

This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.



For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

Mathematicians' Perspectives on the Teaching and Learning of Proof

Lara Alcock

ABSTRACT. This paper reports on an exploratory study of mathematicians' views on the teaching and learning that occurs in a course designed to introduce students to mathematical reasoning and proof. Based on a sequence of interviews with five mathematicians experienced in teaching the course, I identify four modes of thinking that these professors indicate are used by successful provers. I term these instantiation, structural thinking, creative thinking and critical thinking. Through the mathematicians' comments, I explain these modes and highlight ways in which students sometimes fail to use them effectively. I then discuss teaching strategies described by the participants, relating these to the four modes of thinking. I argue that teaching aimed at improving structural thinking tends to dominate, and that courses that introduce proof, regardless of classroom organization, should address all four modes in a balanced and integrated way.

1. Introduction

The research reported here addresses the teaching and learning of mathematical proof through the perspectives of mathematicians who teach a course entitled Introduction to Mathematical Reasoning. This is a "transition course," designed to facilitate the move from calculation-oriented mathematics courses to proof-oriented courses such as real analysis and abstract algebra. Like other such courses, it involves elementary work on logic, set theory, relations, functions, and various types of proof (contradiction, mathematical induction, etc.), with some of this done within mathematical topics such as number theory and combinatorics. One goal is to familiarize students with the mechanics of proof, so that when they enter upper level courses they will be able to concentrate on the content.

Research on transition courses has focused largely on two issues. One is student learning in such contexts, with a frequent finding that this is less than we might hope (Moore, 1994). The second is alternative pedagogical approaches to introducing proof, with researchers investigating collaborative classes (Blanton & Stylianou, 2003). The study reported here took a different approach, investigating the teaching and learning that occurs in this particular transition course from the

The author wishes to thank the mathematicians who took part in this research, along with Keith Weber and Adrian Simpson for their valuable comments on earlier drafts.

perspective of the mathematicians for whom it is part of their regular teaching. It involved in-depth interviews with five such mathematicians, and I aimed to investigate the participants' experiences *as they reported them*, with no attempt to establish whether their comments constituted an accurate report of their own or their students' behavior in the classroom. This was a deliberate decision, taken for three reasons. First, because I wanted to establish a relationship with the professors based on good faith and respect, which would have been difficult if they felt that their comments and their teaching stood to be negatively evaluated. Second, because I believe that if the mathematics education community wishes to engage mathematicians in discussions about their teaching, it needs to understand what they perceive as their most pressing concerns and to take these seriously. Third, because the participants have an enormous amount of collective experience that constitutes a rich source of information about the learning and teaching of proof in its own right.

This paper begins with a review of the mathematics education literature on the teaching and learning of proof and a detailed description of the methods used to collect and analyze the interview data. Following this, the results are reported in two distinct sections. The first of these, Section 5, describes the four modes of thinking that I discerned in the professors' discussions of proving, together with their comments on their students' use of these modes. The other, Section 6 describes the teaching strategies the professors use, relates these to the four modes, and concludes with the main points of an ongoing debate among the mathematician participants about what should be the nature of the course. In Section 7, I discuss ways in which I believe these results should influence our thinking about how we teach transition-to-proof courses and other advanced mathematics courses. The paper concludes with a short discussion of some remaining issues and directions of ongoing research.

2. Previous Work on the Teaching and Learning of Proof

There has been much research on university students' difficulties with both constructing and evaluating proofs. Students' standards of justification may differ from those of mathematical proof, so that they offer empirical arguments where deductive ones are required (Harel & Sowder, 1998; Recio & Godino, 2001). Also, students may be unaware of the importance of definitions in constructing proofs (Edwards & Ward, 2004; Vinner, 1991), and they often find it difficult to understand or work with many of the definitions of advanced mathematics (Dubinsky, Elterman, & Gong, 1988; Moore, 1994). This may be partly attributed to the fact that they have less familiarity with examples of concepts than do their teachers (Moore, 1994) and less inclination to generate examples in order to facilitate understanding of a new definition (Dahlberg & Housman, 1997). When evaluating a proof given by others, students may base their judgment on its surface resemblance to formal arguments rather than on its content (Segal, 2000). They may not have the skills necessary to identify the logical structure of a mathematical statement or proof, particularly when this involves interpreting quantifiers and conditional statements (Dubinsky & Yiparaki, 2000; Durand-Guerrier, 2003; Weber & Alcock, 2005). As a result, students cannot reliably judge whether a proof establishes a given result (Selden & Selden, 2003). Overall, they may come to view proofs and

proving as unrelated to their own ways of thinking about mathematical concepts and relationships (Moore, 1994; Raman, 2003).

There has been rather less research on how students can come to understand the nature of proof and master the art of proving. Much of the extant work involves theoretical arguments (e.g., Simpson, 1995) and case studies (e.g., Pinto & Tall, 2002), both of which suggest that there may be at least two “routes” to proof. One of these is a “formal” route in which formal routines and manipulations are first established and later come to be invested with meaning, and the other a “natural” route in which informal ideas and images are progressively refined and linked to formal definitions (Pinto & Tall, 2002; Simpson, 1995; Weber, 2001). Suggestions for ways to improve the teaching of proof tend to reflect one or other of these routes. Some suggest a systematic approach based on a solid grounding in logic and its associated linguistic expressions (Epp, 1998, 2003; Selden & Selden, 1999), or on a presentation or sequence of problems that can lead students to more easily see the structure of certain proof types (Harel, 2002; Leron, 1985). With a similar focus on structure, empirical studies have specified skills students need to produce simple proofs in a given area (Gholamazad, Lijedahl, & Zazkis, 2003; Weber, 2006), and have investigated technological environments that can focus students’ attention on the axiomatic development of a mathematical theory (Cerulli & Mariotti, 2003). Suggestions and research in keeping with the “natural” route to proof tend to focus on classroom environments in which students and teacher work collaboratively, investigating problems, formulating conjectures, debating the validity of arguments and so forth (Alibert & Thomas, 1991; Blanton & Stylianou, 2003; Rasmussen, Zandieh, King, & Teppo, 2005; Yackel, Rasmussen, & King, 2000).

While we have these suggestions for how to improve the teaching of proof-oriented mathematics, there have been few empirical studies on how transition courses are usually taught. Both Moore (1994) and Weber (2004) have reported on the teaching of individual professors, and on the articulated rationales these professors give for presenting proof-based material in a certain way. However, I am not aware of any research that seeks to synthesize the knowledge, views and methods developed by mathematicians who teach introductory proof courses. This study takes a step toward addressing this.

3. Research Context

The Introduction to Mathematical Reasoning course described here is offered in a primarily lecture-based format at a large state university in the United States and is a prerequisite for the majority of the upper level mathematics courses. Around 100 to 120 students per semester take it, approximately 40 of whom will go on to be mathematics majors. The course is taught in classes of 20 to 25 students that meet for two 80-minute periods per week during one 14-week semester. The small class size allows professors to have relatively close contact with individuals and to become familiar with their work. The five mathematicians who took part in this exploratory research were invited to participate because they were teaching the course during the 2002-2003 academic year, when the interviews took place. Four of the five had taught the course several times and had also participated in more and less formal discussions with each other about such things as its design and the choice of textbook. All had also taught content-based upper level courses at either the research site or other institutions and they were familiar with the problems

that Introduction to Mathematical Reasoning was introduced to alleviate. Their comments reflect direct experience in teaching the course as well as their knowledge of each others' views on the course.

4. Methods

4.1. Data Collection. The five participants were interviewed individually, and the interviews were audio recorded and transcribed. In an initial interview lasting an hour or more, each was asked first to speak in general about their experience with the course. They were then asked more specifically about their thoughts on:

- the most important things students should learn in the course.
- common student mistakes or misunderstandings, and why these arise.
- strategies they used in trying to teach this material, why they used these, and whether they found them to be successful.
- what they would like to see come out of this research.

It was clear that all of the participants had given considerable thought to the teaching and learning of proof. They offered many well-articulated conceptualizations of the problems they had encountered, and reflections upon their attempts to address these. However, these conceptualizations naturally varied from participant to participant, and the pedagogical approaches they reported were designed in response to individual interpretations of students' behavior. In addition, there appeared to be disagreement among the participants on some points. My goal was to synthesize the issues raised, so that each participant could recognize his or her own experience while relating this to a coherent overall framework. In order to accomplish this I analyzed the transcripts of these first interviews in the manner described by Glaser (1992), as detailed below.

