DEPARTMENT OF MATHEMATICS
TECHNICAL REPORT

TEACHING PROVING BY COORDINATING ASPECTS OF PROOFS WITH STUDENTS' ABILITIES[1]

DR. ANNIE SELDEN
DR. JOHN SELDEN
New Mexico State University

OCTOBER 2007

No. 2007-2

TENNESSEE TECHNOLOGICAL UNIVERSITY
Cookeville, TN 38505

Teaching Proving by Coordinating Aspects of Proofs with Students' Abilities[1]

John Selden and Annie Selden
Department of Mathematical Sciences
New Mexico State University

In this chapter, we introduce some concepts for analyzing proofs, including various structures, and for analyzing undergraduate and beginning graduate mathematics students' proving abilities. We then discuss how the coordination of these two analyses might be used to improve students' ability to construct proofs.

For this purpose, we need a richer framework for keeping track of students' progress than the everyday one. We need to know more than that a particular student can, or cannot, prove theorems by induction or contradiction or can, or cannot, prove certain theorems in beginning set theory or analysis. It will be more useful to describe a student's work in terms of a finer-grained framework including various smaller abilities that contribute to proving and that may be learned in differing ways and at differing periods of a student's development.

Developing a fine-grained framework for analyzing students' abilities is not an especially novel idea. In working with higher primary and secondary students, Gutiérrez and Jaime (1998) developed a fine-grained framework of reasoning processes in order to more accurately and easily assess student van Hiele levels.

For proof construction, there are already a number of abilities suitable for keeping track of students'progress. For example, in comparing undergraduates who had

[1]

completed a course in abstract algebra with doctoral students in abstract algebra, Weber (2001) found the doctoral students more able to use strategic knowledge. When asked if two specific groups, such as □ and Pts, are isomorphic, the undergraduates first looked to see if the groups had the same cardinality; after which they attempted unsuccessfully to construct an isomorphism between them, whereas the doctoral students considered properties preserved by isomorphism (Weber & Alcock, 2004), strategically a better starting point.

The ability to use strategic information is about the proving processes of individual students. But it is also useful to consider the product of such processes -- the proofs themselves -- because differing kinds of, and aspects of, proofs can correspond to various abilities needed to produce them. We begin by analyzing various structures of proofs.

In Section 1, we discuss three different structures of proofs, illustrated through a single real analysis proof that, under the right circumstances, advanced undergraduate or beginning graduate students might construct. In Section 2, we present a similar analysis of a simpler proof -- one that mid-level undergraduates might construct. After that, in Section 3, we make three informal observations about proofs, proving, and the teaching of proofs. Then, in Section 4, we initiate a discussion of proving abilities, and in Section 5, we turn to how all of this might be used in teaching.

1. Three structures of proofs

Structures of proofs have been discussed before. Leron (1983) suggests presenting a proof (for example, in a lecture) in a "top down" way, first describing the main ideas before filling in the details. The structure followed in such a presentation is

reminiscent of what we will call the hierarchical structure of a proof, except that Leron is not concerned with characterizing the proof itself, but rather with providing a kind of advance organizer for its presentation.

Konior (1993) is also concerned with a hierarchical structure of proofs, but he focuses on segmentation, that is, "separation of the whole [mathematical] text into a few parts ... in order to reflect the logical structure of the proof and facilitate the reconstruction of the whole proof." He discusses delimiters, such as words and paragraph breaks that indicate beginnings and endings of proof segments (e.g., subproofs). We see this as an interesting approach, one important to students' understandings of how proofs are written, but an investigation of delimiters is beyond the scope of this chapter.

We discuss: (1) a *hierarchical* structure in which we attend to subproofs and subconstructions (such as finding the δ in an $\varepsilon - \delta$ real analysis proof, provided this requires more than one or two steps); (2) a *construction path*, that is, a linear path describing one ordering of the steps[2] through which a proof could have been constructed by an idealized prover, who never erred or followed false leads and who wrote each step as soon as possible; and (3) a division of proofs into what we call the *formal-rhetorical* part and the *problem-centered* part. For any given proof, these three structures can be combined and reduced to a single three-dimensional diagram that qualitatively reveals the proof's complexity. That complexity, in turn, can provide one indication of the proof's difficulty, and knowledge of a proof's difficulty can be an elusive, but important, tool in teaching.

[2]We speak of steps, or sentences, rather than statements, because proofs can contain sentences that are neither true nor false, but instead give instructions to the reader, such as "Let x be a number."

Of course, a particular theorem can often be proved in several ways. However, there are not likely to be very many such ways (except for relatively minor details) for the theorems most advanced undergraduate and beginning graduate students are asked to prove. There are also several ways one might break up longer, compound sentences into smaller ones and this will occasionally be called for, especially in determining a construction path through a proof. Indeed, there may be points in a construction path where several essentially equivalent choices of direction are possible. None of this variability should remove the usefulness of the qualitative judgments of difficulty arising from the analysis.

Next these three structures and the resulting three-dimensional diagram will be presented and illustrated by analyzing a proof[3] of a single theorem: *f+g is continuous at a point, provided f and g are.*

1.1 The hierarchy and construction path of an idealized prover

First a proof of the above theorem is presented. Then that proof is analyzed and built up hierarchically, attending to subproofs and subconstructions. Simultaneously, we build a construction path giving the order in which the sentences (or subsentences) could have been written by an idealized prover. For reference, and to facilitate the analysis, individual sentences of the proof have been numbered in bold brackets (e.g., **[1]**, **[2]**) in their order in the proof. (The sentence labeled **[12, 13, 14, 15]** is further broken down into four subsentences, one for each statement of equality or inequality.)

[3] Since we will be discussing student proving and proofs, we have chosen a statement of the theorem and one proof that some might not regard as the "best."

Proof. **[1]** Let a be a number and f and g be functions continuous at a. **[2]** Let ε be a number > 0. **[3]** Note that $\varepsilon/2 > 0$. **[4]** Now because f is continuous at a, there is a $\delta_1 > 0$ such that for any x_1, if $|x_1 - a| < \delta_1$, then $|f(x_1) - f(a)| < \varepsilon/2$. **[5]** Also because g is continuous at a, there is a $\delta_2 > 0$ such that for any x_2, if $|x_2 - a| < \delta_2$, then $|g(x_2) - g(a)| < \varepsilon/2$. **[6]** Let $\delta = \min(\delta_1, \delta_2)$. **[7]** Note that $\delta > 0$.
[8] Let x be a number. **[9]** Suppose that $|x - a| < \delta$. **[10]** Then $|x - a| < \delta_1$, so $|f(x) - f(a)| < \varepsilon/2$. **[11]** Also $|x - a| < \delta_2$, so $|g(x) - g(a)| < \varepsilon/2$. **[12, 13, 14, 15]** Now $|f(x) + g(x) - (f(a) + g(a))| = |(f(x) - f(a)) + (g(x) - g(a))| \leq$
$|(f(x) - f(a))| + |(g(x) - g(a))| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. **[16]** Thus
$|f(x) + g(x) - (f(a) + g(a))| < \varepsilon$. **[17]** Therefore $f + g$ is continuous at a. QED.

