DEPARTMENT OF MATHEMATICS

TECHNICAL REPORT

OVERCOMING STUDENTS' DIFFICULTIES IN LEARNING TO UNDERSTAND AND CONSTRUCT PROOFS

DR. ANNIE SELDEN

DR. JOHN SELDEN

New Mexico State University

AUGUST 2007

No. 2007-1

TENNESSEE TECHNOLOGICAL UNIVERSITY

Cookeville, TN 38505

OVERCOMING STUDENTS' DIFFICULTIES IN LEARNING TO UNDERSTAND AND CONSTRUCT PROOFS[1]

Annie Selden and John Selden
New Mexico State University[2]

When a topologist colleague was asked to teach remedial geometry, he used *Schaum's Outline of Geometry* and also wrote proofs on the blackboard. One day a student, who was familiar with two-column proofs having statements such as $\triangle ABD \cong \triangle BCD$ and reasons such as SAS, blurted out in utter surprise, "You mean proofs can have words!"

While this geometry student's previous experience had led him to an unfortunate view of proof. Other students experience epiphanies about themselves and proof. Asked what she (personally) got out of a transition-to-proof course, one of our students answered, "I learned that I could wake up at 3 a.m. thinking about a math problem."

What do responses like this tell us? Almost all undergraduate mathematics courses are about the concepts and theorems of mathematics -- when a matrix has an inverse, how to find it, and when to use it; when a series converges; the distinction between continuous and uniformly continuous; the meaning of compact. However, students in courses like abstract algebra, real analysis, and topology normally demonstrate their competence by solving problems and proving theorems. And, if students go beyond a few lower-division courses such as calculus or first differential equations, this usually involves constructing original proofs or proof fragments. But, often not much time can be devoted to helping students learn how to construct proofs. This might not lead to difficulties, if only students came to university understanding something about the nature of proof and already had some experience constructing simple proofs. Unfortunately, many students, even high-performing ones, do not. And the resulting difficulties they encounter may be one of the reasons many students do not continue in mathematics.

Transition-to-proof courses, also called bridge courses, are meant to ameliorate this situation. Their main focus is not on the concepts and theorems of mathematics, but on helping students learn to construct proofs.[3] This is perhaps best seen as a complex constellation of content knowledge, beliefs, problem solving ability, and skills. These skills include identifying hypotheses and conclusions, locating relevant definitions and theorems, using them appropriately, isolating the mathematical "problem," coming up with "key" ideas to solve it, and finally, organizing them into a logically coherent deductive argument. Their acquisition seems to be considerably aided by practice, and

---

[1] This manuscript will appear as a chapter in an upcoming MAA Notes Volume: *Making Connections: Research to Practice in Undergraduate Mathematics Education*.

[2] Annie Selden has retired from Tennessee Technological University and is now Professor Emerita. John Selden is retired from and also formerly taught at Tennessee Technological University. The Seldens have retired to Las Cruces, New Mexico and now jointly teach one graduate course per semester at New Mexico State University

[3] However, the introduction to one transition-to-proof textbook asserts something different. It states that the purpose of such courses, often called something like Introduction to Mathematical Reasoning, is to gather together results concerning basic set notions, equivalence relations, and functions so teachers will not have to cover them repeatedly at the beginning of courses like abstract algebra and real analysis.

the process of learning to construct proofs may even involve students coming to know themselves better. Indeed, the above student's comment, about waking up with a math problem, suggests that she has learned to persist until she eventually comes up with a solution, even if that's in the middle of the night. Unfortunately, many students believe that they either know how to solve a problem (prove a theorem) or they don't, and thus, if they don't make progress within a few minutes, they give up and go on to something else.

Of course, undergraduate students do not learn to construct proofs only in transition-to-proof courses. They tend to improve their ability to construct proofs throughout the entire undergraduate mathematics program. Some departments do not even offer transition-to-proof courses, and some combine them with mathematics content courses such as discrete structures. Occasionally, students are offered an R. L. Moore type course,[4] that is, a course in which the textbook and lectures are replaced by a brief set of notes and in which the students produce all the proofs. To some extent, the emphasis in such courses is on a deep understanding of the mathematical content -- however, it has been our experience that once students get started in such courses they often improve their proof making abilities very rapidly. Unfortunately, a few students may have great difficulty getting started.

In whatever setting students are to progress in their proving abilities, one might expect the teaching to be somewhat special. In many university mathematics content courses, teachers can profitably explain mathematical theorems and why they are true, but in teaching the skills and problem solving abilities involved in proving, one should also expect to emphasize guiding students' practice. In developing such teaching, it can be useful to ask: What kinds of difficulties do student have, and how might these difficulties be alleviated?

We will describe some results from the mathematics education research literature that address these questions. However, it is important to note that this research typically makes no claim (as one familiar with other social sciences might expect) that all, or even most, students have the described difficulties. Instead, the main point of this literature is to uncover, understand, and describe features of student learning, in this case difficulties that might otherwise go unnoticed, at least to the degree claimed. Evidence in such work is usually directed towards being sure that the observations and descriptions are accurate about the particular (often, small) group of students studied.

As we proceed, we will mention possible pedagogical reasons for the difficulties described, and when they are known, we will give some pedagogical suggestions. In conclusion, we will offer some additional teaching suggestions, based partly on the research literature and partly on our own teaching experience.

The Curriculum and Students' and Teachers' Conceptions of Proof

That the concept of proof in mathematics is a difficult one for students is not surprising given various everyday uses of the word "proof." To a jury, it can mean "beyond a reasonable doubt" or "the preponderance of the evidence." To a social scientist or statistician, it can mean "occurring with a certain (rather large) probability."

---

[4] Such courses are also referred to as modified Moore Method or Texas Method courses. For more information, see Mahavier (1999) or Jones (1977).

And, to a scientist, it can mean the positive results of an empirical investigation. To comprehend the special way that "proof" is used in mathematics can take time and such everyday meanings can get in the way.

Views of High School Geometry Students

A number of studies have documented the finding that the concept of mathematical proof is not quickly or easily grasped. For example, in the middle of a year-long U.S. high school geometry course, after being introduced to deductive proof, students in five classes were given a short instructional unit designed to highlight differences between measurement of examples and deductive proof. Seventeen of the students were interviewed and asked to compare and contrast two arguments (for different theorems) -- a deductive proof and an argument containing four examples using differently shaped triangles. Some of these students had a nuanced "evidence is proof" view. They considered empirical evidence to be sufficient proof for a statement about all triangles, provided one took measurements of each type of triangle -- acute, obtuse, right, scalene, equilateral, and isosceles. Others had a qualified view of deductive proof, believing that a two-column proof only proved a theorem for the type of triangle depicted in the accompanying figure and would need to be reproved, perhaps using the same steps, for other types of triangles. Most surprising and quite disturbing, especially after an instructional unit designed to help them make the distinction between empirical justification and deductive proof, was the result that some of these geometry students simultaneously believed that empirical evidence is sufficient proof and that proof is just evidence for a claim (Chazan, 1993).

The Influence of the Curriculum

It would seem that partly, perhaps even to a large extent, how students see proof is a consequence of how it is portrayed by their teachers and the overall curriculum. To give a well-documented example, we turn to secondary schools in England where a new National Curriculum, for students aged 5-16 years, was adopted following the 1982 publication of the influential Cockcroft Report. The new curriculum was partly "in response to evidence of children's poor grasp of formal proof in the 60's and 70's" (Hoyles, 1997). The intent of the new curriculum was for students to test and refine their own conjectures in order to gain personal conviction of their truth, and to present justifications for their validity, that is, deductive arguments.