4.2. Data Analysis. First, I added *conceptual descriptions* to each transcript. These were short phrases written to capture the primary conceptual content of each of the participant's comments. Concurrently, for each eight paragraphs of the participant's speech, I wrote a *summary memo*. These memos summarized the content of my conceptual descriptions, thereby providing a condensed version of the conceptual content of each interview. These descriptions and memos were written on a rotating basis, switching from one participant to the next after each summary memo, in order to facilitate synthesis of the ideas raised and to avoid becoming focused on the opinions of a single participant.

During this process, I also wrote *theoretical memos* about questions arising from the data. These typically contained my thoughts on points that needed clarifying, on possible theoretical relationships between issues raised, and on links to extant theoretical concepts and empirical studies. All memos were written in an informal style, the aim being to exhaust ideas rather than to produce a polished product (Glaser, 1992). However, I took care to indicate, through the language used, which ideas were grounded in the data and which were my own conjectures, questions or speculations. At this stage, all memos were kept in one file in the order in which they were written, and each was headed with a reference to the interview and line(s) that prompted its writing. An example of a section of transcript and my accompanying conceptual descriptions and summary and theoretical memos can be seen in the appendix.

Next, I sorted the file of memos. I identified the main substantive and/or theoretical content of each, and placed it in a new file under one or more short headings describing this content. Each memo added was assigned a new heading, or added under one or more existing headings, possibly prompting some reconsideration of the wording used for these. As this file became longer, these headings and their associated memos were progressively subsumed under higher-level headings. This resulted in a flexible system of categories that could be viewed at different levels, and in which lower level headings, together with their memos, could be moved around as necessary. A separate file was also kept of specific questions that I wanted to ask in subsequent interviews.

I found this to be a more satisfactory way of managing this exploratory data than straightforward “coding” of the transcripts. Such coding usually involves classifying data as it is read, and can tempt the researcher to assign short codes at an early stage, so that later data becomes skewed into prematurely-specified categories and the focus is upon labeling rather than upon the relationships between the concepts that the codes describe. The method described here separates memoing and sorting into two distinct stages, so that at the time of writing the initial summary and theoretical memos I was free to express each conceptual description and theoretical question as accurately as possible. The sorting then focuses on the conceptual and theoretical content of the memos, with the eventual category headings playing the role of short codes. Note that by the time these headings are fixed, they are automatically attached to references to the relevant data. Naturally, this does not mean that another researcher would use the same descriptions or ask the same theoretical questions. But it does avoid an early decision either to focus on some parts of the data at the expense of others or to prematurely specify terminology. It thus allows the data as a whole to drive the construction of the category system.

After this process was completed for the initial interviews, I designed second and third interviews based on the question file, and on reading the main file for any areas that seemed under-represented given their importance in the emerging category system. With one exception (Professor 4, who had retired), all of the professors took part in these follow-up interviews, which again lasted approximately one hour. I analyzed each new interview in a similar way to that described above: a file was created of summary and theoretical memos, and once it was completed these were sorted into the existing system, modifying this system where appropriate.

Finally, a two-hour seminar was arranged in which I presented my main findings to the four remaining professors and a colleague from mathematics education. This was an opportunity for all of the participants to respond both to each other's views and to my overall formulation of a theoretical framework on the basis of the data. The seminar was also audio recorded and the participants' comments analyzed.

4.3. Form of the Results. The participating professors spoke about many practical issues including ongoing feedback and assessment, the difficulty of finding an appropriate textbook, the lack of a coherent body of material in the course, and the need to maintain a certain level of emotional comfort on the part of the students. However, the discussions were primarily centered around how to account for the fact that students would frequently make errors in surprisingly simple logical situations, and how to help them to avoid such errors and make progress in writing correct proofs. As the analysis progressed, therefore, the category that emerged as core was “reasoning.” This was subdivided into four modes of thinking

that the professors, as a group, expected a person to engage in while proving, but with respect to which they often found their students to be lacking. Following language used by the professors, I term these modes *instantiation*, *structural thinking*, *creative thinking*, and *critical thinking*. This is a coarse breakdown, and I do not claim that the list is exhaustive, based as it is on self-reports of only five (albeit experienced) participants. Indeed, it was not the case that each professor placed equal emphasis on all four modes; each had developed a range of different emphases in thinking about these issues. However, each recognized all four in the seminar discussion, and none remarked that any major issue had been neglected. I therefore put the classification forth as appropriate for this paper’s aims of describing these professors’ experiences (their goals for student learning, their teaching strategies, and their concerns about these) and providing a framework for further discussion.

5. Results: Four Modes of Thinking

This section describes the four modes of thinking. Each is defined, as in Table 1, by its purpose for the prover, that is, by the interim goal during a proof attempt that its use addresses. To facilitate later discussion, each is also classified as semantic or syntactic, where I define *semantic* modes to be those that involve thinking about the mathematical objects to which a statement refers, and *syntactic* modes to be those that involve thinking about and manipulating a statement based on its form (for more on this classification, see Weber & Alcock, 2004).

In Sections 5.1 through 5.4, each of the four modes are described in detail through professors’ comments, which also serve to highlight what they would like students to learn in the course. As mentioned above, the classification into modes is coarse, so there are cases in which several different skills or strategies would contribute to the goal of one mode. For the sake of clarity the content of Sections 5.1 through 5.4 is restricted to presentation of interview excerpts. Commentary upon links with other research is postponed until Section 5.5 and teaching strategies are discussed separately in Section 6. Throughout the interview excerpts, “...” indicates a pause or hesitation and “[...]” indicates that one or more short phrases has been omitted.

TABLE 1. Four Modes of Thinking

Mode	Class	Purpose
Instantiation	Semantic	To meaningfully understand a mathematical statement by thinking about the objects (particular or generic) to which it applies
Structural thinking	Syntactic	To generate a proof for a statement by using its formal structure (making formal deductions based on the statement and/or associated definitions and known results)
Creative thinking	Semantic	To examine instantiations of mathematical objects in order to identify a property or set of manipulations that can form the crux of a proof
Critical thinking	Semantic	To check the correctness of assertions (by checking for counterexamples and for properties that are implied or should be preserved)

5.1. Instantiation. Instantiation is a semantic mode, the goal of which is to meaningfully understand a mathematical statement by thinking about its referent objects. I use the term broadly, to include thinking about generic examples or images as well as particular objects. The professors talked about instantiation as: a natural response to a new definition, an activity that is necessary for making sense of a new definition, a way to understand the implications of choosing a certain property as a definition, and a way of making the material seem less abstract and more accessible. They noted that students often do not spontaneously instantiate for such purposes and may resist inducements to do so. The majority of the quotations in this section come from Professor 1, who had given the most thought to this mode. Both creative and critical thinking build on the use of instantiations, and their respective sections include more comments from other professors.

Professor 1 described instantiation as a natural activity that is necessary for making sense of a new definition.

P1: So one of the things, again, that's second nature to me but it's not to them, is that if I see a definition, I immediately instantiate it. You know, just try some examples of this definition, and try to fit it in.

P1: What happens is, you know, that you describe a new definition, you say "let f be a function, let x be a real number, we say that..." and then "some relationship between f and x holds if ... blah blah blah." So then what they have to do, they have to realize that this definition only makes sense in the context of, I have to have a function in mind and I have to have a [number] in mind.

Professor 5 described the way in which instantiations (in this case, imagined sets of integers and apples) can be used to understand the implications of using a certain property as a definition.

P5: We started thinking through, um, you know, what are the implications of defining a set to be simply a collection of unordered objects? [...] So we think through, what happens with integers, and what happens with apples and different collections of things, and, um, you know if you have five apples in a set, is that the same as having one apple, for example?

Professor 4 described a way in which an instantiation can make the abstract concept of function seem more accessible.

P4: My guess is that the notion of functions between arbitrary sets is too abstract for them. I think they can't... picture it. I mean I keep telling them to try picturing a function as like shooting the elements of one set to the other set. Make it very physical. Because, if they can see these things, these things are not hard.

Regarding students' behavior, Professor 1 noted that his students often do not instantiate mathematical statements at times when this would be appropriate, and that he encounters resistance when he encourages them to do so.