The first step of the construction path, which will appear as **[H1]** for "hidden," is not in the proof or its hierarchical structure. It is obtained by "clarifying" the statement of the theorem, that is, by rewriting the statement in a more formal way, explicitly mentioning variables and quantifiers and using standard logical connectives. This clarification yields the statement: *For all real-valued functions f, all real-valued functions g, and all real numbers a, if f is continuous at a and g is continuous at a, then $f + g$ is continuous at a.*

This version of the statement of the theorem exposes its logical structure. It is essential to understand this logical structure in order to be sure the proof proves this theorem, as opposed to some other theorem. The logical structure, independent of the meaning of "function," "+," and "continuous," yields the first and last sentences of the proof. Together they form what we have called a proof framework[4] (Selden & Selden,

[4] Proof frameworks can be independent of the meanings of certain content words. Selden and Selden (1995, p. 130) illustrated this with a proof framework for a theorem about semigroups that can also serve as a proof framework for a theorem about real intervals. This was accomplished by replacing the words "semigroup" with "set of numbers," "group" with "interval," and "a subgroup of" with "a subinterval of" and retaining the usual meaning of the other words.

1995). The resulting (partial) construction path of our idealized prover is now **[H1]**, **[1]**, **[17]**, **…** and the resulting (partial) hierarchical structure is shown in Figure 1.

Proof: **[1]** Let a be a number and f and g be functions continuous at a.

. . .

[17] Therefore $f + g$ is continuous at a. QED.

Figure 1. The top level of the hierarchical structure

Next we clarify line **[17]** by applying a definition of continuous to $f + g$ at the number a. This clarification yields a new statement, **[H2]**, that needs a subproof: For every number $\varepsilon > 0$, there is a $\delta > 0$, so that for every number x, if $|x - a| < \delta$ then $|f(x) + g(x) - (f(a) + g(a))| < \varepsilon$. Writing a proof framework for this subproof yields the (partial) construction path **[H1]**, **[1]**, **[17]**, **[H2]**, **[2]**, **[7]**, **[8]**, **[9]**, **[16]**, **…** and the resulting (partial) hierarchical structure is shown in Figure 2.

Proof: **[1]** Let a be a number and f and g be functions continuous at a.

[2] Let ε be a number > 0. Let $\delta =$ **…**
[7] Note that $\delta > 0$. **[8]** Let x be a number. **[9]** Suppose that $|x - a| < \delta$.

…

[16] Thus $|f(x) + g(x) - (f(a) + g(a))| < \varepsilon$.

[17] Therefore $f + g$ is continuous at a. QED.

Figure 2. Adding the second level of the hierarchical structure

Next we add in the construction of the δ and finish the proof, yielding the hierarchical structure in Figure 3.

Proof: **[1]** Let a be a number and f and g be functions continuous at a.

[2] Let ε be a number > 0.

[3] Note that $\varepsilon/2 > 0$. **[4]** Now because f is continuous at a, there is a $\delta_1 > 0$ such that for any x_1, if $|x_1 - a| < \delta_1$, then $|f(x_1) - f(a)| < \varepsilon/2$. **[5]** Also because g is continuous at a, there is a $\delta_2 > 0$ such that for any x_2, if $|x_2 - a| < \delta_2$, then $|g(x_2) - g(a)| < \varepsilon/2$. **[6]** Let $\delta = \min(\delta_1, \delta_2)$.

[7] Note that $\delta > 0$. **[8]** Let x be a number. **[9]** Suppose that $|x - a| < \delta$. **[10]** Then $|x - a| < \delta_1$, so $|f(x) - f(a)| < \varepsilon/2$. **[11]** Also $|x - a| < \delta_2$, so $|g(x) - g(a)| < \varepsilon/2$. **[12, 13, 14, 15]** Now $|f(x) + g(x) - (f(a) + g(a))| = |(f(x) - f(a)) + (g(x) - g(a))| \leq |(f(x) - f(a))| + |(g(x) - g(a))| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. **[16]** Thus $|f(x) + g(x) - (f(a) + g(a))| < \varepsilon$.

[17] Therefore $f + g$ is continuous at a. QED.

Figure 3. Adding the third level of the hierarchical structure and finishing the proof

In Figure 3, we first added lines **[3] – [6]** of the construction path and then added lines **[10] – [15]** of the construction path. The first addition was treated as a new level in the hierarchical structure, but the second was not. This is because lines **[3] – [6]** can stand alone as the construction of δ, but lines **[10] – [15]** cannot stand alone as a subproof. Rather, they form a part of **[2] – [16]**, the proof of **[H2]**.

The construction path of our idealized prover is now: **[H1]**, **[1]**, **[17]**, **[H2]**, **[2]**, **[7]**, **[8]**, **[9]**, **[16]**, **[3]**, **[4]**, **[5]**, **[6]**, **[10]**, **[11]**, **[12]**, **[13]**, **[14]**, **[15]**.

Although the linear and hierarchical structures illustrated above can contribute to observing student abilities, one might also ask: Do undergraduate or graduate students just starting to construct proofs even tacitly understand these structures? And does this matter? Perhaps a hint at an answer can be obtained from considering two incidents. First, when the first author was a young mathematician, a colleague gave an amusing

account of Professor M, one of his respected graduate teachers. He said M was known among the graduate students for being so brilliant that when he proved a theorem he sat down at his desk and wrote the statement of the theorem and the first line of its proof. Then, its second line. He continued in this fashion until he reached the last line of the proof, wrote QED, and never looked at the proof again. But why was this tale amusing? Perhaps it was because the audience of mathematicians knew from personal experience that mathematicians do not construct original proofs errorlessly or linearly from the top down. Of course, unlike our idealized prover, most mathematicians probably do not actually write later lines of a proof, or subproof, well before earlier ones. They may not even consciously focus on a specific goal or conclusion. However, we suggest that mathematicians do have, and need, easy conscious access to information about where they are trying to go in a proof or subproof. William James (1890) might well have viewed information available in this way as part of the *fringe* of consciousness[5] which he saw as very important, perhaps because it is instrumental in determining one's future focus.