Somehow, in that new National Curriculum, proof was relegated to only one of several strands, called student Attainment Targets (ATs), namely, to AT1: Using and Applying Mathematics.[5] Unfortunately, as was found several years later, even high-attaining secondary students came to see proof in terms of the "investigations" that were undertaken in this applied strand of the curriculum (Coe & Ruthven, 1994; Healy & Hoyles, 1998, 2000). While such investigations were undertaken with the intention of having students see proof as less of a formal ritual demanded by teachers and more as a

---

[5] The new National Curriculum has undergone several changes in the number of attainment targets (ATs). In 1995, the other three ATs were AT2: Number and Algebra; AT3: Shape, Space, and Measures; and AT4: Handling Data. Relegating proof to AT1 (Using and Applying Mathematics) had the effect of separating the function of proof away from the other mathematical strands (Hoyles, 1997, p. 8).

natural outgrowth of testing, refining, and verifying their own conjectures, the results were disappointing, even disastrous, for students entering England's universities. Apparently, the verifications, that had been intended to be student constructed deductive arguments, were instead turned into standardized templates and empirical arguments (Coe & Ruthven, 1994).

By 1995, the situation had caused so much concern that the London Mathematical Society issued a report on the problems as mathematicians perceived them. The report stated that recent changes in school mathematics "have greatly disadvantaged those who need to continue their mathematical training beyond school level." In particular, the following problems were cited: "serious lack of essential technical facility -- the ability to undertake numerical and algebraic calculation with fluency and accuracy," "a marked decline in analytical powers when faced with simple problems requiring more than one step," and "a changed perception of what mathematics is -- in particular of the essential place within it of precision and proof" (London Mathematical Society, 1995, p. 2).

After the public outcry of mathematicians, a large-scale study, called *Justifying and Proving in School Mathematics*, was undertaken. The study surveyed 2,459 high-attaining Year 10 students (14-15 years old, that is, comparable to U.S. high school sophomores) in 94 classes from 90 English and Welsh schools. In a series of papers and reports, it was convincingly documented that it was the new National Curriculum, as implemented by teachers, that was, in large part, responsible for the perceived decline in U.K. students' notions of proof and proving (Hoyles, 1997; Healy & Hoyles, 1998, 2000).

What did this large, mostly quantitative, but partly qualitative, study find? In the Executive Summary of the report (Healy & Hoyles, 1998), one finds the following conclusions, amongst others. (1) Students' performance on constructing proofs was "very disappointing." These better-than-average[6] students were asked to judge whether a number of empirical, narrative, and algebraic arguments were correct and convincing, as well as which they would produce and which would get the best marks from their teachers. They were also asked to prove a familiar result, *the sum of any two odd numbers is even*, and an unfamiliar result, *if p and q are two odd numbers, then* $(p+q)\times(p-q)$ *is a multiple of 4.* Only 40% of students showed evidence of deductive reasoning for the familiar result, and just 10% did so for the unfamiliar result. (2) Even so, most students (84% in geometry, 62% in algebra) were aware that once a statement is proved "no further work was necessary to check if it applied to a particular range of instances," for example, the sum of two odd numbers that are squares is also even. (3) Students who expected to take the higher level GCSE examination at age 16, rather than the middle-level examination, were better at constructing and identifying correct arguments.[7] In conclusion, the report stated:

> The major finding of the project is that most high-attaining Year 10 students after following the National Curriculum for 6 years are unable to distinguish and describe mathematical properties relevant to a proof and use deductive reasoning in their arguments. . . . at least some of the poor performance in proof of our

---

[6] They had scored an average of 6.56 on a national test (the Key Stage 3) whose overall average is normally between 5 and 6.

[7] British students can opt for one of three levels of examination: foundation-, middle-, and higher-tier. It has been argued elsewhere that the acceptability (for entrance to university) of good results on the middle-tier has discouraged many students from taking the higher-tier GCSE mathematics examination.

> highest-attaining students may simply be explained by their lack of familiarity with the process of proving. (Healy & Hoyles, 1998, p. 6)

Thus, the way a curriculum conveys proof and proving is clearly crucial, but skilled and knowledgeable teachers are also critical for implementing such a curriculum. The current *NCTM Standards* (2000) advocate reasoning and proof across the K-12 curriculum. For example, in a section describing the Reasoning and Proof Standard for Grades 9-12, one reads, "Students should understand that having many examples consistent with a conjecture may suggest that the conjecture is true but does not prove it, whereas one counterexample demonstrates that a conjecture is false." (NCTM, 2000, p. 345.) Are current U.S. secondary school teachers capable of providing the rich opportunities and experiences with proof that would enable students to come to such an understanding?

Secondary Teachers' Views and Knowledge of Proof

In one recent study (Knuth, 2000a, 2000b), seventeen secondary school mathematics teachers, with from 3 to 20 years teaching experience, some with master's degrees, were interviewed on their conceptions of proof and its place in secondary school mathematics. They were asked such questions as: What does the notion of proof mean to you? Why teach proof in secondary school? When should students encounter proof? They were also presented with arguments for several mathematical statements and asked to evaluate them. Some of these arguments were proofs; others were not.

Although all these teachers professed the view that a proof establishes the truth of a conclusion, several also thought it might be possible to find a counterexample or some other contradictory evidence to refute a proof. The interviews produced no evidence to suggest the teachers saw proof as promoting understanding or insight. Three teachers did talk about the role of proof in explaining why something is true, but by this they meant understanding how one proceeded step-by-step from the premise to the conclusion.

In the context of secondary school, the teachers distinguished formal proofs, less formal proofs, and informal proofs. For some teachers, two-column geometry proofs were the epitome of formal proofs. Less formal proofs were not as mathematically rigorous, and informal proofs were explanations or empirically-based arguments. All the teachers considered proof as appropriate only for those students in advanced mathematics classes and those intending to pursue mathematics-related majors in college. All indicated that they accepted informal proofs from students in lower-level mathematics classes. However, doing only this may have the unfortunate consequence that students develop the belief that checking several examples constitutes proof (Knuth, 2002a).

The teachers were given five sets of statements with 3 to 5 arguments purporting to justify them; in all, there were 13 arguments that were proofs and 8 that were not. The teachers rated each argument on a four-point scale with 1 not a proof and 4 a proof and provided rationales for their ratings. Ratings of 2 or 3 were included to allow teachers to express alternative views of validity. In general, the teachers were successful in recognizing proofs, with 93% of the proofs rated as such. However, the number of nonproofs they also rated as proofs was surprising -- a third of the nonproofs were rated as proofs. In fact, every teacher rated at least one of the eight nonproofs as a proof and eleven teachers rated more than one as a proof. Indeed, ten teachers considered an argument demonstrating the converse of the statement, *If* $x > 0$, *then* $x + \frac{1}{x} \geq 2$, to be a proof of it; these teachers seemed to focus on the correctness of the algebraic manipulations, rather than on the validity of the argument (Knuth, 2000b).

Given this result regarding some better and more committed secondary mathematics teachers, can one expect that beginning U.S. university students would be reasonably skilled at proof and proving? Would they, for example, understand the distinction between proof and empirical argument? Probably not.

University Students' Views of Proof

Undergraduate students sometimes come to see proofs and proving as unrelated to their own ways of thinking. In order to cope, they may employ mimicking strategies with the result that they develop various views of proof that are unusual from a mathematician's viewpoint; Harel and Sowder (1998) have classified some of these "proof schemes." These are *not* techniques of (mathematical) proof, but rather kinds of arguments, sometimes incorrect or incomplete, that some university students find convincing, and may even think of as proofs.[8] An example of preservice elementary teachers' views of proof follows.

