more complex in advanced mathematics.

In the majority of high school geometry proofs, the conclusion of the proven theorem is directly derived after assuming its hypotheses are true. That is, high school geometry proofs of a conditional statement “if *p,* then *q*” generally assume all conditions

[3] In this sense, this aspect of proof comprehension sometimes differs from the “surface level” described by Yang and Lin. This difference is due, in part, to the content of proofs in high school geometry and advanced mathematics. It is uncommon to create new terminology for a specific proof in high school geometry.

stated in p and deduce that q is also the case. In advanced mathematics, some proofs assume the negation of q and deduce the negation of p (proof by contraposition), other proofs make additional assumptions to those mentioned in the proven statement (e.g. proof by cases, proofs by induction), and yet others assume the negation of the proven statement and deduce a contradiction (proof by contradiction).

Thus, in the context of advanced mathematics a reader needs to not only identify the logical status of statements in proofs, but also recognize the logical relationship between the statement being proven and the assumptions and conclusions of the proof. Recognizing this relationship involves understanding what Selden and Selden (1995) called the *proof framework*: "the 'top-level' structure of the proof, which does not depend on detailed knowledge of the relevant concepts". For instance, the proof framework of a proof by contraposition of the claim, "If n^2 is odd, then n is odd" would take the claim "n is even" as an assumption and "n^2 is even" as the conclusion. A reader who understands this proof framework would not only identify the logical status of these two statements in the proof ("n is even" is an assumption and "n^2 is even" is a conclusion), but also recognize the logical equivalence between the original conditional statement and its contrapositive. Notice that in this case, the reader recognizes the proof framework by studying the logical status of only two statements in the proof (initial hypothesis and conclusion), potentially ignoring most of its other statements. In this sense, even though the proof framework refers to the "top-level" structure of the proof, it could be understood by studying a *local* aspect of the proof.

In general, assessing a reader's comprehension of the proof framework and the logical status of statements in the proof may involve asking him or her to:

1. *Identify the purpose of a sentence within a proof framework* (e.g., "In the proof, what is the purpose of making this particular assumption?", "what is the purpose of considering these particular cases"?). In our example, one could ask: "In line 3, what is the purpose of assuming that the theorem is false and there are only finitely many triadic primes?"
2. *Identify the type of proof framework.* This type of question could simply ask students to name the kind of proof being employed. For example, "is the proof a proof by contradiction, a proof by contraposition, or a proof by cases?"

As Selden and Selden (2003) illustrated, and Weber (2009) replicated, students often fail to consider the proof framework of a proof they are reading, leading students to be unsure of what is being proven and unable to determine if a proof is correct.

4. 1. 3. Justification of claims

In a proof, new statements are deduced from previous ones by the application of accepted mathematical principles (e.g., theorems, logical rules, algebraic manipulations). However, as with all scientific texts, a proof would be impossibly long if all of its logical details were explicitly stated (Davis & Hersh, 1981). In many cases, the reader needs to infer what previous statements and mathematical principles are used to deduce a new assertion within a proof. Research has illustrated that undergraduates sometimes fail to do this when reading a proof (e.g., Alcock & Weber, 2005; Weber & Alcock, 2005). Thus, one dimension of proof comprehension involves being able to provide justifications for how new assertions follow from previous ones. This dimension is essentially Yang and Lin's (2008) *chaining level* of proof comprehension.