The second, more recent, incident concerns a chance discussion of the above theorem, its proof, and the above construction path and hierarchical structure with a graduate student in mathematics education. The student was surprised that constructing proofs might involve the kind of structures illustrated above, rather than progressing linearly from top down as one might write a letter. It turned out that this graduate student had studied only a little advanced mathematics and much or all of that had been presented

[5] One can think of consciousness at any particular moment as consisting of *focus* and *fringe*, rather like vision consists of central and peripheral vision. This sensory phenomenon i.e., the separation of sensory processing into small high resolution and large low resolution parts, is fairly common in animals; in addition to vision in humans, a variety of bat has it for hearing and a variety of mole has it for touch.

in the traditional definition-theorem-proof style. The proofs had also been presented in the traditional way from top to bottom, just as one might read them from a textbook.

These two incidents suggest that at least a tacit understanding of the above proof structures matters and that, early in their exposure to advanced mathematics, some students do not grasp much about them. Such students are likely to have a view of proof construction that will prevent them from succeeding. We wonder how many students start avoiding further mathematics courses and eventually abandon hopes of a mathematical career, as a result of such an unfortunate view.

1.2 The formal-rhetorical and problem-centered parts of a proof

Notice that writing steps **[H1]**, **[1]**, **[17]**, **[H2]**, **[2]**, **[7]**, **[8]**, **[9]**, **[16]** in the construction path of our idealized prover (Fig. 2) calls on an understanding of the logical structures of the theorem to be proved and of the definition of continuous at a point (as applied to $f + g$). This is made clear in the rather formal statements **[H1]** and **[H2]**. In addition, one needs to know, and to act on, how parts of a statement relate to parts of its proof; we call knowledge of this kind *behavioral knowledge*[6] because a tendency to behave or act is a part of it. It is closely related to "knowing-to-act in the moment" (Mason & Spence, 1999), as well as to the ideas of "concepts-in-action" and "theorems-in-action" (Vergnaud, 1982). For example, if a formal version of a theorem started, "For all real numbers x," then in a direct proof of it one might start by "introducing" the variable x, with a statement like "Let x be a real number." Although for some authors of proofs, this statement may be left understood instead of made explicit, especially when x

[6] Behavioral knowledge should not be confused with behaviorism, the idea that physically observable, and perhaps measurable, behavior should be the only basis for a scientific treatment of psychology. Taking a behaviorist view would prevent most of the study of the mind, and hence, much current research in mathematics education.

appears in the statement of the theorem. Furthermore, such an x is usually said to be "fixed, but arbitrary" and does not "vary."

Quite a lot of such behavioral knowledge is required in constructing proofs. While it is not so important that a student be able to articulate such behavioral knowledge, it is important that he/she tend to act on it by constructing the next portion of a proof, for example, through writing, speech, or inner speech. Although behavioral knowledge neither implies, nor is implied by, the corresponding procedural (knowing how, without necessarily acting) or conceptual knowledge (knowing why), these may also be useful.

For a student with a reasonable repertoire of proof-related behavioral knowledge, writing the steps in Figure 2 can be quite straightforward. Such writing can be viewed as invoking a schema, much like an expert solves a high school algebra equation. It does not depend on a deep understanding of, or intuition about, the concepts involved or on genuine problem solving in the sense of Schoenfeld (1985, p. 74). We call this part of the proof the *formal-rhetorical* part and contrast it with the remainder of the construction path of our idealized prover, **[3]**, **[4]**, **[5]**, **[6]**, **[10]**, **[11]**, **[12]**, **[13]**, **[14]**, **[15]** (Figure 3), that we call the *problem-centered* part. This problem-centered part of a proof does involve problem solving, although it might be useful to emphasize that a problem can be viewed as a matter of degree, and speak of degrees of non-routineness (Selden, Selden, Hauk, & Mason, 2000, pp. 148-149). The steps in the problem-centered part may call on conceptual knowledge, mathematical intuition, and the ability to bring to mind the "right" resources at the "right time."

Constructing the formal-rhetorical parts of proofs seems to call on a different kind of knowledge than constructing the problem-centered parts of proofs. Helping students acquire these kinds of knowledge probably involves different kinds, or aspects, of teaching, even though these might be blended into a single course and taught by a single teacher. Such a blending of teaching seems especially appropriate for learning to construct proofs because the two parts of a proof interact. Indeed, there are theorems for which constructing the formal-rhetorical part of a proof can be very helpful in revealing the "real problem" to be solved in the rest of the proof. Our sample theorem (Section 1.1) is such a theorem. Since it is about continuous functions, one might expect some kind of visual-spatial intuition about functions would be useful for constructing a proof. However, the formal-rhetorical part (Fig. 2) reveals that the problem is to find a δ that will yield the inequality $|f(x)+g(x)-(f(a)+g(a))| < \varepsilon$ in step **[16]**. The solution does not arise in any obvious way from intuition about functions. Instead, it involves use of $\varepsilon/2$ in the definition of continuity for each of f and g, a choice of δ as the smaller of the resulting two δ's, application of the triangle inequality, and some algebraic rewriting.

In the construction path of our sample proof (Section 1.1), as in many short proofs, the entire formal-rhetorical part of the proof comes first, followed by the problem-centered part. However, in longer proofs, each subproof may have its own formal-rhetorical part and subsequent problem centered part, and the order of subproofs within a construction path can vary.

1.3 A three-dimensional diagram of complexity

We will use the above three structures to build a three-dimensional diagram, or model, of our sample proof (Section 1.1) that will reveal its complexity in a qualitative way (see Figure 4). We see complexity as being a contributor to difficulty, and judgments of difficulty as being useful in teaching. We expect such diagrams to be useful in making rough estimates of the difficulty of constructing proofs.