In the 10th week of a sophomore-level mathematics course, 101 preservice elementary teachers were asked to judge verifications of a familiar result, *if the sum of the digits of a whole number is divisible by 3, then the number is divisible by 3,* and an unfamiliar result, *if a divides b and b divides c, then a divides c*. For each of these, students were given, in randomized order, inductive arguments based on examples, patterns, and specific large numbers, and deductive arguments -- a general proof, a false "proof," and a particular (or generic[9]) proof. These students, who had met the idea of proof in their high school geometry courses and in the current course, rated these arguments on a four-point scale, where 4 indicated they considered the argument to be a mathematical proof and 1 indicated it was not a proof. The results showed that both inductive and deductive arguments were acceptable to the students. Apparently the current course that had given "extensive and explicit instruction about the nature of proof and verification in mathematics" had not achieved its goal. In particular, each of the inductive arguments was rated high (3 or 4) by more than 50% of the students. For both familiar and unfamiliar contexts, 80% gave a high rating (3 or 4) to at least one inductive argument, and over 50% gave a very high rating (4) to at least one inductive argument. Also, while over 60% accepted a correct deductive argument as a valid mathematical proof, 52% also accepted an incorrect deductive argument (Martin & Harel, 1989).

Nonstandard views of mathematical proof can be seen as obstacles to overcome. While it is not clear precisely how to bring students' views of proof in line with mathematicians' views, it seems plausible that working towards mathematical sense-making, explanation, and justification on the part of students would be one possible route,

---

[8] The taxonomy of "proof schemes" also includes various axiomatic proof schemes, that is, arguments that mathematicians would consider proofs.

[9] A generic proof is a proof of a particular case that can be generalized in a straightforward way. For example, see Rowland (2002).

provided one avoids the pitfall, described above, of allowing mathematical "investigations" to conclude with purely empirical justifications.

Understanding and Using Definitions and Theorems

Not only are there everyday uses of "proof" that might compound students' difficulties in coming to know what a mathematical proof is, students can be confused about the role of definitions in mathematics.

Mathematical Definitions

Everyday descriptive, or dictionary, definitions[10] describe both concrete and abstract things, already existing in the world, such as trees, love, democracy, or epistemology. They can be both redundant and incomplete, and it is never clear whether all aspects of a definition must apply for its proper use. In contrast, mathematical definitions[11] bring concepts into existence; the concept, say of group, means nothing more and nothing less than whatever the definition says. While all parts of a mathematical definition definitely need to be considered when producing examples and nonexamples, other features of prospective examples need not be considered. This point is often missed. When asked whether $F = 151 \times 157$ is prime, a number of preservice elementary teachers correctly, but irrelevantly noted that both 151 and 157 are prime, before going on to conclude that their product is composite (Zazkis & Liljedahl, 2004). Furthermore, in proving theorems, one should consider all parts of a definition.

Students may not be aware of, or may not make, the distinction between everyday definitions and mathematical definitions. One could help them become aware of this distinction by discussing it with them and by engaging them in the act of defining.

Interpreting and Using Theorems

Undergraduate students often fail to use relevant theorems or interpret the content of theorems incorrectly [see Rasmussen and Ruan in this volume for a notable exception]. Below, we provide some examples that illustrate students' difficulties in using and interpreting theorems.

The Fundamental Theorem of Arithmetic, guaranteeing a unique prime decomposition of integers, is part of the core mathematics curriculum for preservice elementary teachers, but in practice some of these students appear to deny the uniqueness. Zazkis and Campbell (1996) asked preservice elementary teachers whether 17^3 was a square number or whether K = 16,199 = 97 x 167 could have 13 as a divisor, given both 97 and 167 are primes. These students took out their calculators -- in the first instance, to multiply out and extract the square root, and in the second instance, to divide by 13. When asked to determine (and explain) whether $M = 3^2 \times 5^2 \times 7$ was divisible by

---

[10] Dictionary definitions are also referred to as *descriptive, extracted*, or *synthetic* definitions.

[11] Mathematical definitions are also referred to as *stipulated* or *analytic* definitions. Such definitions apply in an "all or nothing" sense, that is, a given set, together with an operation, is a group or is not a group. In contrast, one can say that two countries are democracies, yet that one is more democratic than the other.

2, 3, 5, 7, 9, 11, 15, or 63, a majority (29 of 54) stated that 3, 5, 7 were divisors since those were among the factors in the prime decomposition. However, sixteen were unable to apply similar reasoning to 2 and 11, some noting instead that "M is an odd number" so "2 can't go into it" or resorting to calculations (like the above) for 11. In addition, many of these students believed that prime decomposition means decomposition into *small* primes (see also Zazkis & Liljedahl, 2004).

Undergraduate students often ignore relevant hypotheses or apply the converse when it does not hold. A well-know instance is the use, by Calculus II students, of the converse of: *If* $\sum a_n$ *converges, then* $\lim_{n\to\infty} a_n = 0$, as an easy, but incorrect, test for convergence. Some calculus books go on to point out that this theorem provides a Test for Divergence. But, perhaps it would be better to explicitly state the contrapositive, *If* $\lim_{n\to\infty} a_n \neq 0$*, then* $\sum a_n$ *diverges.*

Sometimes undergraduate students use theorems, especially theorems with names, as vague "slogans" that can be easily retrieved from memory, especially when they are asked to answer questions to which the theorems seemingly apply. For example, Hazzan and Leron (1996) asked twenty-three abstract algebra students: *True or false? Please justify your answer. "In* S_7 *there is no element of order 8."* It was expected that students would check whether there was a permutation in S_7 having 8 as the least common multiple of the lengths of its cycles. Instead, 12 of the 16 students who gave incorrect answers invoked Lagrange's Theorem[12] or its converse. Seven of them incorrectly invoked Lagrange's Theorem to say the statement was false -- there is such an element since 8 divides 5040. Another two students inappropriately invoked a contrapositive form of Lagrange's Theorem to say the statement was true because 8 doesn't divide 7. The authors go on to point out that students often think Lagrange's Theorem is an existence theorem, although its contrapositive shows that it is a non-existence theorem: *If k doesn't divide o*(*G*)*, then there doesn't exist a subgroup of order k.* Perhaps it would be good to state this version explicitly for students.

The above examples refer to students' misuse of theorems when they are asked to solve specific problems, for example, determine whether a number is prime or a series converges, or decide whether a group has an element of order 8. However, it is not hard to imagine similar difficulties when students attempt to use theorems in constructing their own proofs.

A Positive Result on Improving Students' Mathematical Reasoning

That even young children can make remarkable strides towards proof, when challenged with appropriate tasks and probing questions, can be seen from one noteworthy longitudinal study (Maher & Martino, 1996a, 1996b, 1997). The study consisted of a series of relatively small, but coherent, long-term interventions with one group of children over a number of years. It led to some extraordinary instances of mathematical sense-making, explanation, and justification, including the development by children, on their own, of the idea of proof in a concrete case.

---

[12] Lagrange's Theorem states: Let G be a finite group. If H is a subgroup of G, then $o(H)$ divides $o(G)$. Here $o(H)$ stands for the order of the group H, i.e., its cardinality.