In general, in order to assess the extent to which readers comprehend the justification of claims in the proof, one may ask them to:

1. *Make explicit an implicit warrant in the proof.* Generally, proofs include expressions of the form "Since A, then B", where the claim B is justified simply by citing statement A. In this type of expression, the general rule according to which statement A is sufficient to conclude statement B is left implicit. In some cases (e.g., proofs appearing in specialized publications) this may be done under the assumption that this general rule is obvious to the reader, and therefore it would be superfluous to mention it. In other cases (e.g., proofs appearing in undergraduate mathematics textbooks) this may be done with the expectation that readers work out for themselves what these general rules are. Demonstrating an understanding of such justifications would involve making the rules explicit (e.g., "what general rule justifies that [*claim of the form 'Since A, then B'*] in the proof?"). In our particular proof, one could ask, "In line 5, why does the fact that 3 does not divide $4p_2 \cdots p_n$ imply that 3 does not divide M?"
2. *Identify the specific data supporting a given claim.* It is also common for proofs to include expressions of the form "Hence, C", in which the claim C is justified by some unspecified subset of all the previous statements in the proof. Often, these claims are justified by statements immediately preceding them in the proof, but they may also be justified by statements from different parts of the proof (as well as other implicit warrants). Assessing the extent to which a reader understands how a claim presented in this manner is justified may involve asking him or her to identify the specific statements within the proof that provide the basis for a claim

(e.g., "which of the following statements in the proof allow the conclusion [*claim of the form 'Hence, C'*]?"). In the proof above, one may ask, "What statement can be used to justify the claim in line 9, that *M* is clearly triadic?"

3. *Identify the specific claims that are supported by a given statement.* Proofs often reach specific results or state specific assumptions without explicitly specifying how these results or assumptions are later used in the proof. Therefore, another way of assessing this dimension of proof comprehension involves asking a reader to identify the exact place(s) in the proof where a given piece of information is employed as justification of new claims (e.g., "which of the following claims in the proof logically depend on the assertion [*specific proposition in the proof*]?"). For instance, for the proof above, one could ask, "which claims in the proof logically depend on line 2 of the proof, the claim that the product of monadic numbers is monadic?"

4. 2. Assessing the holistic comprehension of a proof

The fourth level in Yang and Lin's (2008) RCGP model, termed *encapsulation,* deals with students' interiorization of the proof as a whole. This level addresses a more *holistic* comprehension of the proof, in which the proof as a whole is understood in terms of its main ideas, methods, and application to other contexts. In this section we develop four different ways of assessing this type of comprehension in undergraduate mathematics.

4. 2. 1. Summarizing via high-level ideas

A common complaint about the linear way in which proofs are conventionally written is that it masks the higher-level ideas contained in the proof: in studying the specific logical details of the proof, one can lose track of the big picture (e.g., Leron, 1983; Anderson, Boyle, & Yost, 1986; Alibert & Thomas, 1991). One way that a proof can be understood is in terms of the overarching approach that is used within a proof. Several mathematicians that we interviewed indicated that this is one way that proof could be read or understood and, further, that this is different than checking the logical details of a proof.

> M2: There's two ways you want people to read proofs. First way is to scan it for the main idea […] And the second way to read it is to understand the logic of it, which is a different kind of reading.

> M5: When I read the theorem, I think, is this theorem likely to be true and what does the author need to show to prove it's true. And then I find the big idea of the proof and see if it will work. If the big idea works, if the key idea makes sense, probably the rest of the details of the proof are going to work too.

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

In these excerpts, the mathematicians valued understanding the proof in terms of the main idea being applied. Seeing the big idea of the proof, as opposed to seeing the proof solely as a chain of specific logical assertions, is a large part of the motivation behind Leron's (1983) structured proofs. These proofs are presented in a hierarchical fashion, where "the top level gives in very general (but precise) terms the main line of the proof" (p. 174), allowing the proof to "be grasped at a glance, yielding an overview of the proof" (p. 175). These proofs have the goal that "the ideas behind the formal proof are better communicated" (p. 183). Leron also advocates a good learning activity and a test of students' understanding is to take a ten-page proof and describe it in one page (p. 184). Since Leron proposed his notion of structured proofs, mathematics educators have commonly cited this as a way to improve students' proof comprehension, implying that having a top-level overview, or understanding the "big idea" of the proof, is an important part of proof comprehension.