P1: And what they'll do is typically if you have a sequence, you know, if I have a sequence definition to use in the rest of the problem, and they don't understand the definition, they'll just skip that sentence and go on. I will... they will come in for help on a problem, and five or ten minutes into the discussion I'll realize that, that they never bothered

to process this particular definition. They have no idea what this means.

P1: I also tell them, things like, before you try to prove a statement, let's say it's a "for all, something," then, you should look – illustrate. Write down a few illustrations of this theorem. So pick some examples, just so you can see how, how this theorem works. So maybe you don't see exactly how it's going to help you do what you have to do, but trust me, it will. [...] And they just refuse to do this. They just, they just won't do it.

5.2. Structural Thinking. Structural thinking is a syntactic mode, the goal of which is to generate a proof for a statement by using its form; that is, by introducing appropriate definitions and making deductions from either these or the statement itself according to the rules of logic. The professors talked about structural thinking as: allowing the logic to drive the steps of a proof, a way to approach proving systematically, a way to reduce the complexity of the proving process, a sensible first approach when trying to construct a proof, and a way to handle tasks such as negating statements involving multiple quantifiers. With regard to students' work, they noted that students often lack skills that are needed for effective structural thinking. These include knowledge of, and inclination to introduce, relevant definitions, ability to identify the logical structure of a sentence, ability and inclination to make mathematically correct interpretations of connectives such as "and" and "or," care in the precise use of symbols and rules, and ability to write one's own statements clearly. In this section we hear from all five professors, with particular focus from Professor 3 on the importance and advantages of systematicity and precision, and from Professor 2 on the errors students make in the use of mathematical language.

Professor 1 remarked that it is often possible to make progress on a proof using structural thinking.

P1: Particularly early on when you're doing proofs, there's a certain amount of steps of the proofs which are just sort of... driven by the logic. You know if you just, carefully write down the definitions of what you're doing, and you look at what you're trying to show, then you almost have to do certain things, as long as you've mastered the logic.

Professor 3 aims to teach students to systematically make use of this logical structure. He sees this as a way to make the task of proving easier by reducing the number of decisions that need to be made.

P3: So what I try is to have things very systematic. [...] We should really get to the point, where if I give you, if I give the students the definition of limit, "for all epsilon blah blah blah..." just mechanically we want to prove, for all epsilon, blah blah blah. There is only one rule that enables us to "prove for all epsilon, blah blah blah." Therefore there is only one way to start the proof. "Let epsilon be arbitrary, we will prove blah blah blah."

P3: My dream is always – I mean it seems to me that when one makes a thing very minimal in this way, this has a very positive side, which is

that...you know when there is only one thing you can do, then it's much easier to know what you can, what you should do.

With regard to the students' work, Professor 3 noted that students may not state definitions correctly and Professor 4 commented on the fact that they may not introduce definitions when this would be appropriate.

P3: For instance, Question: "Define even." Answer: "Even is when it's multiplied by 2." Things like that.

P4: I mean like for example...ah...show that two sets are equal. The only way you can really show that the – you give a definition, sets are equal if they have the same elements, okay? And...so that's what they have to show. And they sort of don't get that. They start with some formulas, they start some manipulations and they don't get the fact that...there's a definition.

Professor 2 commented on students' lack of experience in parsing sentences, and offered a variety of instances in which students' uses of connectives (in this case, "and" and "so") are inconsistent with their use in formal mathematics.

P2: There is no formal grammar taught. I – next to none of my students had ever heard the word "parse" used about English language. They'd heard it, some of them knew it from computer programming, but almost none of them had ever...they don't, they're unfamiliar with the names for parts of a sentence.

P2: I was once again surprised at how many kids don't hear the difference between... "the solutions to this equation are 3 and 4," and " x is a solution to this equation if x is 3 and x is 4."

P2: But students who would say, "we want to prove this, so that..." - it was never clear to me whether they meant "so we can conclude that," or "so we need to verify that." And there were a lot of proofs which would be correct, if I could only infer that they meant "so we need to verify...so we need to verify...so we need to verify...and then this last one is clearly true so everything stacks up."

Professor 3 commented on students' laxity regarding mathematical precision, in particular their tendency to write symbols incorrectly and to misquote given rules.

P3: When I tell you that the symbol is this, it's exactly that. Not the symbol turned around, or upside down, or – okay. If I tell you that the rule says this, like there is a rule that says that if a is less than b and c is less than d , then a plus c is less than b plus d , that's it. You cannot change the b plus d into b times d . [...] Whereas the students seem to think that there is an enormous amount of flexibility.

Finally, both Professors 1 and 3 commented on the way in which an absence of precision in students' writing impedes their own efforts to evaluate students' proof attempts.

P1: [I say] "So now I want you to read your argument, and I want you to identify – so there is a place in your argument where you lied to me. And there must be a critical place." [...] So of course at this point they're not writing proofs. They're not even writing arguments that

you can apply this kind of analysis to. I mean it's sort of so vague or so diffuse.

P3: Most of the proofs, ideas in the homework. . . I cannot even tell what's wrong with them. I cannot even pinpoint the step where it goes wrong.

5.3. Creative Thinking. Creative thinking is a semantic mode, the goal of which is to examine instantiations of mathematical objects in order to identify a property or set of manipulations that can form the crux of a proof. The professors spoke of two ways to do this: a direct method in which one experiments with an example in the hope of finding an argument or sequence of manipulations that will generalize, and an indirect method in which one attempts to build a counterexample and tries to identify a reason why this cannot be done (an informal analogue of proof by contradiction). Once again, the professors' comments indicate that they consider such thinking natural, but that their students do not use this mode effectively. Reasons include a lack of inclination to instantiate (described in Section 5.1), incorrect understanding of the kinds of objects for which certain properties might hold, and the tendency to rely on empirical rather than property-based arguments. Professors 1, 2, 3, and 5 made comments of this type.

The professors often cited the direct method as a strategy for constructing an induction argument, as in this comment from Professor 5.

P5: See if you can get from 2 to 3, if you can't get from n to n plus 1.

The indirect method was described in the context of universal statements, by Professors 1 and 2.

P1: The way I often think about a proof is that, you know you imagine this as, try to beat this. Meaning, try to find a counterexample. [...] If you think about the reason why you were failing to find a counterexample, okay, then, that sometimes gives you a clue, to why the thing is true.

P2: Then if it's an existential statement I look to see whether I can produce an example. And if it's a universal statement I probably try to show that I can't find a counterexample.

When the interviewer commented that it seems non-obvious that one would try to prove a statement by thinking about how it could never not hold, Professor 1 explained why he thinks this a natural way to proceed.

P1: So why do I think that this is a natural way to go about it? Um. . . well I, I'll make a statement which I haven't thought about before. The natural, the sort of natural thing that our brains can do, is sort of build examples and check them. Okay, and. . . all you, you know if one thinks of universal statements as saying that it's really a statement of impossibility, it's the negation, right? It's a statement that you can't. . . do something. And the way you understand that you can't do it, is by thinking about doing it.

Regarding students' work, both Professors 3 and 5 cited instances in which a student did at least cursorily examine instantiations of appropriate objects (numbers, in these cases), but failed to make an argument in terms of properties. In

Professor 3's case, the student had been asked to establish that 1997 is a prime number. In Professor 5's case, a student was addressing an examination question requiring a proof that if the sum of two primes is odd, then one of the primes must be 2.

P3: So then I asked her in class, "How do you know that 1997 is prime?" The idea being that she would say "Well I checked," in which case I was going to answer, I was going to say "Okay, you tell me that you have checked. But why should I accept that?" Okay? That was going to be the second part, but then I didn't even get there because then she said, "It looks like a prime to me." That's what she said.

P5: I mean there was one student who just added...thirty things together...No, excuse me, she added 2 – she added 2 to every other integer, and got something prime. Got something odd, or prime or something like that.

5.4. Critical Thinking. The goal of critical thinking is to check the correctness of assertions in a proof. In theory this may be achieved syntactically, by checking the form of the deductions and perhaps providing subproofs (see, for example, Selden & Selden, 2003). However, the professors in this study rarely made comments that related to such a process. Rather, critical thinking emerged as a semantic mode involving searching for possible counterexamples, checking for implied properties that are false and/or checking for properties that should be preserved. Regarding students' work, the professors' comments indicate that students may not perform checks using any of these methods, and they express some exasperation about this. This section includes comments from Professors 1, 2, 3 and 5.

Professor 5 described the process of checking for potential counterexamples as she tries to teach it to her students.