We describe the reasoning progress of Stephanie, one of the children with whom Maher and Martino (1996a, 1996b, 1997) began their long-range, but occasional, interventions commencing in Grade 1. By Grade 3, the children had begun building physical models and justifying their solutions to the following problem: How many different towers of heights 3, 4, or 5 can be made using red and yellow blocks? Stephanie not only justified her solutions, she validated or rejected

> her own ideas and the ideas of others on the basis of whether or not they made sense to her. . . . She recorded her tower arrangements first by drawing pictures of towers and placing a single letter on each cube to represent its color, and then by inventing a notation of letters to represent the color cubes. (Maher & Speiser, 1997, p. 174)

She used spontaneous heuristics like guess and check, looking for patterns, and thinking of a simpler problem, and developed arguments to support proposed parts of solutions, and extensions thereof, to build more complete solutions. Occasional interventions continued for Stephanie through Grade 7. Then in Grade 8 she moved to another community and another school and her mathematics was a conventional algebra course. The researchers interviewed her that year about the coefficients of $(a+b)^2$ and $(a+b)^3$. About the latter, she said "So there's a cubed . . . And there's three a squared b and there's three a b squared and there's b cubed. . . . Isn't that the same thing?" Asked what she meant, she replied, "As the towers." It turned out, upon further questioning, that Stephanie had been visualizing red and yellow towers of height 3 in order to organize the products $a^i b^j$. At home, before the interview, she had written out the coefficients for the first six powers of $a+b$. Also, in a subsequent interview, she could explain how counting the towers related to the binomial coefficients and used Pascal's triangle to predict the terms of $(a+b)^n$. (For a more complete discussion, see Maher & Speiser, 1997.) Her understanding prompted Speiser to remark, "I wish some of my [university] students were able to reason that well." The case of Stephanie illustrates possibilities for developing solid mathematical reasoning early on.

However, most undergraduate students do not come to university with such experiences. Rather, as described above, many come with nonstandard views of proof. In addition, they often need to acquire much of the complex constellation of knowledge and skills used in proving theorems.

Understanding the Structure of a Proof and the Order in which it Might be Written

Transition-to-proof course students often say they don't know what the teacher wants them to do or where to begin. This is especially true of intuitively obvious results such as $A \cup B = B \cup A$, where one "follows one's nose" logically and there is no "trick" or mathematical problem-solving aspect. Rather it is a matter of knowing how to use, and sometimes unpack, the relevant definitions, including using them in the "right" order. For mathematicians, this has become automatic, whereas students often don't know what to do. For example, on the final exam in a transition-to-proof course, students were asked to prove: *Let f and g be functions on A. If $f \circ g$ is one-to-one, then g is one-to-one*. In the course, the definition of f one-to-one was that if $f(x) = f(y)$ then $x = y$. However, all but one student unsuccessfully began with the hypothesis -- $f \circ g$ is one-to-one -- rather than assuming that $g(x) = g(y)$. (Moore, 1994). They did not appear to know how to use the definition of one-to-one and relate that to the structure of their proofs.[13]

Unpacking the Logical Structure of Statements of Theorems

Another difficulty students have when constructing their own proofs is an inability to unpack the logical structure of informally stated theorems -- theorems that depart from a natural language version of predicate calculus. That is, theorems that omit specific mention of some variables or depart from the use of *for all, there exists, and, or, not, if-then*, and *if-and-only-if* in a significant way. For example the statement, *Differentiable functions are continuous*, is informal because a universal quantifier and the associated variable are understood by convention, but not explicitly indicated. Similarly, *A function is continuous whenever it is differentiable* is informal because it departs from the familiar *if-then* expression of the conditional as well as not explicitly specifying the universal quantifier and variable.

Being able to unpack the logical structure of such informally stated theorems is important because the logical structure of a mathematical statement is closely linked to the overall structure of its proof. For example, knowing the logical structure of a statement helps one recognize how one might begin and end a direct proof of it. When asked to unpack the logical structure of four informally worded syntactically correct statements, two true and two false, undergraduate mathematics students, many in their third or fourth year, did so correctly just 8.5% of the time. Especially difficult for them was the correct interpretation of the order of the existential and universal quantifiers in the false statement: *For* $a < b$, *there is a* c *so that* $f(c) = y$ *whenever* $f(a) < y$ *and* $y < f(b)$[14] (Selden & Selden, 1995).

Furthermore, the ability to unpack the logical structure of the statement of a theorem also allows one to know whether an argument proves that statement, as opposed to some other statement. For example, eight mid-level undergraduate mathematics and mathematics education majors were asked to judge the correctness of student-generated "proofs" of a single theorem.[15] Upon finding a proof of the converse particularly easy to follow, four initially incorrectly stated that it was a proof of the original statement, and two of these maintained this view throughout the interview (Selden & Selden, 2003).

Understanding the Effect of Existential and Universal Quantifiers

One source of students' difficulties in discerning the logical structure of theorems is a lack of understanding of the meaning of quantifiers and that their order matters.

---

[13] In a rather formally written proof, one might begin something like, "Suppose $f \circ g$ is one-to-one." But (with this definition), the hypothesis is not used until one attempts to prove that g is one-to-one by assuming $g(x) = g(y)$. An alternative definition, $x \neq y$ implies $f(x) \neq f(y)$, might have made this particular theorem easier to prove, but apparently the students did not think of using it.

[14] If f were continuous and if it were stated that $a < c < b$, this would be the Intermediate Value Theorem as stated in most beginning calculus textbooks.

[15] The theorem was: *For any positive integer n, if* n^2 *is a multiple of* 3, *then n is a multiple of* 3.

Undergraduate students often consider the effect of an interchange of existential and universal quantifiers as a mere rewording. For example, in another study, when given the two statements:

- *For every positive number* a *there exists a positive number* b *such that* $b < a$.
- *There exists a positive number* b *such that for every positive number* a $b < a$.

24 of 54 students in undergraduate mathematics courses, such as linear algebra and multivariable calculus, and 3 of 9 students in a beginning graduate abstract algebra course said they were "the same" or were merely reworded (Dubinsky & Yiparaki, 2000).

While understanding the logical structure of a definition or a theorem is certainly not sufficient for constructing a proof, it is definitely necessary. In other words, if you do not understand what something really says, you certainly cannot prove it.

Knowing How to Read and Check Proofs

An integral part of the proving process is being able to tell whether one's argument is correct and proves the theorem it was intended to prove. For this, one must check one's own proof. How do undergraduates read and check proofs?

An Exploratory Study

We conducted an exploratory study of how eight undergraduates (four secondary education mathematics majors and four mathematics majors) from the beginning of a transition-to-proof course validated, that is, evaluated and judged the correctness of, four student-generated "proofs" of a very elementary number theory theorem (Selden & Selden, 2003). The "proofs" were real student work from a similar transition-to-proof course. The theorem was: *For any positive integer* n, *if* n^2 *is a multiple of* 3, *then* n *is a multiple of* 3. Unbeknownst to the students, for later reference, we had dubbed the student-generated "proofs": (a) Errors Galore, (b) The Real Thing, (c) The Gap, and (d) The Converse, indicating our view of them. Each of the eight students was interviewed individually for about one hour in a semistructured interview consisting of four phases. In Phase 1, the students were asked to explain the statement of the theorem in their own words, give some examples of it, and try to prove it. Two were successful, and after some time, those who could not complete a proof were asked to proceed with the other portions of the interview. In Phase 2, the students were shown the four "proofs," one after the other, and asked to think out loud as they read each one and decided whether it was, or was not, a proof. In Phase 3, having seen and thought about the "proofs" one after the other, they were given an opportunity to reread them all together and rethink their earlier decisions. In Phase 4, they were asked some general questions about how they read proofs. For example: When you read a proof is there anything different you do, say, than in reading a newspaper? How do you tell when a proof is correct or incorrect? How do you know a proof proves this theorem instead of some other theorem?