In general, assessing readers' comprehension of a proof's high-level ideas may involve asking them to:

1. *Identify or provide a good summary of the proof.* In this case, a reader is asked to provide a summary of the proof. Because objectively grading the quality of a proof summary might be difficult, a teacher may instead provide several summaries of the proof and ask the student to choose which summary captures the main idea of the proof. For our example proof, one could ask students to select the better of the following two summaries, where the first summary captures the main idea of the proof and the second includes too much detail for some ideas and leaves others out:
    - It assumes there are finitely many triadic primes and uses them to construct a triadic number *M* that has only monadic prime factors, which would imply *M* is also monadic. This is a contradiction.

- It lets $M = 4p_2 \cdots p_n + 3$, where p_i are prime numbers and $p_i \neq 3$. Thus, 2 does not divide *M* because *M* is odd. Further p_i does not divide *M* because it leaves a remainder of 3. This produces a contradiction.
- I don't know which summary would be better.

2. *Identify a good summary of a key sub-proof in the proof.* In this type of assessment, the reader is asked to provide or identify a summary of a key sub-proof of the proof (e.g., "which of the following best summarizes why [*significant result within the proof*]?").

4. 2. 2. Identifying the modular structure

Several mathematicians indicated that understanding a proof entailed breaking the proof into components or modules and then specifying the logical relationship between each of the modules, as the following quotes illustrate:

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

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

> proof a lot because, you know, he's carrying along all these hypotheses which were really unnecessary.

Breaking a proof into manageable modules was one of the motivations behind Leron's (1983) structured proofs, as his method allows the reader to clearly see how the proof can be partitioned and what the purpose of each module is. In a sense, proofs that contain lemmas can also be thought of as partially structured proofs, as the lemma can be thought of as a separate entity from the proof, which can be used both within the proof and elsewhere. The following are two ways of assessing students' understanding of the relationships between different modules in a proof:

1. *Ask students to partition a proof into modules.* One could assess this dimension of proof comprehension by asking a reader to partition a proof into modules. Leron (1983) explicitly suggested that asking students themselves to structure a proof would help them understand it.
2. *Identify the purpose of a module of a proof*. In this case, one could ask a student to identify the role of a given part of the proof within its general framework. These parts could be a lemma, a specific case in a proof by cases, or any type of subproof within the original proof (e.g., in a proof by cases one could ask: "which of the following statements best justifies the consideration of [*specific case*] in the proof?"). In the proof above, one could ask, "why was the sub-proof that the product of monadic numbers is monadic included in the proof?"
3. *Identify the logical relation between modules of a proof.* One could also assess this dimension of proof comprehension by asking a reader to identify the logical relation between two or more modules of a proof (e.g., "which of the following

best describes the logical relation between [*two or more parts of the proof framework*]"?). In the proof above, one question that addresses this point might be, "What is the logical relation between lines 1-2 in the proof, which establish that the product of monadic primes is monadic, and lines 3-7, which establish that *M* is not divisible by a triadic prime?" (In this case, the two modules are logically independent of one another).

4. 2. 3. Transferring the general ideas or methods to another context

An important aspect of comprehending a proof involves identifying the procedures used in the proof and the ways in which these procedures can be applied (or re-interpreted) to solve other proving tasks. Every mathematician that we interviewed mentioned this as a primary reason for reading the proofs of others, as the following representative excerpts illustrate:

M1: You read the proof carefully and you discover these are things that you can use. That's certainly from a pragmatic point of view, that's an important part of reading proofs, that you steal good ideas out of good proofs.

M6: Well, I would say most often is to get some ideas that might be useful to me for proving things myself. […] usually I'm reading something because it seems to have some connection to some problems that I was interested in. I'm hoping that if that the tools they're using or ideas they're using might connect up to some of these problems that I have thought about.

M4: [When reading a proof,] I want to understand the proof technique in case I can use bits and pieces of that proof technique to prove something that they haven't yet.