P5: I taught them to, you know, [make a] mathematical claim, [then] stop. [...] Think through examples. Have a couple of examples, and then try to see if there's something fishy about them, that might not extend to all possible integers. Ah...look for a counterexample.

Professor 1's comment indicates that in his own work, the choice of what to check is based on a somewhat sophisticated classification within the domain of referents.

P1: I guess I classify numbers into numbers which are prime, let's say, prime powers, and then there are sort of, "far from [prime]," you know so they have a few factors. But then they should have...but then there are also square-free numbers. [...] So I have a sort of classification in mind of different kinds of integers, and I just make sure that whatever I'm trying, I'm trying with these different kinds.

Professor 2 described the way in which her sensitivity to implied properties that cannot be true manifests itself.

P2: [If] I've said something or written something that is wrong, I will get a sense of disquiet. [...] Part of it may be that at a semi-conscious or unconscious level, I'm in the habit of looking for implications of what I've written down. And if those implications are sufficiently absurd, then it gets up into the conscious level.

Professor 3 commented upon checking for preservation of properties in the following remarks about a student's mis-writing of the binomial formula as

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^k.$$

He first noted that there is no doubt that the student had worked hard to memorize the formula, before continuing:

P3: Of course, anybody who has a moderate amount of understanding of the question, should ah... realize for example that if you multiply a and b by 2, the right side should be multiplied by 2 to the n , and that – doesn't quite happen. In other words, there is something very wrong about those two powers k in there. It just cannot be.

This last is arguably a rather sophisticated way of checking the correctness of such a formula, and perhaps we would not expect students to use it often. However, the professors regularly noted that students often appear not to perform checks using any of these methods. Professor 3 gave the following instance in which checking for counterexamples would reveal a flaw in the argument.

P3: For instance, problem: "Express the number 30 as the difference of two squares, or show that it cannot be done." Answer: "It cannot be done because 30 is divisible by 6 and a number that is divisible by 6 cannot be written as the difference of two squares." Well, 12 is 16 minus 4. Ah... take any number that's a difference of two squares, multiply it by 36, you'll get a number that's the difference of two squares and is divisible by 6, so I mean... again I could give you loads of examples of the same kind.

Professor 3 expressed exasperation at such responses, considering them "obviously false" on the basis that counterexamples are readily accessible.

P3: I don't have a clue as to, what gets them to, to say things like that. In other words I would say, things that are obviously false. To a normal person with a little bit of mathematical education it would seem obvious that you could never say such a thing because it's so obvious that it's false. Take any example that you want, you see clearly that it's false.

5.5. Relationship to Other Research and Theories. We have now seen many comments on skills that the professors think students need to develop in order to produce mathematical proofs effectively. Many of their concerns are reflected in the mathematics education research literature, and in this section some specific links are discussed.

Others have observed that instantiation is used by mathematically successful individuals. Dahlberg and Housman (1997) note that more successful students spontaneously generate examples in response to a new definition, and in Weber and Alcock (2004) we observe that mathematicians are able to instantiate concepts like group and isomorphism in many ways. Pedagogically, Watson and Mason (2002) argue that student-generated examples can allow students to experience the range of variation in a concept and explore its boundaries, and Winicki-Landman and Leikin (2000) and van Dormolen and Zaslavsky (2003) use examples to explore the implications of basing instruction on different choices of definition. The observation

that students often do not instantiate is also consistent with the considerable body of research on students learning mathematics as meaningless manipulations. It is well recognized that a student who knows “what to do” (Skemp, 1976) can often correctly complete a wide variety of algebraic calculations without considering their referent objects (Cerulli & Mariotti, 2003; Sfard & Linchevsky, 1994).

The component skills for structural thinking are widely discussed. In Alcock and Simpson (2002) we observe that students may not readily interpret “show that x is an X ” to mean “show that x satisfies the definition of X ,” and it is recognized that they often reason instead in terms of their concept images (Edwards & Ward, 2004; Vinner, 1991). It is also recognized that in everyday situations, context can play a large role in allowing an individual to interpret a statement containing constructions such as “if... then...” or “for all... there exists...” (Dubinsky & Yiparaki, 2000; Epp, 2003; Zepp, Monin, & Lei, 1987) and that this can interfere with students’ abilities to make use of logical structures in constructing and validating proofs (Selden & Selden, 1995, 1999). It is worth noting that effective structural thinking is “syntactic” in the sense in which we use the term in Weber and Alcock (2004) – to refer to correct reasoning based on form and logic – and not as this term is sometimes used to refer to the meaningless manipulation of symbols or ritual use of a certain format.

The skills associated with creative thinking are not so widely discussed in the literature. There are many reports about students who accept or offer perceptual or empirical evidence without going on to abstract some property and construct a general argument (Chazan, 1993; Harel & Sowder, 1998), but there is less discussion of how they might make this next step. Certainly the issue is considered important; Raman (2003) observes that mathematicians sometimes think about proofs in terms of “key ideas” that serve to link their formal and informal understanding, and Rowland (2001) suggests specific principles for using generic examples to help students understand and construct general proofs. However, I would contend that in general we who teach mathematics tend to believe that the creative aspect of proof construction, the “having a good idea,” is to some degree a matter of talent and is less open to teaching than other modes. Rowland’s work and the indirect strategy suggested by the mathematicians in this study suggest that this need not be the case, and I believe that this might be a fruitful avenue for further investigation.

As noted in Section 5.4, although the goal of critical thinking might in theory be achieved syntactically, in this study it appeared as a semantic mode, relying heavily on consideration of an assertion’s referent objects and their properties. Checks based on preserved or implied properties are not widely discussed, though I would speculate that such an approach requires a more sophisticated and organized knowledge base than checking particular example objects. Counterexample checking is discussed in the psychological literature; for example, Johnson-Laird and Hasson (2003) report that people commonly use it to evaluate inferences based on statements about everyday situations. Research in mathematics education indicates that students do not necessarily make such checks when validating mathematical proofs, although relatively straightforward questions can prompt an improvement (Alcock & Weber, 2005; Selden & Selden, 2003). Part of the problem here may be that of deciding which objects might potentially serve as counterexamples. As we argue in Weber and Alcock (2005), evaluating a statement like “1007 is prime because 7 is prime,” requires one to infer a warrant (in the sense of Toulmin, 1969) such as the

general statement “If x is prime then $1000 + x$ is prime,” and evaluate it across the range of possible prime numbers. This highlights the complexity of counterexample checking, especially given the prior structural question of identifying a conditional statement if not expressed in the standard “if...then...” way.

6. Teaching Emphasis

This section discusses teaching strategies used by the participants and their reflections on these. It is organized so as to highlight which strategies may be thought of as addressing which mode, although the participants did not divide up their discussions in this way. The principal observation is that the majority of strategies are directed at developing the skills needed for effective structural thinking (the other three modes are therefore discussed first, departing from the previous order). The participants were aware of this emphasis and expressed mixed feelings about it. Their opinions are discussed in Section 6.5.

6.1. Teaching Strategies: Instantiation. Of the teaching strategies discussed, few seem to pertain directly to instantiation. Those that do involve informally encouraging students to examine particular instantiations, and in one case (Professor 1), use more formal tasks that require instantiating a range of objects with different properties.

P4: I mean I keep telling them to try picturing a function as like shooting the elements of one set to the other set. Make it very physical.

P1: So, what I’ve been trying to do is to have these exercises where the whole purpose of the exercise is just for them to process a mathematical definition. [...] I have one where I, where I just define what [...] a partition of a set means. I define it formally [as] a collection of subsets, such that [1.] the empty set is not one of the subsets, [2.] for every element of the underlying set there is a subset that contains it, [3.] for any two sets in the partition the intersection is empty. [...] And then I just ask okay, construct three examples of a partition on the set $\{1, 2, 3, 4, 5\}$. And then, okay, construct an example of a collection of sets on $\{1, 2, 3, 4, 5\}$ which satisfies the first two properties but not the third. The first and the third properties but not the second, the second and the third properties but not the first.

6.2. Teaching Strategies: Creative Thinking. The five professors cited few pedagogical strategies that seem to pertain directly to creative thinking. One device that might provide an appropriate context is an “exploration” phase preceding a proof attempt, as used by Professor 2.

P2: Getting them to write down that thing following the caption “analysis” before they have to write down something following the caption “proof” seems to help. Giving them permission to just “think out loud,” without having to jump immediately to a properly structured proof seems to help.