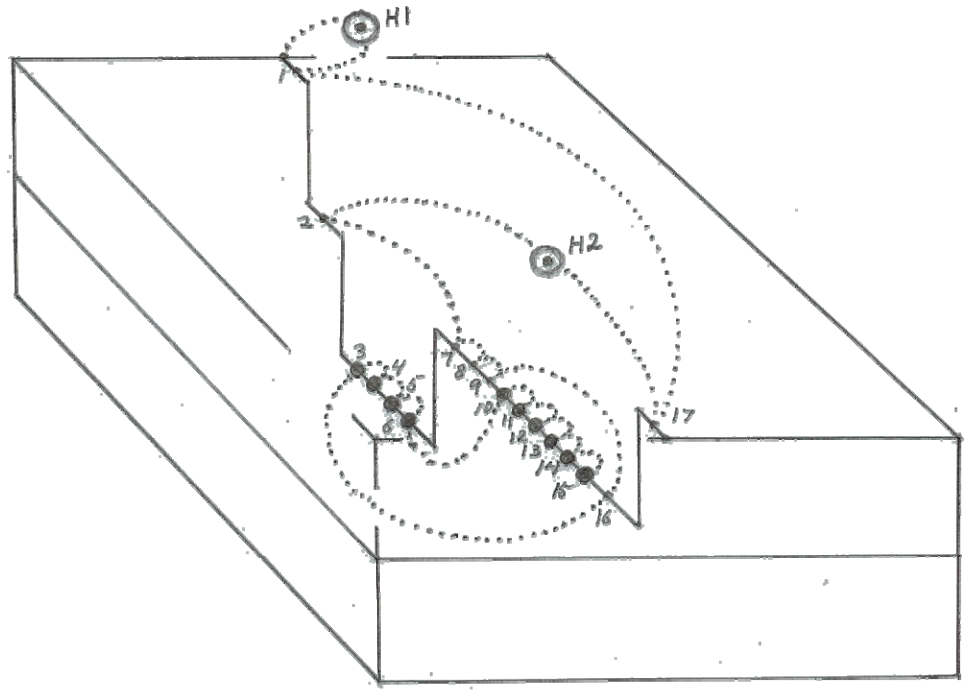

Figure 4. The diagram for the first sample proof

In such diagrams, a proof will be represented by a (possibly branching) rectilinear line through a rectangular solid of several levels, like a layer cake. The line (and proof) starts at the top back and ends at the top front of the diagram. The line drops down a level to represent a subproof or subconstruction. Nodes on the line represent steps in the order they occur in the proof (i.e., **[1]**, **[2]**, **[3]**, …), and proofs involving cases can be

represented by having the line divide into several "parallel" branches where case arguments occur. The construction path will be represented by a curved dotted line from node to node that contains occasional "hidden" nodes off the rectilinear path. The nodes of the problem-centered part of the proof will be indicated by heavier type. This is illustrated in Figure 4 for the sample theorem analyzed in Sections 1.1 and 1.2.

2. A second simpler sample proof

In this section we will analyze the structure of a second simpler sample proof, that of the theorem: *If f and g are functions from A to A and $f \circ g$ is one-to-one, the g is one-to-one*. This theorem is more suitable for an undergraduate transition-to-proof course, while the previous theorem might be better for more advanced students.

The proof with its steps numbered and divided into its hierarchical levels is given in Figure 5.

Proof: **[1]** Let f and g be functions from A to A. **[2]** Suppose $f \circ g$ is one-to-one.

[3] Let x and y be in A. **[4]** Suppose $g(x) = g(y)$.
[5] Then $f(g(x)) = f(g(y))$, **[6]** i.e., $(f \circ g)(x) = (f \circ g)(y)$.
[7] But $f \circ g$ is one-to-one. **[8]** So $x = y$.

[9] Therefore g is one-to-one. QED.

Figure 5. The hierarchical structure of the second sample proof

We will include a clarification, **[H1],** of the statement of the theorem, although it consists only of adding "For all f and g" at the beginning. If-then theorems involving variables, but not quantifiers, invariably are meant to be universally quantified. Also, in order to go beyond the first level proof framework to construct the subproof, one needs a clear definition of g is one-to-one. Here is such a definition, which we call **[H2]**: For all x and y in A, if $g(x) = g(y)$ then $x = y$. The construction path then becomes: **[H1]**, **[1]**, **[2]**, **[9]**, **[H2]**, **[3]**, **[4]**, **[8]**, **[5]**, **[6]**, **[7]**. (One also needs to know the meaning of

$f \circ g$, which allows transforming **[5]** into **[6]**.) The problem-centered part of the proof is then just **[5]**, **[6]**, **[7]** and Figure 6 gives the three-dimensional diagram of the proof.

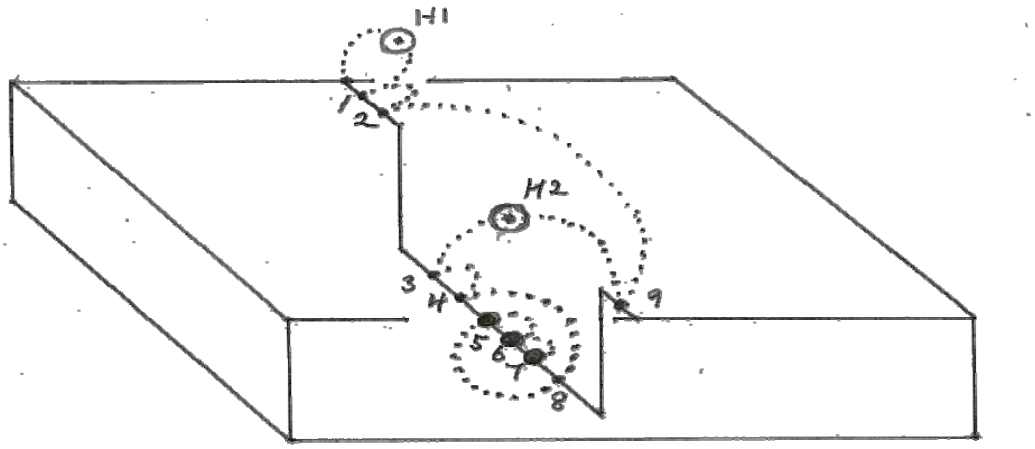

Figure 6. The diagram for the second sample proof.

We now turn to the way proofs are currently often taught and bring up three informal observations.

3. Informal observations

3.1 The genre of proof

When students first start constructing their own proofs, they may inquire: What is a proof? They are occasionally told just to write a convincing argument. However, proofs are not just convincing, deductive arguments, they are also texts composed in a special genre. If beginning students are unaware of the need to write in this genre when asked to write "easy" proofs, they may suffer from what might be called the "obviousness obstacle."

For example, Moore (1994, pp. 258-259) reports that when one student in a transition-to-proof course was asked on a test to prove: *If A and B are sets satisfying*

$A \cap B = A$, *then* $A \cup B = B$, she drew a Venn diagram with one circle, labeled A, contained in a larger circle, labeled B, and gave an intuitive argument based on her "understanding of set equality, subset, intersection, and union" using informal language, rather than a "proof based only on definitions, axioms, previously proved results, and rules of inference." According to the professor, she had not learned "the language and culture of how we write these things down." When asked what was wrong with her proof, the student said "I didn't explain it well enough." Without understanding that there is a genre of proof, such obvious theorems may be very difficult for students.