Several of the interviewed mathematicians also mentioned that they hoped students would be able to apply the ideas of the proofs that they presented in class to other contexts as well. In the mathematics education literature, Weber (2002) noted that the reason for presenting proofs of seemingly obvious statements (such as "$f(x)=x^2$ is a continuous function" in real analysis) is to illustrate for students how the method in the proof can be used to prove a class of theorems. Building on the philosophical work of Rav (1999), Hanna and Barbeau (2008) argued that illustrating methods is an important role of proof in the mathematics classroom. In our example, the method[4] of proof involves using a (presumed) finite list of primes to construct a number with contradictory properties (in this case, being both monadic and triadic).

In general, there are at least three ways in which this dimension can be assessed:

1. *Transfer the method.* This involves being able to successfully apply the method in the solution of a different task (e.g., "Using the method in the previous proof, prove [*similar theorem*]"). In our specific example, one might ask, "Using the method of the proof above, prove that there are infinitely many primes of the form $6k+5$".
2. *Identify the method.* This involves being able to identify the *general manner* in which the method of the original proof can be applied in a different proving task (e.g., given a theorem T that can be proven using the method displayed in the original proof, ask "which of the following general procedures would you follow

[4] What counts as a method of proof is a subtle issue that is beyond the scope of this paper. A proof might have several methods. However, we anticipate that many proofs for which one would test students' comprehension have a fairly well-delineated method.

to prove T?", or "how would you start proving T?"). In the example above, rather than asking a student to write a complete proof of the theorem that there are infinitely many primes of the form $6k+5$, one could simply ask what would be an appropriate value of M if one were to use such an approach.[5] One benefit of this type of question over a strict transfer one is that students' performance on this item would be less time-consuming and would be less dependent on students' domain knowledge (i.e., one could understand the general method needed to prove a theorem but lack the logical or algebraic proficiency to complete the proof).

3. *Appreciate the scope of the method*. This involves recognizing the assumptions that need to be in place to allow the method to be carried out (e.g., "explain why the method used in proof P of theorem T would not be useful for proving statement S?"). In the example above, one could ask, "Why can't the ideas used in this proof be used to prove there are infinitely many monadic primes?" (a product of triadic numbers need not be triadic).

4. 2. 4. Illustrating with examples

Comprehending a proof often involves understanding how the proof relates to and could be illustrated by specific examples—that is, being able to follow a sequence of inferences in the proof in terms of a specific example. Many of the mathematicians that we interviewed emphasized that this was an indispensable tool that they used to gain an understanding of a proof that they were reading, as the following excerpts illustrate:

---

[5] To avoid students answering this question by matching symbols rather than using a deep understanding, one could also ask the student to justify why this choice of M was appropriate, or to present them with a poor choice of M and ask why the proof would not work in that case.

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

M5: [When asked if he used examples to understand proofs] Always. Always. Like I said, I never just read a proof at an abstract level. I always use examples to make sure the theorem makes sense and the proof works. […] When I'm looking through a proof, I can go off-track or believe some things that are not true. I always use examples to see that [it] makes sense.

This idea is also behind the generic proof presentation that Rowland (2001) and others (e.g., Weber, Porter, and Housman, 2008) advocate. A similar tool discussed by some mathematicians was the importance of relating a proof to a diagram to better understand it. For instance, M2 declared that a student could not understand the proof that monotonic bounded sequence converges if he or she did not draw a picture of the statement first. The idea of linking formal proofs to pictures (or other informal models) is something that both mathematics educators (e.g., Raman, 2003) and mathematicians (e.g., Thurston, 1994) argue is important for developing understanding.

In general, questions that can be employed to assess a reader's understanding of this dimension require the reader to:

1. *Illustrate a sequence of inferences with a specific example.* This involves being able to identify the way in which a sequence of inferences in the proof applies to a given specific example (e.g., "Using the ideas in the proof, how would you show [*statement about a particular example*]"?). In the proof used in this paper, one

might ask, “If 3, 7, 11, and 19 were the only triadic primes, what would *M* be and why would we be certain that *M* was triadic and none of the listed triadic primes divided *M*?”