However, when asked about the specifics of what a student would do in such a phase, the professors found it difficult to articulate any general principles. Even Professor 1, who described the heuristic of trying to build a counterexample and

attending to why this is impossible, noted that he had not systematized this in the classroom.

P1: I talk to the students about this somewhat, although I haven't quite figured out how to make it more systematic. But, the way I often think about a proof is that, you know you imagine this as, try to beat this. Meaning, try to find a counterexample. . .

6.3. Teaching Strategies: Critical Thinking. Teaching that pertains to critical thinking was described more regularly and systematically, at least as regards searching for possible counterexamples. Most of the professors commented that they model this search in class, and in one case this is institutionalized by designating a student whose job it is to try to give such counterexamples (Professor 3 reported that he assigns this role to a specific student, although he does not require the writing that he described):

P3: There is supposed to be this person over there that's designated "cat" - "creator of arbitrary things," for the semester. The cat's job, is that every time there is an arbitrary thing, he picks one, and writes it down somewhere, and puts it in a sealed envelope and we don't know what it is. [...] The cat's job, this is made very clear, the cat's job is to try to prove me wrong.

Another strategy that pertains to critical thinking is having students consider false statements and state what it would take to exclude counterexamples. Again Professor 3 goes further, deliberately including such statements on assignments with the injunction that students should always be alert to the possibility that something they are asked to prove could be false.

P1: And so, on the first homework assignment one of the questions gives six or seven universal propositions, and these are all false. So find a counterexample, and if possible, find a simple modification so, that makes it true.

P3: [I] give them lots of problems which they call trick questions. [...] We have a general principle that ah...every time you are asked to do something, either you do it, or you say that it cannot be done and you give a reason why.

The professors were also concerned that students should be alert to implied properties, although the more hesitant nature of these comments suggests that this is addressed less systematically.

P2: I guess sometimes when we're trying to construct a proof, and they suggest something which I think goes in the wrong direction, my reaction is that "Wouldn't that imply something that...something seriously wrong?"

6.4. Teaching Strategies: Structural Thinking. The professors' teaching strategies aimed at structural thinking were considerably more numerous and well-articulated, suggesting that the participants use them more systematically. Some involve rather informal modeling of the introduction of appropriate definitions and suggesting that this be done in an "exploration" phase.

P4: If I were to do those in class, the idea would be, how do we approach a problem like that? Well if f is injective, how do we show f is injective? Well we say, let's go back to the definition. If f of a is equal to f of a -prime, then we have to show that a is equal to a -prime.

P2: They are encouraged to write in rather free-form all the relevant definitions and all the apparently relevant theorems, in whatever order they come. And then when they see...think out from that maybe they can see how to put a proof together.

However, strategies pertaining to structural thinking are predominately based on a variety of writing guidelines or "rules." These guidelines pertain to correctly formulating definitions, to correctly introducing mathematical objects, to laying out proofs according to certain templates and/or clear general formats, and to restricting the use of certain terms. Examples of comments on each are given below.

On correctly formulating definitions:

P3: Every time you're asked to define [...] something...the first thing you need to ask yourself, because it's going to matter for what you write, is what kind of a predicate, that is how many arguments it has. For example, if you want to define prime, prime is about one number, it's something that can be true or false for one integer, and therefore you should start by saying "Let x be an integer. We say that x is prime if..." And then comes the condition, right? Whereas ah...divisible is a two-variable predicate. So you would start with "Let x and y be integers. We say that x is divisible by y if..." and then comes the mathematical content.

On correctly introducing objects that will be used in a proof:

P1: If you want to talk about a letter in more than one sentence, then the first sentence in which that letter appears must be "Let x be a blah." And I give them a bunch of rules, which are that they have to introduce them, they have to give everything a different name, to introduce it you must list all the properties that that thing has, you can't later on sneak in something.

On "templates" for writing certain types of proof:

P5: [I] put out a template for writing inductive proof. You write out the inductive hypothesis, then you write out, "Show P of n implies P of $n + 1$." So it's there in front of you. And then you, somehow start with the left hand side of P [of] $n + 1$ and then you make logical deductions and you use the inductive hypothesis. And you work at it, and you arrive at the conclusion of P [of] $n + 1$.

On writing of proofs in general:

P5: I gave them some very...general guidelines for proofs. [...] One mathematical claim per line. Um...and, and at the very beginning, to write every logical step from one argument to the next. So that you can go back and check if the logic is correct.

On restricting the use of certain words or phrases:

P3: For example, never use the word “any,” because in mathematics it tends to mean too many things. “Any” can sometimes mean “some,” it sometimes means “all” . . . Now the good student would know how to use the word “any” correctly. But a not very good student is likely to use it in a way that is ambiguous, so better never to use it.

P2: It’s different from high school English, in which one is warned never to say “he said” and “she said,” you have to use a much more precise synonym than “said”. Whereas in mathematical English I’d like them to just use “implies,” and stay away from all the synonyms for implies.

6.5. Reflections on Strategies for Structural Thinking. The professors expressed a variety of opinions about the use of guidelines and rules to support structural thinking. Professors 1 and 3, in particular, had long debated the nature of the course, with Professor 3 believing strongly in a systematic, rule-based approach and Professor 1 believing the approach should be more investigational, but feeling compelled to introduce guidelines in response to student errors. Epp (2003) discusses similar debates. Each professor, however, had seriously considered the arguments for the opposing view, and the other professors also viewed this as a complex issue. This section has a different tone as at this point I believe it is more illuminating to move from a synthesis to a contrast of opposing points of view.

Professor 3 introduced the idea of a very systematic approach, and cited one of its advantages, as follows.

P3: My own idea as I said is to try to, to tell them. . . to try to get them to operate in a very systematic way, and systematically tell them that it’s for their own good. For example, one of the purposes of having all these rules, these very precise rules, is that you know that a proof is supposed to be an argument that you give, that should convince me.

He considers it reasonable to ask for this level of rigidity, since students have not previously learned to communicate with the precision needed for mathematical proofs.

P3: And I insist very much on minimizing the number of rules because I have become convinced that as an antidote to the general . . . amorphousness of their work, of the work of their ideas, there is a need to put a lot of rigidity. [...] Maybe it doesn’t hurt to go to the other extreme for a while. Once, in one course.

He acknowledged that mathematicians do not think entirely in this way, but considers it unrealistic to hope to convey this in the time allowed.

P3: Doing proofs, mathematical proofs, is not a matter of analyzing logical structures and statements, and applying the logical rules of inference. [...] That may be the frame, the, the thing that supports the proof. But the real proof and the real mathematics is not that. And that should also be conveyed in the course. But of course it’s too much. That’s wishful thinking.

He also acknowledged arguments in favor of a more investigation-based approach, but considers this unrealistic given students’ tendency to give inadequate consideration to their claims and justifications.

P3: I have doubts, because even when I put moderately challenging problems, what I always find is... that the main obstacle is not that they don't know what to do, the main obstacle is that they will do anything... and don't seem to think that it needs justification.

Professor 1, in contrast, originally began the course with the intent of structuring it around investigations of “interesting problems where the answer is not obvious.” He considered a rule-based approach to be unrepresentative of the way that mathematicians think when they write proofs.

P1: Although we know in principle that it can be done syntactically, we can bring it down to some axiomatic system... that, in fact, you think very semantically. I mean you know sort of... you're thinking about the objects themselves, and you are confident with your ability to do... to do legal reasoning, on this thing. [...] And then you have some pictures of it, and you are confident of your ability to translate those mental pictures into things that we all accept as real mathematics.

He argued against treating proving as a syntactic enterprise.

P1: [Professor 3] wants them to understand that proof is basically a syntactic thing which we then, sort of, as we become more sophisticated we take shortcuts, but he wants them to really understand the syntax of what they're doing, and to work on writing it. My problem is that it's more of the same. In other words, it encourages the students to think of mathematics as a syntactic enterprise, rather than trying to assign meaning to what they're doing.

However, Professor 1, too, has felt compelled to issue guidelines in response to student errors (such as that quoted in the previous subsection regarding correct introduction of objects to be used in a proof, which he hopes will head off “scope” errors in which a student uses the same letter to refer to two potentially different objects in a single proof). While he hopes these rules will help students to write good mathematics, he considers their proliferation to be a problem.