The students made judgments regarding the correctness of each student-generated "proof" four times, at the beginning and end of each of their two readings in Phases 2 and 3. At each time, there were 32 person-proof judgments. At the beginning (Time 1), these were just 46% (15 of 32) correct, but by the end (Time 4) 81% were correct. We attribute this difference to the students (at Time 4) having thought about the "proofs" several times, and perhaps, to the interviewer's no longer accepting "unsure" as a response. Most of the errors detected were of a local/detailed nature rather than a global/structural nature, with only the two students who had proved the theorem themselves observing that the converse had been proved in (d).

When asked how they read proofs, the students said they attempted careful line-by-line checks to see whether each mathematical assertion followed from previous statements, checked to make sure the steps were logical, and looked to see whether any computations were left out. Several said they went through the proofs using an example. Also, for these students, a feeling of personal understanding or not--that is, of making sense or not--seemed to be an important criterion when making a judgment about correctness of a "proof." Thus, what students say about how they read proofs seems a poor indicator of whether they can actually validate proofs with reasonable reliability. While these students tended to "talk a good line," their judgments at Time 1 were no better than chance (46% correct).

On the other hand, even without explicit instruction, the reflection and reconsideration engendered by the interview process eventually yielded 81% correct judgments, suggesting that explicit instruction in validation could be effective (Selden & Selden, 2003). Indeed, several transition-to-proof textbooks include "proofs to grade,"[16] but we think it would also be helpful to have students validate actual student-generated proofs.

Knowing and Using Relevant Concepts

In addition to the ability to unpack, understand, and interpret definitions and theorems correctly, and check one's logic, one must have some relevant content knowledge. Constructing all but the most straightforward of proofs involves a good deal of persistence and problem solving to put together relevant concepts. And in order to use a concept flexibly, it is important to have a rich *concept image*, that is, a lot of examples, non-examples, facts, properties, relationships, diagrams, and visualizations, that one associates with that concept.[17]

In many upper-level mathematics courses, students are given definitions together with a few examples, after which they are expected to use these definitions reasonably flexibly. To do this, students may need to find additional examples and non-examples and to prove or disprove related conjectures, more or less without guidance. One can think of these activities as helping to build students' concept images. How does one go about building a rich concept image for a newly introduced concept? Do undergraduate

[16] Textbooks for such courses have from none to just a few to a moderate number of exercises involving critiquing "proofs." The directions vary -- students may be asked to: (a) find the fallacy in a "proof;" (b) tell whether a "proof" is correct; (c) grade a "proof," A for correct, C for partially correct, or F; or (d) evaluate both a "proof" and a "counterexample." Most of these "proofs" have been carefully constructed by the textbook authors so there is just one error to detect. See, for example, Smith, Eggen, and St. Andre (1990; p. 39).

[17] The idea of concept image was introduced by Vinner and Hershkowitz (1980), elaborated by Tall and Vinner (1981), and is now a much used notion in mathematics education research.

students actively try to enhance their concept images, for instance, by considering examples and nonexamples?

Getting to Know and Use a New Definition

In one study conducted by Dahlberg & Housman (1997), eleven students, all of whom had successfully completed introductory real analysis, abstract algebra, linear algebra, set theory, and foundations of analysis, were presented with the following formal definition. A function is called *fine* if it has a root (zero) at each integer. They were first asked to study the definition for five to ten minutes, saying or writing as much as possible of what they were thinking, after which they were asked to generate examples and nonexamples. Subsequently, they were given functions and asked to determine whether these were fine functions and, if so, why. Next, they were asked to determine the truth of four conjectures, such as "No polynomial is a fine function."

Four basic learning strategies were used by the students on being presented with this new definition – example generation, reformulation, decomposition and synthesis, and memorization. Examples generated included the constant zero function and a sinusoidal graph with integer x-intercepts. Reformulations included $f(-1) = 0$, $f(0) = 0$, $f(1) = 0$, $f(2) = 0$, . . . , and $f(n) = 0 \ \forall n \in Z$. Decomposition and synthesis included underlining parts of the definition and asking about the meaning of "root." Two students simply read the definition – they could not provide examples without interviewer help and were the ones who most often misinterpreted the definition.

Of these four strategies, example generation, together with reflection, elicited the most powerful "learning events," that is, instances where the authors thought students made real progress in understanding the newly introduced concept. Students who initially employed example generation as their learning strategy came up with a variety of discontinuous, periodic continuous, and non-periodic continuous examples and were able to use these in their explanations. Those who employed memorization or decomposition and synthesis as their learning strategies often misinterpreted the definition, for example, interpreting the phrase "root at each integer" to mean a fine function must vanish at each integer in its domain, but that its domain need not include all integers. Students who employed reformulation as their learning strategy developed algorithms to decide whether functions they were given were fine, but had difficulty providing counterexamples to false conjectures (Dahlberg & Housman, 1997).

Thus, it seems that while students are often reluctant, or unable, to generate examples and counterexamples, doing so helps enrich their concept images immensely and enables them to judge the probable truth of conjectures.

Dealing with Various Symbolic Representations

Another aspect of understanding and using a concept is knowing which symbolic representations are likely to be appropriate in certain situations; this can be very important for success in proving. Concepts can have several (easily manipulated) symbolic representations or none at all. For example, prime numbers have no such representation; they are sometimes defined as those positive integers having exactly two factors or being divisible only by 1 and themselves. It has been argued that the lack of an (easily manipulated) symbolic representation makes understanding prime numbers especially difficult, in particular, for preservice teachers (Zazkis & Liljedahl, 2004). Similarly, irrational numbers have no such representation; thus, in proving results such as $\sqrt{2}$ is irrational or the sum of a rational and an irrational is irrational, one is led to consider proofs by contradiction -- something often difficult for beginning students.

Symbolic representations can make certain features transparent and others opaque.[18] For example, if one wants to prove a multiplicative property of complex numbers, it is often better to use the representation $re^{i\theta}$, rather than $x + iy$, and if one wants to prove certain results in linear algebra, it may be better to use linear transformations, *T*, rather than matrices. Students often lack the experience to know when a given representation is likely to be useful.

It has been argued that moving flexibly between representations (e.g., of functions given symbolically or as a graph) is an indication of the richness of a student's understanding of a concept (Even, 1998). Also, understanding an abstract mathematical concept can be regarded as possessing "a notationally rich web of representations and applications" (Kaput, 1991, p. 61).

Bringing Appropriate Knowledge to Mind

No one questions the need for content knowledge, sometimes referred to as resources,[19] in order to solve problems and prove theorems. But students can have such resources and not be able to bring them to bear on a problem, or proof, at the right time.

Knowing, but not Using, Factual Knowledge

In two companion studies, 19 volunteer third quarter A and B calculus students, and later 28 volunteer differential equations students, took a one-hour paper-and-pencil test (without calculators) asking them to solve five moderately non-routine first calculus problems, that is, problems somewhat, but not very, different from what they had been taught. Immediately afterwards, they took a half-hour routine test, covering the resources needed to solve the non-routine problems. For example, one non-routine problem was: *Find at least one solution to the equation* $4x^3 - x^4 = 30$ *or explain why no such solution exists*. Two-thirds of the calculus students failed to solve a single problem completely and more than 40% did not make substantial progress on a single problem. Also, more than half of the differential equations students were unable to solve even one problem and more than a third made no substantial progress toward a solution. Of those non-routine problems for which the students had full factual knowledge, just 18% of the calculus students' solutions and 24% of the differential equations students' solutions were completely correct (Selden, Selden, & Mason, 1994; Selden, Selden, Hauk, & Mason, 2000).