However, students who see themselves as learning to write in a special genre will have something positive to do, and hence, be less frustrated. This seems to be so, even if no detailed description of the genre of proof has been given, and the students must learn by trial and error or as apprentices. Indeed, it is unlikely that a very useful description of the genre could currently be given because none seems to have been developed yet.

We have been collecting commonalities that can distinguish proof from other forms of deductive argument, for example, arguments that occur within lawyers' speeches to juries, political treatises, or advertising. We will mention just one here, namely, that mathematicians, in research papers and reference books, tend not to restate previously articulated definitions within proofs. For example, the definition of compact is unlikely to be found in a proof that compact Hausdorff implies normal. This is not to say that a definition is never stated within a proof, especially if it is of interest principally in that proof.

So far, the characteristics of the genre of proof appear to be part of the rather sparse style in which proofs are often written.

3.2 Convince yourself

Students are sometimes encouraged to prove theorems by first convincing themselves intuitively, then making the argument more and more precise, eventually arriving at a proof. This is sometimes helpful, but there are theorems that for most students cannot be proved in this way, and those who insist on basing their work on intuition cannot prove such theorems. Our first sample theorem (Section 1.1) is such a theorem. However, for contrast, we will discuss the proof of another theorem that can be obtained by a refinement of intuition.

The theorem is: *For all real-valued functions of a real variable, if* f *is continuous at* a*, and* g *is continuous at* $f(a)$*, then* $g \circ f$ *is continuous at* a. At first glance, this theorem seems to be rather like the theorem about the sum of continuous functions whose proof was discussed in Section 1.1. However, one might develop some visual-spatial intuition about it by examining a sketch (Figure 7).

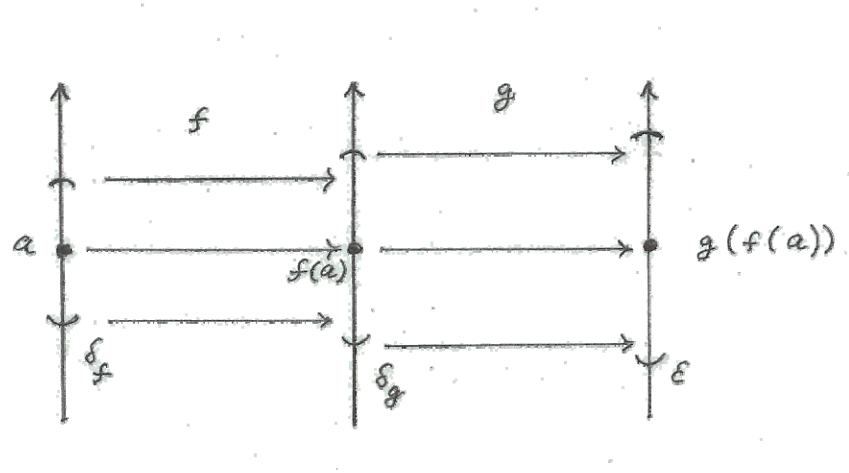

Figure 7. A representational sketch of $g \circ f$

Drawing a representation like that in Figure 7 might well be within the grasp of an advanced undergraduate student who started with the intuitive view that continuous at a

means "points close to a, map close to $f(a)$." This representation would suggest that the theorem is true, and extending the idea of closeness to an $\varepsilon - \delta$ argument would lead to a proof.

3.3 *Logic*

In the United States, most undergraduates, who receive any significant explicit instruction in how to write proofs, are provided it in one mid-level transition-to-proof course. Typically, such courses include logic and a little about such topics as sets, equivalence relations, and functions, as well as some especially accessible mathematics to provide theorems to prove. The logic is likely to consist of an abstract, symbolic, decontextualized treatment of propositional and some predicate calculus, including truth tables and the validity of arguments. This is usually taught early in the course, largely outside of the context of proofs, presumably because logic is considered prerequisite to understanding proofs.

Although some parts of logic, such as understanding how to negate a statement, are certainly important for constructing some proofs, we doubt that formal logic often plays a large role in the proofs beginning students typically construct. This somewhat counterintuitive view is supported by an analysis of student-generated proofs from a transition-to-proof course at a large Southwestern university. The course was similar to that described above, except that most of the lectures occurred in its first half and the second half was devoted to having groups of 3 or 4 students prove theorems in class. Most of the 30 students were mathematics or mathematics education majors and the teacher provided occasional advice to the groups as they attempted to construct proofs. Although the course was not intended to be the basis of a study, all of the 62 correct

student-generated proofs were preserved and later analyzed for a master's thesis (Baker, 2001). These proofs consisted of 926 lines of text whose analysis included noting uses of logic mentioned in the teacher's lecture notes or in the textbook (Velleman, 1994). Logic was used in just 29 lines. The remaining lines invoked definitions and subject matter knowledge or consisted of advance organizers, assertions of hypotheses or conclusions, and statements of whether the proof was direct, by contradiction, or by cases.

We will also examine our first sample proof (Section 1.1) for uses of the kind of logic often taught in transition-to-proof courses. In writing this proof, we doubt that a student would call on much formal logic. However, we can find three places where a valid logical argument, together with information from outside the proof, could have been used. These occur in deducing step **[3]** from **[2]**; step **[7]** from parts of **[4]**, **[5]**, and **[6]**; and the premise of **[10]** from **[9]** (and similarly, the premise of **[11]** from **[9]**).

We will illustrate how deducing step **[7]** from parts of **[4]**, **[5]**, and **[6]** can be viewed as the result of the valid logical argument: $(\mathrm{P} \wedge \mathrm{Q}) \rightarrow \mathrm{R}$, P, Q, $\therefore$ R First take $(\mathrm{P} \wedge \mathrm{Q}) \rightarrow \mathrm{R}$ to be an instance of a statement from outside the proof, namely, *For all real numbers* a *and* b, *if* $a > 0$ *and* $b > 0$, *then* $\min(a,b) > 0$. If one then interprets P as $\delta_1 > 0$ (from **[4]**) and Q as $\delta_2 > 0$ (from **[5]**), this statement becomes: *If* $\delta_1 > 0$ *and* $\delta_2 > 0$, *then* $\min(\delta_1, \delta_2) > 0$, where R is interpreted as $\min(\delta_1, \delta_2) > 0$. Finally, invoking the valid logical argument $(\mathrm{P} \wedge \mathrm{Q}) \rightarrow \mathrm{R}$, P, Q, $\therefore$ R and identifying δ with $\min(\delta_1, \delta_2)$ (from **[6]**), one gets $\delta > 0$, which is **[7]**.