P1: Every time the students demonstrate to me that they, that some completely natural logical thing to me, that I couldn't imagine how anybody could fail – could make this kind of logical error. But they make it. Okay. So what that does is it causes me to institute some rule that I announce to the class. Not being able to do this, or must do this or something like this. [...] The problem is the more, the longer that I teach the course, the more of these I end up putting in.

As a caveat, it should be stressed that neither of the above professors applies their rules with absolute rigidity. Both allow for more flexible use of mathematical language, provided that a certain level of competence is demonstrated.

P1: And I tell them that, as you demonstrate the maturity to me that you understand these, then we'll let you relax those rules. And since it's a small enough class, well I know your work, and I know which students have earned the right to relax the rules.

P3: So I give you the rules, and that means that I am making a commitment to accept then what you do, without questioning it, as long as

you obey the rules. [...] If you don't do it that way, then it might still be okay, but now you are taking a risk.

Other professors, too, expressed mixed feelings about a rule-based approach. There was a concern that proof templates can be useful but that students can come to rely excessively on them. Professor 5 commented on this.

P5: I don't think it's an open and shut case. I think that it's useful sometimes to bring in [...] a set of guidelines, or... almost a template, for writing a proof. And there is a kind of a template for proof by contradiction, and there is a template for induction, and there is a template for... proving things using this method of smallest counterexamples and the well-ordering, property. [...] I think that it's very instructive to... to lay these out and repeat them, so that the students have a starting point, and are not completely intimidated and put off and, to the point where they can't even start. But if you overdo it, you're going to end up with... sort of robots.

Similarly, Professor 2 expressed concerns about giving the impression that there is a unique right way to proceed in writing a given proof.

P2: I really got worried about the fact that I was giving the one – the impression that there was only one right way. Especially in proofs where if you follow your nose there is a, a single energy-minimizing path.

In addition to these philosophical issues, there were also concerns about the effectiveness of a teaching approach that relies heavily on writing guidelines. One problem is choosing and sticking to rules, as in this comment from Professor 2:

P2: One of the things is to show students what it means to articulate a bunch of starting facts, and a scheme for deductive reasoning, and to use those facts to investigate problems. [...] Inevitably we end up telling them they're allowed to use everything they know about some previous material. And "everything they know" is not a well-defined body of facts. So we get promptly confused as to what, what have the status of axioms.

Another problem is the difficulty of sticking to precise use of terms in one's own speech.

P2: For mathematicians, much of the time, when they say "In general you can't count on something being true," it means there exists at least one counterexample. Whereas "In general you can't count on blah blah" in the ordinary language means that, somehow more than half the time it's wrong. And... I mean I've found myself making that mistake even though I've noticed that it's a dangerous one.

Finally there were questions about the degree to which students are able to use the rules effectively in constructing proofs. Professor 2 noted that what is learned in abstract discussions of predicate calculus may not readily transfer to contexts with more mathematical "substance."

P2: What I'm learning is difficult is moving from stuff which has just the form of predicate calculus, to real live sentences in mathematics. Where the structure of the sentence is hidden, a little – the student

gets distracted by the substance of the mathematics, from seeing the structure.

In all, it seemed that all the participants use some rules or guidelines as scaffolding to help their students to avoid logical errors and to write with the required level of precision. However, they had both philosophical and practical concerns about the efficacy of this approach.

7. Discussion

So far, I have treated the four modes of thinking separately, covering their meanings, associated teaching strategies, and professors' concerns about their teaching and learning in the classroom. In this section, I argue that the four modes may sensibly be viewed as interdependent and that successful provers switch flexibly between all four according to their changing goals within a proof attempt. I therefore suggest that in teaching proof we should balance attention to the four modes, highlighting not only how a person is thinking at a given time, but also what they hope to accomplish by thinking in that way.

7.1. Interdependence of the Four Modes. First, instantiation is arguably the basis for any semantic work. One needs to instantiate in order to have objects to think creatively about and in order to subject inferences to counterexample checks. Instantiation is not strictly necessary for structural thinking, but I would argue that it can help to clarify the structure of a statement, especially in cases involving suppressed quantifiers (Selden & Selden, 1995). For example, in a recent real analysis class, I had a student argue that the statement “Every bounded sequence contains a convergent subsequence” is false because the sequence $1, 0, 1, 0, 1, 0, 1, \dots$ has a subsequence $1, 1, 0, 1, 1, 0, \dots$ that does not converge. Discussing the fact that it also has a subsequence $1, 1, 1, 1, 1, \dots$ that does converge helped to clarify the fact that the statement contains an implicit existential quantifier.

In the opposite direction, some of the skills associated with structural thinking are necessary for instantiation. One has to decide what to instantiate, and doing so involves being able to identify what objects a statement is about and what it says about those objects. Structural thinking also supports creative and critical thinking in similar ways. Further, when writing down the results of creative thinking, one needs to be able to formulate these in appropriately precise language.

Creative thinking does not so much directly support the other modes as fill gaps in what those modes can achieve. Doubtless there are cases in which a competent structural approach could generate an entire proof and critical thinking could adequately check its validity. However, as Weber (2001) argues, in many situations there are numerous correct deductions that could be made from a given premise and numerous reasonable approaches that one could try (direct proof, indirect proof, various proofs by cases, etc.), but few that would lead to a proof. In general, an individual needs some mechanism for deciding which direction to pursue. Examining instantiations can facilitate such decisions.

Critical thinking is also not strictly necessary for any given proof attempt. It might be possible to produce a proof either structurally or creatively without doing any sort of checking. However, mathematicians would surely advise against this, especially for students who are still learning a mathematical topic. Particular

risks are that creative thinking can easily lead to over-generalization because instantiations often incorporate properties that do not hold for the whole set under consideration. Similarly, formal rules can easily be confused; for a true conditional statement, the contrapositive is always true but the converse need not be. Critical thinking can help to catch potential errors from such sources.

7.2. Pedagogical Implications. Based on the reasoning of the previous section, I argue that each of the four modes of thinking described in this paper is important in the construction of proofs and should be taught in introductory proof courses. However, this leaves open the question of how this should be achieved. Should we teach instantiation first, in order that students develop facility in meaningfully understanding statements? Should we teach structural thinking first, since students will clearly need to parse statements correctly in order to work sensibly in the other modes? The latter is implicit in much of the teaching that is discussed in this study, as well as in the majority of textbooks designed for such courses. However, I would suggest that in separating out any of the modes and attempting to teach it in isolation for some length of time, we are neglecting an important part of the thinking of successful provers – that is, the flexible use of all four modes in response to changing demands during a proof attempt. To illustrate the significance of this, I offer the following outline of a hypothetical proof construction for the statement that if the sum of two primes is odd, then one of the primes must be 2 (a question from the course, as mentioned by Professor 5). This is not intended to be a prescription or a description of an actual attempt. Not every step is necessary and an experienced mathematician or student might skip steps or move very quickly through some, might make similar steps in a different order, etc. However, I believe that it is a reasonable characterization of how a competent student at this level might sensibly respond.

- Step 1, Structural:* Identify the statement as a conditional with premise “the sum of two prime numbers is odd” and conclusion “one of the primes must be 2.”
- Step 2, Instantiation:* Experiment with examples, confirming that adding two primes other than 2 gives an even number ($5 + 7 = 12$, $3 + 11 = 14$, etc.). Perhaps also try adding 2 to other primes, to get a feel for the fact that this gives odd numbers ($5 + 2 = 7$, $3 + 2 = 5$ etc.).
- Step 3, Critical:* Consider larger primes to confirm that this still holds ($29 + 19 = 58$, $53 + 17 = 70$, $53 + 2 = 55$ etc.).
- Step 4, Structural:* Set up the first line(s) of a proof, writing “Let x and y be prime numbers, and suppose $x + y$ is odd, that is, for some integer n , $x + y = 2n + 1$.” Experiment with manipulating this equation ($x = 2n + 1 - y$ etc.) and decide that there is no obvious way forward from here.
- Step 5, Creative:* Try to construct a case in which neither x nor y is 2 but their sum is odd. Notice that this is impossible because all primes other than 2 are odd, and that adding two such together will always give an even number. Notice also that 2 is even so that when added to another prime the result is odd.
- Step 6, Critical:* Check that there are no other primes that are even, so that no others would also allow an odd sum.
- Step 7, Structural:* Formulate the first part of argument in appropriate mathematical language, writing “If $x + y$ is odd, then x and y cannot both

be odd.” Perhaps justify this by adding “since if x and y are both odd, then $x = 2a + 1$ and $y = 2b + 1$ where a and b are integers, so $x + y = 2a + 1 + 2b + 1 = 2(a + b + 1)$, which is even.”