---

[18] Representations can be *transparent* or *opaque* with respect to certain features. For example, representing 784 as 28^2 makes the property of being a perfect square transparent, but representing 784 as $(13 \times 60) + 4$ makes that property opaque. For more details, see Zazkis and Liljedahl (2004, pp. 165-166).

[19] Schoenfeld (1985) described good mathematical problem-solving performance in terms of resources, heuristics, control, and belief systems.

To solve the above non-routine problem, one needs to know (1) that one might set the derivative of $4x^3 - x^4 - 30$ equal to zero to find its maximum -3 and (2) that solutions of the given equation are where this function crosses the x-axis (which it does not). Many of the students had these two resources, but apparently could not bring them to mind at an appropriate time. We conjectured that, in studying and doing homework, the students had mainly followed worked examples from their textbooks and had thus never needed to consider various different ways to attempt problems. Thus, they had no experience at bringing their assorted resources to mind. It seems very likely that a similar phenomenon could occur in attempting to prove theorems.

How does one think of bringing the appropriate knowledge to bear at the right time? To date, mathematics education research has had only a little to say about the difficult question of how an idea, formula, definition, or theorem comes to mind when it would be particularly helpful, and probably there are several ways. In their study of problem solving, Carlson and Bloom (2005) found that mathematicians frequently did not access the most useful information at the right time, suggesting how difficult it is to draw from even a vast reservoir of facts, concepts, and heuristics when attempting to solve a problem or to prove a theorem. Instead, the authors found that mathematicians' progress was dependent on their approach, that is, on such things as their ability to persist in making and testing various conjectures.

Our own personal experience of eventually bringing to mind resources that we had -- but did not at first think of using -- suggests that persistence, over a time considerably longer than that of the Carlson and Bloom interviews, can be beneficial. We conjecture that certain ideas get in the way of others, and that after a good deal of consideration, such unhelpful ideas become less prominent and no longer block more helpful ideas. This may be related to a psychological phenomenon that can take several forms; for example, in vision, if one fixates on a single spot in a picture, it will eventually disappear.

While coming to mind at the right time can be seen as an idiosyncratic, individual act, it may sometimes be related to the idea of transfer of one's knowledge. How does one come to see a new mathematical situation as similar to a previously encountered situation and bring the earlier resources to bear on the new situation?

Knowing What's Important and Useful

In addition to knowing what a proof is, being able to reason logically, unpack definitions, and apply theorems, and having a rich concept image of relevant ideas, one needs a "feel" for the content and what kinds of properties and theorems are important. Knowing what's important should go a long way towards bringing to mind appropriate resources.

Not Seeing that Geometry Theorems are Useful when Making Constructions

Seeing the relevance and usefulness of one's knowledge and bringing it to bear on a problem, or a proof, is not easy. Schoenfeld (1985, pp. 36-42) provides an example of two beginning college students who had completed a year of high school geometry and were asked to make a construction: *You are given two intersecting straight lines and a* *point P marked on one of them. Show how to construct, using straightedge and compass, a circle that is tangent to both lines and that has the point P as its point of tangency to one of the lines.* During a 15-minute joint attempt, they made rough sketches and conjectures, and tested their conjectures by making constructions. When asked why their constructions ought or ought not to work, they responded in terms of the mechanics of construction, but did not provide any mathematical justification. Yet the next day they were able to give the proof of two relevant geometric theorems within five minutes. Apparently, these students simply did not see the relevance of these theorems at the time.

Knowing to Use Properties, Rather than the Definitions, to Check Whether Groups are Isomorphic

In another study, four undergraduates who had completed a first abstract algebra course and four doctoral students working on algebraic topics were observed as they proved two group theory theorems and attempted to prove or disprove whether specific pairs of groups are isomorphic: $\mathbf{Z}_n$ and $\mathbf{S}_n$, $\mathbf{Q}$ and $\mathbf{Z}$, $\mathbf{Z}_p \times \mathbf{Z}_q$ and $\mathbf{Z}_{pq}$ (where p and q are coprime), $\mathbf{Z}_p \times \mathbf{Z}_q$ and $\mathbf{Z}_{pq}$ (where p and q are not coprime), $\mathbf{S}_4$ and $\mathbf{D}_{12}$. Nine times these undergraduates, who were successful in only two of twenty instances, first looked to see if the groups had the same cardinality; after which they attempted unsuccessfully to construct an isomorphism between the groups. They rarely considered properties preserved under isomorphism, despite knowing them (as ascertained by a subsequent paper-and-pencil test). For example, they all knew **Z** is cyclic, **Q** is not, and a cyclic group could not be isomorphic to a non-cyclic group, but they did not use these facts and none were able to show **Z** is not isomorphic to **Q**, until afterwards. These facts did not seem to come to mind spontaneously, or in reaction to this kind of question.

In contrast, the doctoral students, who were successful in comparing all of the pairs of groups, rarely considered the definition of isomorphic groups. Instead, they examined properties preserved under isomorphism. When the groups were not isomorphic, they showed one group possessed a property that the other did not; for example, **Z** is cyclic, but **Q** is not. To prove $\mathbf{Z}_p \times \mathbf{Z}_q$ is isomorphic to $\mathbf{Z}_{pq}$, where p and q are coprime, three of them noted that the two groups have the same cardinality and showed $\mathbf{Z}_p \times \mathbf{Z}_q$ is cyclic. None tried to construct an isomorphism (Weber & Alcock, 2004).

Knowing which Theorems are Important

In comparing the proving behaviors of four undergraduates who had just completed abstract algebra and four doctoral students who were writing dissertations on algebraic topics, it was found that the doctoral students had knowledge of which theorems were important when considering homomorphisms. For example, in considering the proposition: *Let G and H be groups. G has order pq (where p and q are prime). f is a surjective homomorphism from G to H. Show that G is isomorphic to H or H is abelian,* all four doctoral students recalled the First Isomorphism Theorem within 90 seconds. In contrast, two undergraduates did not invoke the theorem, while the other two invoked its weaker form only after considerable struggle. When the doctoral students were asked why they used such sophisticated techniques, a typical response was, "Because this is such a fundamental and crucial fact that it's one of the first things you turn to” (Weber, 2001).

Another four undergraduates, who had recently completed their second course in abstract algebra, and four mathematics professors, who regularly used group-theoretic concepts in their research, were interviewed about isomorphism and proof (Weber & Alcock, 2004). They were asked for the ways they think about and represent groups, for the formal definition and intuitive descriptions of isomorphism, and about how to prove or disprove two groups are isomorphic. The algebraists thought about groups in terms of group multiplication tables and also in terms of generators and relations, as well as having representations that applied only to specific groups, such as matrix groups. Each algebraist gave two intuitive descriptions of groups being isomorphic: that they are essentially the same and that one group is simply a re-labeling of the other group. To prove or disprove two groups are isomorphic, they said they would do such things as "size up the groups" and "get a feel for the groups," but could not be more specific. In addition, they said that they would consider properties preserved by isomorphism and facts such as $\mathbf{Z}_n$ is *the* cyclic group of order n.

In contrast, none of the undergraduates could provide a single intuitive description of a group; for them, it was a structure that satisfies a list of axioms. While all four undergraduates could give the formal definition of isomorphic groups, none could provide an intuitive description. To prove or disprove that two groups were isomorphic, these undergraduates said they would first compare the order (i.e., the cardinality) of the two groups. If the groups were of the same order, they would look for bijective maps between them and check whether these maps were isomorphisms (Weber & Alcock, 2004).