2. *Interpret a statement or its proof in terms of a diagram*. This involves being able to state how the ideas of the theorem statement or its proof relate to a carefully chosen diagram (e.g., “Given the diagram below, label a graphical interpretation of [*a variable within the proof*] and explain how you know that [*this variable has a pertinent property*]”)

5. Discussion

5. 1. Summary of assessment model

In this paper we present a multi-dimensional model for assessing proof comprehension in undergraduate mathematics. Our model describes ways to assess students’ understanding of seven different aspects of a proof. The first three types of assessment address students’ comprehension of only one, or a small number, of statements within the proof. These types of assessment, which we term *local*, are:

1. *Meaning of terms and statements:* items of this type measure students’ understanding of key terms and statements in the proof.
2. *Logical status of statements and proof framework:* these questions assess students’ knowledge of the logical status of statements in the proof and the logical relationship between these statements and the statement being proven.
3. *Justification of claims:* these items address students’ comprehension of how each assertion in the proof follows from previous statements in the proof and other proven or assumed statements.

The remaining four types of assessment, which we call *holistic*, address students' understanding of the proof as a whole. These assessment types are:

4. *Summarizing via high-level ideas:* these items measure students' grasp of the main idea of the proof and its overarching approach.
5. *Identifying the modular structure:* items of this type address students' comprehension of the proof in terms of its main components/modules and the logical relationship between them.
6. *Transferring the general ideas or methods to another context:* these questions assess students' ability to adapt the ideas and procedures of the proof to solve other proving tasks.
7. *Illustrating with examples:* items of this type measure students' understanding of the proof in terms of its relationship to specific examples.

Although we believe that each of these types of questions measures a different facet of proof comprehension, there may be interesting relationships between them. For instance, in certain cases, being able to summarize the proof in terms of its high-level ideas may be necessary in order to successfully transfer these ideas and methods to another context. The study of these relationships is an interesting avenue for further research.

Also, there is not a uniform use of these assessments, since some proofs may not be amenable to all types of questions. For instance, our example proof did not have an obvious visual interpretation and therefore it would not be appropriate to relate that proof to a diagram. Other shorter and straightforward proofs, such as a standard proof that $\sqrt{2}$ is irrational, would not be possible to break into modules or sub-proofs.

5. 2. Contributions of this paper

This paper builds upon the work of Yang and Lin (2008) and Conradie and Frith (2000) in a significant way. Yang and Lin (2008) made an important contribution toward delineating students' comprehension of a proof by proposing a four-level hierarchy of understanding. Their focus was on the first three levels, which dealt with the meaning of the statements within a proof and the logical structure of the proof. This is quite appropriate for proofs in a high school geometry course, which is the scope that Yang and Lin assigned to their model. However, in undergraduate mathematics courses, where the proofs become longer and different proof strategies are employed, there are more sophisticated ways in which a proof can be understood. We incorporate Yang and Lin's contributions into our model and we elaborate on other aspects of understanding a proof at the undergraduate level.

Conradie and Frith (2000) proposed specific questions that they used to assess students' understanding of proof in their own classrooms. Our model generates many of the types of questions that Conradie and Frith (2000) advocated, justifies why these types of questions are important, and suggests methods that other instructors and researchers can use to generate questions for other proofs.

5. 3. Implications

In this paper, we have presented a multi-dimensional model for assessing proof comprehension at the undergraduate level. Although we believe that teachers of advanced mathematics can also benefit from the model, we shall focus on its significance for educational researchers.

We propose that our model can contribute to research on proof comprehension in various ways. First, despite the importance of proof comprehension in advanced undergraduate mathematics courses and widespread complaints that undergraduates do not understand the proofs that they read, there is little empirical research on proof comprehension (Mejia-Ramos & Inglis, 2009). Examining how undergraduates answer the types of questions we propose can reveal how much (or how little) undergraduates gain from reading a proof as well as indicating the specific types of questions that typically give undergraduates difficulty.