Step 8, Structural: Formulate the second part of the argument in appropriate mathematical language, writing “So one of x and y must be even, so one of x and y must be 2, since 2 is the only even prime.” Perhaps justify the last statement by adding “since if $c \neq 2$ is even then 2 divides c , so c cannot be prime.”

Step 9, Critical: Check each deduction in the proof.

An initial focus on the skills needed for structural thinking does not seem to be misguided for several reasons: students need to learn the linguistic conventions of mathematics, there are many simple proofs at this level that can be tackled entirely in the structural mode, it is impossible to correctly prove any statement without understanding its structure, and teaching syntactic validation of proofs may avoid potential confusion about when checking a small number of examples is an appropriate action. However, if the classification into the four modes of thinking is taken as a reasonable basis for further discussion, then the argument above raises the question of the degree to which, in the longer term, any approach to teaching proof:

- (a) is balanced with respect to the four modes and
- (b) fosters the integrated and flexible use of all four.

It certainly seems reasonable to claim that collaborative classroom environments, in which students investigate, refine, and prove mathematical conjectures, address both of these issues reasonably well. Indeed, such environments may directly impact a number of the failings cited by the professors here because research indicates that in such situations students will find meaningful ways to understand statements and to check their own and others’ assertions (Alibert & Thomas, 1991; Blanton & Stylianou, 2003; Rasmussen, Zandieh, King, & Teppo, 2005; Yackel, Rasmussen, & King, 2000). However, this does not mean that the potential is always maximized in such classrooms. Nor does it mean that a lecture-based presentation could not provide effective instruction in all four modes. I view this as important because my position as a mathematics educator is not that I would like to see all mathematicians drastically alter their teaching practices. Rather, I would prefer to see people think carefully in evaluating the ways in which their own preferred instructional methods can be used to maximum advantage and to consider alternatives if and when they identify something that cannot be done well within their current range of approaches. I hope that this framework provides a coherent way of making this evaluation of introductory proof courses. I also suggest that it could provide a basis for metacognitive discussions with students about what they need to do in order to successfully construct and validate mathematical proofs, and about how the tasks that they are asked to complete should contribute to this. In particular, I suggest that an emphasis on the purposes of the modes might help students to distinguish between stages of proving activity at which instantiating examples can be useful and stages at which it is insufficient.

8. Directions for Future Research

In addition to the above pedagogical questions, this study also raises questions for future research. For me the most interesting of these is the question of what

might account for a perceived split in the student population, as described by Professor 4:

P4: Basically the class consists of two groups. There are groups that understand it, and probably hardly need it, and then there are those who... ah... really need it, and are not learning it.

This is a controversial comment, but I consider it important to include it as several of the professors spoke of similar observations. Neither I nor they take it as a reason to “give up,” and as the evidence above suggests, all those interviewed were committed to finding ways to help all of their students improve their mathematical reasoning as much as possible. However, it is a reality that students do fail transition courses despite having succeeded in previous mathematics, and even the most dedicated teacher may suffer from doubts regarding what they are actually able to teach:

P3: I think there are many people in this department and in every math department who think that those things [how to write definitions correctly, that statements may be proved but noun phrases may not, etc.] are somehow so obvious that anybody who is not a complete moron and has some grasp, some ability to do mathematics should be able to see that. Therefore it is a waste of time to teach that because, the people who need to be taught are the people who are hopeless anyhow. [...] And I would have been vehemently opposed to that idea a few years ago. And now I'm less sure because you know... I'm certainly trying to teach that and I cannot... say that it works, so... maybe these people have a point!

It is clearly important that we continue to improve our understanding of factors that contribute to student difficulties with mathematical proof. My own hypothesis on completing this study was that some students who normally might fail this course could improve their performance considerably by learning to use instantiations effectively. Not doing so would be consistent with research suggesting that in earlier mathematics, students can be successful by manipulating mathematical notation without thinking about its meaning (Sfard & Linchevsky, 1994). At this level, attempting such an approach would automatically cut a student off from all three of the semantic modes of thinking described above. With this in mind, a colleague and I conducted research to investigate whether a tendency to instantiate forms a consistent distinction between the work of those who fail and those who succeed in Introduction to Mathematical Reasoning. Preliminary results from this research (Alcock & Weber, 2005; Weber, Alcock, & Radu, 2005) suggest that the situation is more complicated: there were both successful and unsuccessful students among those who tended to instantiate and those who did not. We are continuing our work on this data in order to better identify factors that contribute to successful use of both syntactic and semantic modes of thinking during proof construction.

References

- Alcock, L. J., & Simpson A. P. (2002). Definitions: Dealing with categories mathematically. *For the Learning of Mathematics*, 22(2), 28–34.
- Alcock, L. J., & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. *Journal of Mathematical Behavior*, 24(2), 125–134.

- Alibert, D., & Thomas, M. (1991). Research on mathematical proof. In D. O. Tall (Ed.), *Advanced Mathematical Thinking* (pp. 215–230), Dordrecht: Kluwer.
- Blanton, M. L., & Stylianou, D. A. (2003). The nature of scaffolding in undergraduate students' transition to mathematical proof. In N. A. Pateman, B. J. Dougherty & J. T. Zilliox (Eds.), *Proceedings of the 27th conference of the International Group for the Psychology of Mathematics Education held jointly with the 25th annual conference of PME-NA, Vol. 2* (pp. 113–120). Honolulu, HI: University of Hawai'i.
- Cerulli, M., & Mariotti, M. A. (2003). Building theories: Working in a microworld and writing the mathematics notebook. In N. A. Pateman, B. J. Dougherty & J. T. Zilliox (Eds.), *Proceedings of the 27th conference of the International Group for the Psychology of Mathematics Education held jointly with the 25th annual conference of PME-NA, Vol. 1* (pp. 181–188). Honolulu, HI: University of Hawai'i.
- Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics*, 24, 359–387.
- Dahlberg, R. P., & Housman, D. L. (1997). Facilitating learning events through example generation. *Educational Studies in Mathematics*, 33, 283–299.
- Dubinsky, E., Elterman, F., & Gong, C. (1988). The student's construction of quantification. *For the Learning of Mathematics*, 8(2), 44–51.
- Dubinsky, E., & Yiparaki, O. (2000). On student understanding of AE and EA quantification. In E. Dubinsky, A. H. Schoenfeld, & J. Kaput (Eds.), *Research in collegiate mathematics education. IV* (pp. 239–289). Providence, RI: American Mathematical Society.
- Durand-Guerrier, V. (2003). Which notion of implication is the right one? From logical considerations to a didactic perspective. *Educational Studies in Mathematics*, 53, 5–34.
- Edwards, B. S., & Ward, M. B. (2004). Surprises from mathematics education research: Student (mis)use of mathematical definitions. *American Mathematical Monthly*, 111, 411–424.
- Epp, S. S. (1998). A unified framework for proof and disproof. *The Mathematics Teacher*, 91, 708–713.
- Epp, S. S. (2003). The role of logic in teaching proof. *American Mathematical Monthly*, 110, 886–899.
- Gholamazad, S., Lijedahl, P., & Zazkis, R. (2003). One line proof: What can go wrong? In N. A. Pateman, B. J. Dougherty & J. T. Zilliox (Eds.), *Proceedings of the 27th conference of the International Group for the Psychology of Mathematics Education held jointly with the 25th annual conference of PME-NA, Vol. 2* (pp. 437–444). Honolulu, HI: University of Hawai'i.
- Glaser, B. (1992). *Emergence vs. forcing: Basics of grounded theory analysis*. Mill Valley, CA: Sociology Press.
- Harel, G. (2002). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. R. Campbell & R. Zazkis (Eds.), *Learning and teaching number theory: Research in cognition and instruction* (pp. 185–212), Westport, CT: Ablex.
- Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in*