It may be that undergraduates mainly study completed proofs and focus on their details, rather than noticing the importance of certain results and how they fit together. That is, they may not come to see some theorems as particularly important or useful. The mathematics education research literature contains few specific teaching suggestions on how to help students come to know which theorems are likely to be important in various situations. But, it might be helpful to discuss with them: (1) which theorems and properties you (the teacher) think are important and why, (2) your own intuitive, or informal ideas, regarding concepts, and (3) the advantages and disadvantages of various representations.

Teaching Proof and Proving

Some Suggestions Emanating from Research

One very positive finding, which was described earlier, is the remarkable sophistication of reasoning reached by some average school students who received brief interventions over a number of years (Maher & Martino, 1996a, 1996b, 1997). As described above, these students used a variety of spontaneously developed heuristics. Eventually, in order to come to agreement, these students, more or less, invented the idea of proof in a concrete case. If grade school students can be encouraged in this way, why not university students? Perhaps this could be done in part with relatively short "interventions" spread across the entire undergraduate program.

Another result is that younger students seem to prefer explanatory proofs written with a minimum of notation. This was certainly the case for U.K. Year 10 students (Healy & Hoyles, 1998). For example, instead of using mathematical induction to prove that sum of the first n integers is $n(n+1)/2$, one could use a variant of Gauss's original argument. Namely, for any n, one can write the sum in two ways as $(1 + 2 + 3 + \cdots + n)$ and as $(n + (n-1) + (n-2) + \cdots + 1)$, then add corresponding terms to obtain n identical summands equal to $n+1$, so twice the original sum equals $n(n+1)$. Hence, the original sum must equal $n(n+1)/2$ (Hanna, 1989, 1990). It seems plausible that undergraduates, and people more generally, might prefer proofs that provide insight to proofs that just establish the validity of a result.[20]

It also appears that great care should be taken to distinguish empirical reasoning from mathematical proof. Exactly how this can be done effectively is not especially clear, since merely giving high school geometry students a short instructional unit on this distinction left some of them very unclear as to the difference between empirical evidence and proof (Chazan, 1993). Perhaps secondary and university teachers need to stress this distinction often and also get students to discuss and reflect on situations where simple pattern generalization does not work.

Since current secondary teachers' conceptions of proof are somewhat limited and they sometimes accept non-proofs as proofs (Knuth, 2002a, 2002b), one way to enhance preservice secondary teachers' abilities to check the correctness of proofs might be to have them consider and discuss, in groups, a variety of student-generated "proofs," as well as having them provide feedback on each other's proofs.

In addition to explaining the difference between descriptive definitions in a dictionary and mathematical definitions, one can engage students in the defining process. For example, when using Henderson's (2001) investigational geometry text, one can begin with a definition of triangle initially useful in the Euclidean plane, on the sphere, and on the hyperbolic plane, but eventually students will notice that the usual Side-Angle-Side Theorem (SAS) is not true for all triangles on the sphere. At this point, they can be brought to see the need for, and participate in developing, a definition of "small triangle" for which SAS remains true on the sphere.

Perhaps it would also be possible to create classroom activities to improve students' ability to enhance their concept images and deal with representations flexibly. One suggestion is that upon introducing a new definition, one could ask students to generate their own examples, alternatively, to decide whether professor-provided instances are examples or non-examples, "without authoritative confirmation by an outside source" (Dahlberg & Housman, 1997, p. 298). Another possibility might be to engage students in conjecturing which kinds of symbolic representations might be useful for solving a given problem or proving a specific result. Also, one could point out that when a theorem has a negative conclusion (e.g., $\sqrt{2}$ is irrational), a proof by contradiction may be just about the only way to proceed.

For certain theorems in number theory, it has been suggested that the transition to formal proof can be aided by going through a (suitable) proof using a generic example

[20] It has been suggested that proofs have various functions within mathematics: explanation, communication, discovery of new results, justification of a new definition, developing intuition, and providing autonomy (e.g., Hanna, 1989; de Villiers, 1990; Weber, 2002).

that is neither too trivial nor too complicated (Rowland, 2002). Gauss's proof that the sum of the first *n* integers is *n*(*n*+1)/2, done for *n* = 100 is one such generic proof. Done with care, going over generic proofs interactively with students could enable them to "see" for themselves the general arguments embedded in the particular instances. If the theorem involves a property about primes, 13 and 19 are often suitable, provided the proof is constructive and that prime (e.g., 13) can be "tracked" through the stages of the argument. A generic proof, but not the standard one, can be given for Wilson's Theorem: *For all primes p,* $(p-1)! \equiv p-1 (\mathrm{mod}\ p)$. That argument for *p* = 13 involves pairing each integer from 2 to 11 with its (distinct) multiplicative inverse mod 13, noting the product of each pair is congruent to 1(mod 13), and concluding that $12! \equiv 1 \times 1 \times 12 (\mathrm{mod}\ 13)$.[21] There is one caveat; there is some danger that students will not understand the generic character of the proof. In an attempt to avoid this, one can subsequently have them write out the general proof.

Some Personal Observations and Ongoing Work

We see learning to construct proofs, especially for beginning students, as composed largely of the acquisition of a complex constellation of skills, content knowledge, beliefs, and problem solving ability -- much of which is best learned by doing. As a result, we think university teachers should consider including a good deal of student-student and teacher-student interaction regarding students' proof attempts, as opposed to just presenting their own or textbook's proofs. Trying to teach such a complex constellation entirely by lecture seems like trying to teach someone to tie her/his shoelaces entirely over the telephone. It might be possible, but seems unlikely to be the most effective way.

In that connection, it might be useful, and certainly could do no harm, to discuss with students some of the difficulties mentioned above: in particular, the difference between mathematical proof and other types of arguments and the difference between mathematical definitions and everyday definitions. It might be helpful to stress that mathematicians of today see proofs as consisting of such careful deductive reasoning that, barring mistakes, the results (theorems) require no further evidence, are permanently true, and can be immediately used anywhere that the premises hold.

Another suggestion is that, when presenting proofs, one could take a top-down approach to explanation, first giving a global overview of the proof's structure to avoid the appearance of "pulling a rabbit out of a hat," followed by introducing and developing concepts as needed (Leron 1983, 1985).

Students often do not appreciate that proofs themselves can have a hierarchical structure -- that there are subproofs (and subconstructions) within proofs, perhaps several levels deep. One could make students aware of this and illustrate how structure comes into thinking about how to prove a theorem. Students need to understand that proofs are not generally conceived of in the order they are written. Not realizing this may result in quite a few students not making use of the hierarchical structure of proofs in their own proving attempts and lead to some of the difficulties mentioned earlier. Students need to

---

[21] For details of this and some other number-theoretic generic proofs, along with a description of how they were used with Cambridge University undergraduates, see Rowland (2002).

be encouraged to write parts of a tentative proof "out of order" (e.g., What will the last line say?), even when they sometimes resist doing so.

There seems to be quite a lot to learn about the way in which proofs are customarily written. If students were taught about this way of writing in some of their courses, they might not be so puzzled about how to begin a proof. Indeed, we take the point of view that proofs are deductive arguments *in an identifiable genre*. They differ from arguments in legal, political, and philosophical works. Within this genre, individual styles can vary, just as novels by Hemingway and Faulkner have differing styles, although their novels are easily seen as belonging to a single genre that clearly differs from newspaper articles, short stories, or poems. As part of some ongoing work, we have been collecting general features of the genre of proof. For example, definitions already stated outside of proofs tend not to be written into them. In teaching, we have found that pointing out such features, especially in the context of a student's own work, can be helpful to students.