We believe that few, if any, students would consciously employ such cumbersome arguments, that needlessly overburden working memory. Thus, from a practical perspective, our sample proof (Section 1.1) does not call on the kind of logic

normally taught. Because there are many such direct proofs that call only a little on logic, it seems that logic is not prerequisite to understanding proofs and it might be possible and helpful to teach logic in the context of proofs, explicitly discussing it mainly as it comes up in proofs that students are constructing.

Instead of **[7]** being the result of a valid logical argument, we see it as the result of an inference generating schema that accepts both conscious and unconscious inputs, such as parts of **[4]**, **[5]**, and **[6]** and additional information, but that is invoked outside of consciousness. Such a schema would not be under conscious control so might best be developed through experiences and reflections on experiences.[7] This supports our previous suggestion that logic meant to support proofs might best be taught in the context of its use in constructing proofs.

4. Coordinating Aspects of Proofs with Students' Abilities

Returning to the idea of developing a richer framework for keeping track of students' progress, we will discuss some proving abilities that a student might have or that a proof might call for. It is the coordination of these two – what a student might be able to do, or not do, and what a proof might call for – that we suggest may facilitate teaching proving, mainly through students' construction of proofs, rather than in some more teacher-centered way such as lecturing. Much of what we will mention may well be familiar to those who have helped numerous students construct their own proofs. Also, our suggestions are not meant to be in any sense complete.

4.1 Kinds of proofs

[7] The development (i.e., construction) of knowledge through experience and reflections on experience is consistent with a constructivist viewpoint.

Authors of transition-to-proof textbooks typically distinguish direct proofs, proofs by contradiction, proofs by mathematical induction, proofs by cases, and perhaps, existence and uniqueness proofs because students need to be aware that a particular theorem might be proved more easily, or more appropriately, using a specific one of these. However, for guiding students' progress, such distinctions are not fine enough. For example, a particular student might be able to construct a number of direct proofs, but not the kind found in Section 1.1. This might be due to its complex structure and the number of quantifiers involved, and knowing this would allow a teacher to guide that student's current work or to provide later theorems of similar complexity.

For another example, we turn to proofs by contradiction. A student might be able to construct several such proofs, but not a proof of: *There is at most one identity element in a semigroup.* This theorem is amenable to a very short proof by contradiction. However, perhaps the student needs to know the meaning of "at most one," in particular, that its negation is "there are (at least) two."

4.2 Formal-rhetorical versus problem-centered reasoning

In Section 1.2, we distinguished between the formal-rhetorical and problem-centered parts of a proof and suggested that writing these two parts calls on different kinds of knowledge, and hence, on different kinds of teaching. We will call the abilities to write these two parts, *formal-rhetorical reasoning* and *problem-centered reasoning*, and apply these ideas to the kinds of exploration preceding and surrounding proofs, as well as to writing proofs.

Here is an example of the kind of problem-centered reasoning that can precede a proof. Two students with fairly strong upper-division undergraduate mathematics

backgrounds were jointly attempting to prove: *If the number of elements in a set is* n, *then the number of its subsets is* 2^n**,** after failing to do so individually. They had been advised: *Don't forget to count the empty set,* $\varnothing$, *and the whole set as subsets. For example the subsets of* $\{a,b\}$ *are* $\varnothing$, $\{a\}$, $\{b\}$, *and* $\{a,b\}$. Together, they first considered $\{a,b,c\}$ and wrote $\{c\}$, $\{a,c\}$,$\{b,c\}$, and $\{a,b,c\}$. Next they wrote$\{a,b,c,d\}$. When asked about this, they indicated that they had "added" c to each of the subsets of $\{a,b\}$ to generate the four additional subsets of $\{a,b,c\}$. They were about to "add" d to each of the subsets of $\{a,b,c\}$, but were puzzled about how to construct a proof. Based on their earlier inability to make progress individually, the teacher offered a stronger than normal intervention, "You can get from 2 to 3, and you can get from 3 to 4. What does that make you think of?" They replied, "Induction," something they had not previously thought of . Despite appropriate exploration, these students could benefit from an opportunity to think of using induction on the proof of a subsequent theorem -- a point we will return to in Section 5.

4.3 Comparing the difficulty of proofs

For an individual, a major determiner of the difficulty of constructing proofs seems to be the nature of his/her own knowledge and habits of mind, and this will be discussed in Section 4.6. However, in the context of a course in which students prove all, or most, of the theorems, one can often see that one proof (or theorem) is more difficult than another by observing that only a few students can construct it, while many students can construct other proofs. This suggests that there are characteristics intrinsic to proofs of theorems that make some more difficult than others, at least with respect to rough judgments of difficulty. Certainly, we have found a need to make such judgments

especially in teaching Moore Method courses,[8] because we prefer to ask a student to prove a theorem that will be challenging, but not so challenging that he/she fails to do so.

Returning to our suggestion that (rough) judgments of the relative difficulties of proofs can sometimes be made independently of specific individuals, we compare our second sample proof with our first. As indicated by their diagrams, the complexity of the second proof (Fig. 6) seems less than that of the first (Fig. 4). Also, while the problem-centered part of the second proof consists of just three consecutive steps, the problem-centered part of the first proof consists of two separate sections, **[3]**, **[4]**, **[5]**, **[6]** and **[10]**, **[11]**, **[12]**, **[13]**, **[14]**, **[15]**, each of which depends on the other. Finally, one might easily believe the first sample proof could be developed from some kind of visual-spatial intuition about continuous functions, but this is unlikely. Instead, the formal rhetorical part of the proof plays a large role, and a kind of technical-algebraic intuition, based on familiarity with previous definitions and theorems, is used. Thus, it would seem clear that the second proof is less difficult than the first.

4.4 Sets and functions

The language of sets and functions occurs widely in proofs, and undergraduate students are often introduced to it in a way that is somewhat abstract. In keeping track of students' progress, it can be helpful to consider whether they can use a concept in constructing proofs. More important than being able to articulate definitions, students need to use them in proofs, that is, be able to carry out appropriate actions effortlessly in order to leave maximum cognitive resources for other parts of proofs, especially the problem-centered parts. In doing this, students need what we are calling behavioral

[8]In such courses, students are typically given notes containing definitions and statements of theorems, or conjectures, and asked to prove them or to provide counterexamples. The teacher provides the structuring

knowledge, and we suspect this is learned as much from practice at constructing proofs as from abstract definitions.