- collegiate mathematics education. III* (pp. 234–283). Providence, RI: American Mathematical Society.
- Johnson-Laird, P. N., & Hasson, U. (2003). Counterexamples in sentential reasoning. *Memory & Cognition*, 31, 1105–1113.
- Leron, U. (1985). A direct approach to indirect proofs. *Educational Studies in Mathematics*, 16, 321–325.
- Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics*, 27, 249–266.
- Pinto, M., & Tall, D. O. (2002). Building formal mathematics on visual imagery: A case study and a theory. *For the Learning of Mathematics*, 22(1), 2–10.
- 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.
- Rasmussen, C., Zandieh, M., King, K., & Teppo, A. (2005). Advancing mathematical activity: A practice-oriented view of advanced mathematical thinking. *Mathematical Thinking and Learning*, 7(1), 51–73.
- Recio, A. M. & Godino, J. D. (2001). Institutional and personal meanings of mathematical proof. *Educational Studies in Mathematics*, 48, 83–99.
- Rowland, T. (2001). Generic proofs in number theory. In S. R. Campbell & R. Zazkis (Eds.), *Learning and teaching number theory: Research in cognition and instruction* (pp. 157–184), Westport, CT: Ablex.
- Segal, J. (2000). Learning about mathematical proof: Conviction and validity. *Journal of Mathematical Behavior*, 18(2), 191–210.
- Selden, A. & Selden, J. (1999). *The role of logic in the validation of mathematical proofs* (Tech. Rep. No. 1999-1). Cookeville, TN: Tennessee Technological University, Department of Mathematics.
- Selden, A., & Selden, J. (2003). Validation of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, 34, 4–36.
- Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, 29, 123–151.
- Sfard, A., & Linchevsky, L. (1994). The gains and pitfalls of reification: The case of algebra. *Educational Studies in Mathematics*, 26, 191–228.
- Simpson, A. P. (1995). Developing a proving attitude. *Proceedings of Justifying and Proving in School Mathematics* (pp. 27–38). London: Institute of Education.
- Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching*, 77, 20–26.
- Toulmin, S. (1969). *The uses of argument*. Cambridge: Cambridge University Press.
- van Dormolen, J., & Zaslavsky, O. (2003). The many facets of definition: The case of periodicity. *Journal of Mathematical Behavior*, 22(1), 91–106.
- Vinner, S. (1991). The role of definitions in teaching and learning mathematics. In D. O. Tall (Ed.), *Advanced mathematical thinking* (pp. 65–81), Dordrecht: Kluwer.
- Watson, A., & Mason, J. (2002). Extending example spaces as a learning/teaching strategy in mathematics. In A. Cockburn & E. Nardi (Eds.), *Proceedings of the 26th annual conference of the International Group for the Psychology of Mathematics Education, Vol. 4* (pp. 378–385). Norwich, UK: University of East Anglia.

- Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, 48, 101–119.
- Weber, K. (2004). Traditional instruction in advanced mathematics courses: A case study of one professor's lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior*, 23(2), 115–133.
- Weber, K. (2005). A procedural route toward understanding aspects of proof: Case studies from real analysis. *Canadian Journal of Science, Mathematics and Technology Education*, 5(4), 469–483.
- Weber, K. (2006). Investigating and teaching the processes used to construct proofs. In F. Hitt, G. Harel, & A. Selden (Eds.), *Research in collegiate mathematics education. VI* (pp. 197–232). Providence, RI: American Mathematical Society.
- Weber, K., & Alcock, L. J. (2004). Syntactic and semantic production of proofs. *Educational Studies in Mathematics*, 56, 209–234.
- Weber, K., & Alcock, L. J. (2005). Using warranted implications to understand and validate proofs. *For the Learning of Mathematics*, 25(1), 34–38.
- Weber, K., Alcock, L. J., & Radu, I. (2005). Undergraduates' use of examples in a transition to proof course. *Proceedings of the 27th annual conference of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Roanoke, VA: Virginia Polytechnic Institute and State University.
- Winicki-Landman, G., & Leikin, R. (2000). On equivalent and non-equivalent definitions. *For the Learning of Mathematics*, 20(1), 17–21 & 20(2), 24–29.
- Yackel, E., Rasmussen, C., & King, K. (2000). Social and sociomathematical norms in an advanced undergraduate mathematics course. *Journal of Mathematical Behavior*, 19(3), 275–287.
- Zepp, R., Monin, J., & Lei, C. L. (1987). Common logical errors in English and Chinese. *Educational Studies in Mathematics*, 18, 1–17.

Appendix A. Sample Transcript, Conceptual Descriptions, and Memos

Throughout, words and phrases likely to become key to the analysis were highlighted in bold in order to facilitate quick re-reading of the large amount of text generated. This is reproduced below as it appeared in the researcher's notes.

A.1. Sample Transcript and Conceptual Descriptions

	<i>Interview transcript</i> (Professor 1, Interview 1, Lines 118-120)	<i>Researcher's conceptual description</i>
P1:	Yes so that's...I mean...the thing that I'm still wrestling with in this course, is there's a lot of stuff that I'm asking them to master really very early on. There's a whole group of skills. Propositional logic, all of these things about, the number of skills before you can even process a new mathematical definition.	Teaching problem that wrestling with is asking them to master a lot of skills – propositional logic e.g. – early on, because needed before can process a definition .
I:	Right.	
P1:	What does a mathematician do, you know? So one of the things, again, that's second nature to me but it's not to them, is that if I see a definition, I immediately instantiate it. You know, just try some examples of this definition, and try to fit it in. And they don't even think, you know – what happens is you know that you describe a new definition, you say, "let f be a function, let x be a real number, we say that..." and then some relationship between f and x holds if, blah blah blah. So then what they have to do, they have to realize that this definition only makes sense in the context of, I have to have a function in mind and then I can make sense out of this definition, but if I don't, then... And what they'll do is typically if you have a sequence, you know, if I have a sequence definition to use in the rest of the problem, and they don't understand the definition, they'll just skip that sentence and go on. I will, they will come in for help on a problem, and five or ten minutes into the discussion I'll realize that, that they never bothered to process this particular definition, they have no idea what this means.	Here we are with instantiation as a thing that mathematicians do . Structure of definition appearing here again as something mathematician can just rattle off. What instantiation does is make sense of definition by having an object in mind . Typically students don't understand the definition , but will skip it and go on. They come for help and he realizes a few minutes in that they didn't bother to process the definition . So they have no idea what it means .

A.2. Examples of Researcher's Summary Memos

P1 I1 114-128

Summary memo: **teaching problem** of asking them to master a lot of **skills** needed to **process definition** and need these **early on**. **Process a definition**

then seems to mean **instantiate examples**, which is a thing that **mathematicians do**. Idea that you **make sense** of a definition by having **objects in mind**, and students don't spontaneously do this. They will **skip a definition** if they don't understand it and try to go on.

P1 I1 114-128

Summary memo: sophisticated **tasks** set for **definition processing**. **Generate examples** then generate more that have **some properties but not others** (Mason ref). **Partitions** as example.

P1 I1 114-128

Summary memo: two more examples of **facility** on part of professor. Both sort of **spontaneously** generating example that **models structure** of what want to describe a **generic definition** and **generic statement that would be hard to negate**. Also **spontaneously restates** more **concisely**.

A.3. Examples of Researcher's Theoretical Memos

P1 I1 114-128

I just don't believe that you need to have these skills in prop logic etc. before you can even begin to process a definition. I believe that you could do it badly and then learn them by means of fixing what you are doing. I believe that he's more on productive lines with example generation as a means of beginning to process what is going on, rather than the meta-math. No arguments that that is useful, of course, but having it as pre-emptive makes me a little uncomfortable because where is the motivation? Unless you're doing that teaching about mathematics thing, but even then I'd say you want something to relate it to. Ah, he says that in the next line. Well, maybe you do need some parsing skills to even do that. I don't know.

P1 I1 114-128

Got a slight implicit laziness here in saying that they didn't bother to process the definition, but this sounds very much like Dahlberg & Housman in that they anticipated being able to do something with the statement without doing that. It's like operating with the statements rather than any objects. Which, of course, on some level is what we want because properties are where it's at. The problem, again, is that having an idea of the objects seems essential to the mathematician, probably for helping you with the decision of what to do. And just it being about something.

P1 I1 114-128

Ah, of course, example generation/instantiation as a way of attacking the abstraction problem head-on. You have the students make it less abstract for themselves. This is sort of obvious really. Of course we consider the task so obvious that we assume they're thinking in that way anyway.

GRADUATE SCHOOL OF EDUCATION, RUTGERS UNIVERSITY, 10 SEMINARY PLACE, NEW BRUNSWICK, NJ 08901, USA

Current address: Department of Mathematical Sciences, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, UK

E-mail address: lalcock@essex.ac.uk