Furthermore, we have found it useful to have students carefully examine the structure of the statement that they are trying to prove, and even to think about how a tentative proof might be structured, before launching into it. For example, consider proving the theorem (mentioned earlier): *Let f and g be functions on A. If $f \circ g$ is one-to-one, then g is one-to-one.* It would be useful for a prover to first unpack the meaning of g being one-to-one. Doing so can direct one to begin the proof by writing, "Let x and y be in the domain of g and suppose $g(x) = g(y)$." This also makes clear that the desired conclusion is "Thus $x = y$." In this way, one exposes the "real, but hidden" mathematical task, namely, to get from $g(x) = g(y)$ to $x = y$. After that, students can concentrate on how the hypothesis that $f \circ g$ is one-to-one might help.

Concluding Remarks

We have tried to provide readers with a coherent organization of some of the mathematics education research on proof and proving, but there is much more.[22] Awareness of the variety of difficulties undergraduates have with proof and proving can make one more sensitive regarding how to help them. The above pedagogical suggestions indicate some steps one might take; however, more information on "what works" is needed.

References

Carlson, M. P., & Bloom, I. (2005). The cyclic nature of problem solving: An emergent multidimensional problem solving framework. *Educational Studies in Mathematics, 58,* 45-76.

[22] Anyone wanting to delve into the considerable literature on proof and proving can go to the bibliography maintained by the *International Newsletter on the Teaching and Learning of Proof* at: http://www.lettredelapreuve.it/. Those with a more philosophical bent might want to consult the annotated bibliography at: http://fcis.oise.utoronto.ca/~ghanna/mainedu.html.

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics*, *24*, 359-387.

Cockcroft, W. M. (Chair). (1982). *Mathematics counts: Report of the Committee of Inquiry into the Teaching of Mathematics in Schools* (Department of Education and Science). London: HSMO.

Coe, R., & Ruthven, K. (1994). Proof practices and constructs of advanced mathematics students. *British Educational Research Journal*, *20*(1), 41-53.

Dahlberg, R. P., & Housman, D. L. (1997). Facilitating learning events through example generation. *Educational Studies in Mathematics*, *33*, 283-299.

de Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras*, 24, 7-24

Dubinsky, E., & Yiparaki, O. (2000). On student understanding of AE and EA quantification. In E. Dubinsky, A. H. Schoenfeld, & J. Kaput (Eds.), *Issues in mathematics education: Vol. 8. Research in collegiate mathematics education. IV* (pp. 239-289). Providence, RI: American Mathematical Society.

Edwards, B. S., & Ward, M. B. (2004). Surprises from mathematics education research: Student (mis)use of mathematical definitions. *American Mathematical Monthly*, *111*, 411-424.

Even, R. (1998). Factors involved in linking representations. *Journal of Mathematical Behavior*, *17*, 105-121.

Hanna, G. (1989). Proofs that prove and proofs that explain. In G. Vergnaud, J. Rogalski, & M. Artigue (Eds.), *Proceedings of the Thirteenth Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 45-51). Paris: CNRS - Paris V.

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

Harel, G., & Sowder, L. (1998). Students’ proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Issues in mathematics education: Vol. 7. Research in collegiate mathematics education. III* (pp. 234-283). Providence, RI: American Mathematical Society.

Hazzan, O., & Leron, U. (1996). Students' use and misuse of mathematical theorems: The case of Lagrange's Theorem. *For the Learning of Mathematics*, *16*(1), 23-26.

Healy, L., & Hoyles, C. (1998). *Technical report on the nationwide survey: Justifying and proving in school mathematic*s. Institute of Education, University of London.

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

Henderson, D. W. (2001). *Experiencing geometry in Euclidean, spherical and hyperbolic spaces* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Hoyles, C. (1997). The curricular shaping of students' approaches to proof. *For the Learning of Mathematics*, *17*(1), 7-16.

Jones, F. B. (1977). The Moore method. *American Mathematical Monthly*, *84* (4), 273-278.

Kaput, J. (1991). Notations and representations as mediators of constructive processes. In E. von Glasersfeld (Ed.), *Radical constructivism in mathematics education* (pp. 53-74). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Knuth, E. J. (2002a). Teachers' conceptions of proof in the context of secondary school mathematics. *Journal of Mathematics Teacher Education, 5*(1), 61-88.

Knuth, E. J. (2002b). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education*, *33*, 379-405.

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

Leron, U. (1985). Heuristic presentations: The role of structuring. *For the Learning of Mathematics*, *5*(3), 7-13.

London Mathematical Society, Institute of Mathematics and its Applications, Royal Statistical Society. (1995). *Tackling the mathematics problem*. London, U.K.: London Mathematical Society.

Mahavier, W. S. (1999). What is the Moore method? *PRIMUS*, *9*, 339-354.

Maher, C. A., & Martino, A. M. (1996a). The development of the idea of proof. A five-year case study. *Journal for Research in Mathematics Education*, *27*, 194-219.

Maher, C. A., & Martino, A. M. (1996b). Young children inventing methods of proof: The gang of four. In L. Steffe, P. Nesher, P. Cobb, G. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 431-447). Hillsdale, NJ: Lawrence Erlbaum Publishers.

Maher, C. A. & Martino, A. M. (1997). Conditions for conceptual change: From pattern recognition to theory posting. In H. Mansfield & N. H. Pateman (Eds.), *Young* *children and mathematics: Concepts and their representation* (pp. 58-81). Sydney, Australia: Australian Association of Mathematics Teachers.

Maher, C. A., & Speiser, R. (1997). How far can you go with a tower of blocks? In E. Pehkonen (Ed.), *Proceedings of the 21st Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 174-183). Jyväskylä, Finland: Gummerus.

Martin, G. W., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education*, *29*, 41-51.

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

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Rowland, T. (2002). Generic proofs in number theory. In S. R. Campbell & R. Zazkis (Eds.), *Learning and teaching number theory: Research in cognition and instruction* (pp. 157-183). Westport, CT: Ablex Publishing.

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

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

Selden, A., Selden, J., Hauk, S., & Mason, A. (2000). Why can't calculus students access their knowledge to solve non-routine problems? In E. Dubinsky, A. H. Schoenfeld., & J. Kaput, (Eds.), *Issues in mathematics education: Vol. 8. Research in collegiate mathematics education. IV* (pp. 128-153). Providence, RI: American Mathematical Society.

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

Selden, J., Selden, A., & Mason, A. (1994). Even good calculus students can't solve nonroutine problems. In J. Kaput and E. Dubinsky (Eds.), *Research issues in undergraduate mathematics learning: Preliminary analyses and results* (MAA Notes No. 33, pp. 19-26). Washington, DC: Mathematical Association of America.

Smith, D., Eggen, M., & St. Andre, R. (1990). *A transition to advanced mathematics* (3rd ed.). Pacific Grove, CA: Brooks/Cole Publishing Company.

Tall, D., & Vinner, S. (1981). Concept image and concept definition with particular reference to limits and continuity. *Educational Studies in Mathematics, 12*, 151-169.

Vinner, S., & Hershkowitz, R. (1980). Concept images and common cognitive paths in the development of some simple geometrical concepts. In *Proceedings of the Fourth International Conference for the Psychology of Mathematics Education* (pp. 177-184). Berkeley, CA.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, *48*, 101-119.

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., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics*, *56*, 209-234.

Zazkis, R., & Campbell, S. (1996). Prime decomposition: Understanding uniqueness. *Journal of Mathematical Behavior, 15*, 207-218.

Zazkis, R., & Liljedahl, P. (2004) Understanding primes: The role of representation. *Journal for Research in Mathematics Education, 35*, 164-186.