For example, the definition of set equality is usually given as $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$, and students are told that this means A and B have the same elements. But in constructing a proof that two sets are equal, it should come to mind easily that this involves showing that an arbitrary element of each set is in the other. That is, normally two subproofs need to be constructed: one beginning *Suppose* $x \in A\ldots$ and ending $\ldots$ *Then* $x \in B$; the other beginning *Suppose* $x \in B\ldots$ and ending $\ldots$ *Then* $x \in A$. Logically the definition involving two inclusions and the above "element-chasing" view of set equality are the same, but psychologically they are different. In our experience, students often begin with the hypotheses and attempt to deduce $A = B$ directly, without considering elements.

Another example is provided by functions. Students can have a reasonably good intuitive grasp of the meaning of one-to-one (1-1), but not know how to prove a function f is 1-1. They may even realize that the definition involves an implication, namely, that $f(a) = f(b)$ implies $a = b$, for all a and b in the domain. However, in our experience, that does not mean they know where to begin a proof. It may not be clear to them that they should normally begin the proof by almost automatically writing,

Suppose $f(a) = f(b)$ $\ldots$ for arbitrary elements a and b in the domain , and only then attempt to use the hypotheses to arrive at $a = b$. Doing this can be taken as a good sign that a student knows how to prove functions are 1-1.

4.5 Logic

---

of the notes and critiques the students' efforts. For more information, see Jones (1977) or Mahvier (1999).

Logic does not occur within proofs as often as one might expect, as we pointed out in Section 3.3. Furthermore, the logic used in proofs is mainly propositional calculus, and there is a tendency, where possible, to avoid variables and quantifiers. That is, typically proofs do not contain arguments about all elements x, but instead are about a fixed, but arbitrary element x. This cannot always be accomplished and our first sample proof (Section 1.1) is an exception in that x_1 and x_2 are universally quantified.

Where logic does occur within proofs, it plays an important role. For example, students should effortlessly be able to convert *not (p and q)* into *not p or not q*; *if p then q* into *(not p) or q*; and *if p then q* into *if not q then not p*. In addition, they should be able to draw inferences such as *q*, when given both *p* and *if p then q* (modus ponens). It seems to us that it would be useful to coordinate a student's abilities to do these various logical activities with proofs that might call on them.

In addition, there are logical and logic-like activities that connect what happens within a proof to the external context. For example, before starting to write the proof of a theorem, a student should be able to unpack its logical structure, making variables and quantifiers explicit, and converting to standard logical connectives, such as if-then. This can be difficult for many students (Selden & Selden, 1995). After that, a student needs a suitable corresponding proof framework. For direct proofs, we have illustrated this with our two sample proofs (Sections 1.1 and 2). In order to handle proofs by contradiction, students also need to formulate negations of quantified statements. They should be able to negate "for all x, $P(x)$" almost automatically to get "there is an x such that not $P(x)$," etc.

Finally, a student must be able to connect a previous theorem, or definition, with a proof in progress. This requires unpacking the theorem's, or definition's, logical structure, taking an instance thereof and adjusting the symbols to fit those used in the proof, seeing that the premises are satisfied within the proof, and writing the conclusion into the proof. This can also be viewed either in terms of student abilities or as called for by various proofs.

4.6 Problem-centered reasoning

Abilities in problem-centered reasoning are more difficult to separate out and observe than other previously discussed abilities (Sections 4.1 - 4.5). However, problem-centered reasoning plays a very large role – ultimately a dominant role – in constructing proofs, so it cannot be omitted. As background, Schoenfeld's (1985) analysis of problem solving should be very helpful, although the time scale of his observations is much less that that usually needed for constructing proofs. This may make a considerable difference. One of the points Schoenfeld makes in regard to control is that students often fail to monitor their work, continuing too long in an unpromising direction. Surely this also happens in student proof construction. However, another kind of control, persistence, can also be observed and can play a very positive role in both problem solving and proof construction. Also, as mentioned in the introduction, strategic knowledge can be very useful (Weber, 2001; Weber & Alcock, 2004).

A major factor in proof construction is the mathematics a student knows, what Schoenfeld (1985) includes in resources. The ability to bring such resources to mind in solving a specific problem or proving a specific theorem can be quite difficult for students (Selden, Selden, Hauk, & Mason, 2000). Bringing to mind appropriate

knowledge depends on both the situation, for example, the comments of others, and the interconnected nature of a student's own knowledge. Perhaps the latter could be observed, or even enhanced, through working with what might be called a *multi-concept map* – an extension of the concept map idea to include more than one "core" concept, in particular, the main concepts needed in the proof at hand. For example, in Section 4.2, we discussed two students who appeared to have some knowledge of proof by induction, but neither was able to bring it to mind until the teacher intervened.

Finally, intuition has a role in problem-centered reasoning and we suggest it would be useful to consider at least two kinds that we will call *visual-spatial* and *technical-algebraic*. By *visual-spatial* we mean intuition based on pictures or diagrams that can be sketched or visualized. These might be realistic, such a graph, or visually metaphorical. Figure 7 is an example of a somewhat realistic sketch of the composition of two functions, although one that does not include their "shape," in the way a Cartesian graph would. Figure 7 might also be regarded as a visual metaphor for a similar situation in higher dimensions that are unavailable to the human senses. In contrast, *technical-algebraic intuition* depends on one's familiarity with the interrelations within sets of definitions or theorems. In Section 4.3, we suggested that part of the difficulty of our first sample proof (Section 1.1) was due to the probable expectation that visual-spatial intuition might be useful, but actually technical-algebraic intuition about manipulating inequalities was called for.

5. Teaching

Teachers of upper division and graduate mathematics courses, such as real and complex analysis, often ask student to produce proofs as a major part of their

assessments, presumably because well-written proofs reflect a good understanding of course content. Thus, a student with only modest proving ability is at a considerable disadvantage in demonstrating understanding. We suspect that often such students do not continue their study of mathematics. Secondary education mathematics majors, in particular, are likely to realize that proofs will be required in graduate school and that they have had less practice at constructing proofs than regular mathematics majors.

How do undergraduate students currently learn to construct proofs -- a topic that is mostly part of the implicit curriculum? The only widely taught courses in the U.S. devoted explicitly to teaching undergraduate students how to prove theorems are mid-level transition-to-proof courses. Such courses typically include topics such as logic, sets, and functions -- information assumed to be a prerequisite for constructing proofs – and by necessity, these topics are taught in a somewhat abstract way, mainly outside the proving context. Only after that, are applications to constructing proofs considered. Such teaching tends to yield conceptual knowledge about proving – knowledge that does not automatically convert to the corresponding behavioral knowledge required to actually construct proofs. This suggests that many transition-to-proof courses are not as effective as they might be. Also, it appears that most of the logic taught in such courses occurs only rarely in proofs (see Section 3.3). Thus, it seems that students should be able to start proving theorems without previous instruction in logic or practice with sets and functions and develop knowledge of logic, sets, and functions through experience and mentoring during their attempts at proof construction.