Second, the model proposed in this paper can be used as a methodological tool for evaluating the effectiveness of instructional interventions designed to increase the comprehensibility of mathematical presentations. In particular, several researchers have suggested novel proof presentation formats that promise to be more meaningful for students (e.g., Alcock, 2009; Leron, 1983; Rowland, 2001). We do not claim that our assessment model can conclusively answer the question of whether these pedagogical recommendations are effective—for instance, even if a particular proof format did not improve proof comprehension, it might be beneficial for students' appreciation of proof in general. However, we contend that researchers could use our assessment models to document specific types of comprehension benefits that their innovative proof presentations may have. For instance, if students perform better on questions related to "transferring the method used in the proof" after reading a generic proof as opposed to a traditional proof, this would provide evidence that generic proofs facilitate students abilities to understand and apply the methods used in the proofs. In recent years there have been several empirical studies assessing the extent to which these novel formats improve proof comprehension (Fuller et al, 2011; Malek & Moshovits-Hadar, 2011; Roy, Alcock, & Inglis, 2010). However, the comprehension tests used in these studies were generated using different frameworks or in an ad hoc manner. We hope that our assessment model can inform future research in this area by suggesting particular types of questions that researchers can use, and by serving as a theoretical framework for comparing the results across different studies of this type.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. DRL0643734.

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 January 20, 2011.

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

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

Anderson, J. R., Boyle, C. F. & Yost, G. (1986). The geometry tutor. *Journal of Mathematical Behavior, 5,* 5-19.

Conradie, J., & Frith, J. (2000). Comprehension tests in mathematics. *Educational Studies in Mathematics, 42*, 225-235.

Davis, P. J. & Hersh, R. (1981). The Mathematical Experience. Boston: Birkhäuser.

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

Fuller, E., Mejia-Ramos, J. P., Weber, K., Samkoff, A., Rhoads, K., Doongaji, D., Lew, K. (2011). Comprehending Leron's structured proofs. In *Proceedings of the 14th Conference for Research in Undergraduate Mathematics Education.*

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

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

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

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

Malek, A., & Movshovitz-Hadar, N. (2011). The effect of using transparent pseudo proofs in linear algebra. *Research in Mathematics Education, 13*(1), 33-58.

Mamona-Downs, J. & Downs, M. (2005). The identity of problem solving. *Journal of Mathematical Behavior, 24*, 385-401.

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.

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

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

Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement, and instruction* (pp. 37-75). Boston: Kluwer.

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). Undergraduates proof comprehension: A comparative study of three forms of proof presentation. 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 (1988) When good teaching leads to bad results: The disasters of "well-taught" mathematics courses. *Educational Psychologist, 23*(2), 145-166.

Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 34*(1), 4-36.

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

Thurston, W. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society, 30*(2), 161-177.

Weber, K. (2002). Beyond proving and explaining: Proofs that justify the use of definitions and axiomatic structures and proofs that illustrate technique. *For the Learning of Mathematics, 22*(3), 14-17.

Weber, K. (2009). Mathematics majors' evaluation of mathematical arguments and their conception of proof. In *Proceedings of the 12th Conference for Research in Undergraduate Mathematics Education*. Available for download at: http://sigmaa.maa.org/rume/crume2009/proceedings.html. Last downloaded April 10, 2010.

Weber, K. (2010). Proofs that provide insight. *For the Learning of Mathematics, 30*(1), 32-36.

Weber, K. (in press). The pedagogical practice of mathematicians. To appear in *International Journal of Mathematics Education in Science and Technology*.

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

Weber, K., Brophy, A., & Lin, K. (2008). How do undergraduates learn about advanced mathematical concepts by reading text? In *Proceedings of the 11th Conference on Research in Undergraduate Mathematics Education.* San Diego.

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

Yang, K.-L., & Lin, F.-L. (2008). A model of reading comprehension of geometry proof. *Educational Studies in Mathematics, 67*, 59-76.