Students may also learn something of proof construction from lectures on mathematical content, such as real analysis or abstract algebra. However, for teaching

proof construction, even well presented lectures may be ineffectual, because a teacher cannot know what features of proofs students are focusing on in class and probably does not know what kinds of proofs students have, or have not, already learned to construct. Take for example, Dr. T's teaching methods in real analysis, as described by Weber (2004). Dr. T, known for very good teaching, first discussed the formal-rhetorical aspects of proving[9] when presenting theorems about sets and functions. Somewhat later when considering sequential limits, Dr. T concentrated on demonstrating how to manipulate absolute value expressions (to find *N*), apparently mistakenly assuming that the students had learned and could supply the required formal-rhetorical parts of such proofs. Unfortunately, they could not.

Homework and tests in courses like real and complex analysis can also provide opportunities to learn proof construction. However, these too are likely to be ineffectual because often they do not focus directly on teaching how to construct proofs or do not include mentoring students' work. In summary, none of the current ways of teaching proof construction – transition-to-proof courses or lectures, homework, and tests in content courses – seems adequate. Indeed, it appears that many beginning graduate students at U.S. universities could benefit from a course designed to improve their ability to construct proofs.

What kind of course is likely to be reasonably effective in helping students improve their proof constructing abilities? We suggest that a good way to teach such a course is from a set of notes containing definitions and statements of theorems to prove, but with little or no additional explanation. Proofs, as well as examples and nonexamples of definitions, should be provided by the students themselves. At least at first, the

[9] Weber (2004) refers to Dr. T's teaching at this point in the course as being in a logico-structural style.

students' proof construction might best be done in class, so the teacher can provide adequate mentoring. By this we do not mean providing many heavy-handed hints, but only enough intervention for a student to succeed reasonably often and with considerable effort. Every intervention, in a sense, deprives a student of the opportunity to succeed without it. Thus it is probably best if mentoring is not available during all, or even most, of a student's proving of a particular theorem, and even that students have to compete for a teacher's attention.

The student proving and the mentoring are likely to occasionally refer to more general kinds of information, such as how to negate a universally quantified statement or how to prove a function is one-to-one. When this occurs, a teacher might organize and extend such information and add it to a set of continuing supplementary notes, so that students can refer to them for future work. Such supplementary notes should be especially useful because the information in them first arose in context in the form of behavioral knowledge (or the lack of it) and was only subsequently developed into conceptual knowledge, rather than the other way round. The supplementary notes should probably be discussed in class only very briefly and when a student brings up a question about them.

In the kind of course we are describing, students seem to do well in small groups. It may be that early on small group discussions alleviate concern over working in an unfamiliar problem-oriented situation. Also, the need to convince one's colleagues of one's ideas may enhance (problem-solving) control (Section 4.6).

In addition, where understanding previous mathematical content (for example, the definition of compactness in terms of coverings) or bringing it to mind are called for,

surely several students must have an advantage over one. All of the above suggests that working in a small group raises the probability of a student's successfully proving a theorem. We see raising the probability of success as important because the literature has little prescriptive to say about teaching students the problem-centered part of proving, but our experience strongly suggests that success breeds success.

There is another way to raise the probability of students proving a theorem that we call "long range priming" and will illustrate by referring to our first sample proof (Section 1.1). It seems clear this is a difficult proof (Section 4.3) and depends on students thinking of using minimum to find δ and the triangle inequality to complete the problem-centered part of the proof. The proof might be rendered somewhat less difficult by inserting two earlier theorems in the notes, one requiring minimum and the other requiring the triangle inequality.

The above illustration suggests the course notes should be written as the course progresses, so that the teacher knows which features of proof students are already familiar with and which kinds of proofs they have successfully constructed thus far. The idea is to provide course notes having "just in time" challenges and information.

Section 4 suggests an additional constraint on the course notes, namely, they should not be too narrow. That is, before writing the notes, one might wish to establish priorities on which abilities to include. This might even be done in a way that would ensure the supplemental notes included much of the background material in a transition-to-proof course. All of this suggests the notes may require considerable time to write and raises a question of practicality for most teachers.

Is such a course practical? Can the notes be written in a reasonable length of time and can a reasonable number of students be mentored in class? Our experience from transition-to-proof courses suggests that the mentoring would become difficult as the class size nears 35, even with groups of four. However, a properly trained student assistant could no doubt be a great help with the mentoring. For more advanced undergraduate or graduate students, the writing of notes appears to be manageable for a very small class with three groups of two students each. This class size could probably be increased to three groups of three to four students.

However, a tool could surely be built that would make writing such notes practical for a larger number of students. It is much easier to select a theorem that requires particular proving abilities than it is to write one "on demand" for a particular student. Thus, what is needed is a reference book, or database, containing branching sequences of definitions, theorems, and proofs, together with an analysis of which previous definitions and theorems, and which abilities, are used in each proof. A teacher could then select a theorem that would "stretch" a particular ability, and working backwards, include any necessary definitions and theorems needed to join the desired theorem to the current notes. Indeed it would probably be possible to have a computer program (controlled by the teacher) generate some of the notes as the course progressed.

References

Gutiérrez, A., & Jaime, A. (1998). On the assessment of the van Hiele levels of reasoning. *Focus on Learning Problems in Mathematics*, *2* & *3*, 27-46.

James, W. (1890). *The principles of psychology*. New York: Holt.

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

Konior, J. (1993) Research into the construction of mathematical texts. *Educational Studies in Mathematics*, *24*, 251-256.

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

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

Mason, J., & Spence, M. (1999). Beyond mere knowledge of mathematics: The importance of knowing-to-act in the moment. *Educational Studies in Mathematics*, 28, 135-161.

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

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

Selden, A., Selden, J., Hauk, S., & Mason, A. (2000). Why can't calculus students access their knowledge to solve non-routine problems? In A. H. Schoenfeld., J. Kaput, & E. Dubinsky, (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.

Velleman, D. J. (1994). *How to prove it: A structured approach.* Cambridge: Cambridge University Press.

Vergnaud, G. (1982). A classification of cognitive tasks and operations of thought involved in addition and subtraction problems. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), Addition and subtraction: A cognitive perspective. Hillsdale, NJ: Lawrence Erlbaum Associates.

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*, 115-133.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, *48*(1), 101-119.

Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics*, *56*(3), 209-234.