

UNPACKING THE LOGIC OF MATHEMATICAL STATEMENTS¹

ABSTRACT. This study focuses on undergraduate students' ability to unpack informally written mathematical statements into the language of predicate calculus. Data were collected between 1989 and 1993 from 61 students in six small sections of a "bridge" course designed to introduce proofs and mathematical reasoning. We discuss this data from a perspective that extends the notion of concept image to that of statement image and introduces the notion of *proof framework* to indicate that part of a theorem's image which corresponds to the top-level logical structure of a proof. For simplified informal calculus statements, just 8.5% of unpacking attempts were successful; for actual statements from calculus texts, this dropped to 5%. We infer that these students would be unable to reliably relate informally stated theorems with the top-level logical structure of their proofs and hence could not be expected to construct proofs or validate them, i.e., determine their correctness.

1. INTRODUCTION

1.1. *Related Research*

Many students arrive at university confused as to the distinction between proof and evidence. Some accept proofs as valid while feeling their conviction would be strengthened by additional empirical evidence (Fischbein and Kedem, 1982). Others simultaneously accept both inductive *and* correct deductive arguments as valid (Martin and Harel, 1989).

Much of the research on proofs has concentrated on increasing students' understandings of teacher and textbook proofs, rather than on enhancing students' ability to construct their own proofs and validate them, i.e., determine their correctness. Some authors have suggested that it is the explanatory power of proofs that makes them meaningful to students and have urged teachers to select "proofs that explain" over "proofs that merely prove" (Movshovitz-Hadar, 1988; Hanna, 1990; Hersh, 1993). Others have observed that "generic proofs" which work at the example level, but are also generalizable, are more understandable to students (Tall, 1979). In addition, it has been suggested that teachers might first present global overviews of proofs, followed by structured top-down approaches introducing concepts as needed, to avoid the appearance of "pulling a rabbit out of a hat" (Leron, 1983, 1985).

In contrast to the above studies which concentrate on proofs given to students, the "method of scientific debate" encourages first-year university

students to produce and evaluate their own proofs and “teaches them that proof is really a tool which may be used to improve ideas and separate false intuition, however natural it may appear, from true mathematical statements” (Alibert and Thomas, 1991, p. 227).

1.2. *The Current Study*

This paper takes the latter perspective, examining just a few of the many abilities university students need when creating and validating proofs themselves. In particular, we examine their ability to clarify the logical structure of mathematical statements and their ability to use such structures in the construction and validation of proofs. Deficiencies in these abilities may go unnoticed, perhaps even when employing the method of scientific debate, as the process of debate itself may clarify the logical structure of mathematical statements.

Our subjects were primarily U.S. undergraduate students who intended to continue their study of mathematics to a level requiring them to make, and determine the correctness of, proofs. They expected to be employed eventually in ways that use mathematics – some in teaching secondary school or beyond, some in industry, and some as professional mathematicians. Our data show that these students could not reliably determine the logical structure of commonly occurring mathematical statements such as calculus theorems similar to the one in Appendix 1. Indeed, their ability to discern such logical structures was much less than we expected and probably less than the expectations of most university mathematics teachers.

We also explain why students, such as our subjects, who cannot reliably determine the logical structure of statements of theorems, cannot be expected to determine the correctness of their proofs – information which is hard to obtain directly.

A student’s ability to determine the correctness of proofs is difficult to observe and is usually inferred indirectly. For example, not long ago, a mid-level undergraduate student of ours in a bridge class “proved” that $f(x)g(x)$ was increasing, given that $f(x)$, $g(x)$, $f'(x)$, and $g'(x)$ were negative, by assuming it sufficed to show $f(x)g(x)$ was positive. We took his submission of this incorrect proof as (partial) evidence that he could not reliably determine the correctness of proofs. However, most mid-level mathematics undergraduate students probably read more mathematics than they write, and what they do write may receive little detailed critique; thus, their teachers have little evidence on which to judge their ability to determine the correctness of proofs. What can one make of a calculus student who says he/she “understands” that a proof of the Fundamental Theorem of Calculus does indeed prove this theorem? It is difficult to

judge whether such a student has actually checked the proof in detail by asking and answering appropriate questions, or even whether he/she knows there are questions to be asked and answered. This paper will argue that students with poor logical unpacking ability often can neither ask nor adequately answer such questions.

1.3. *Underlying Questions*

Our data are simple and do not require sophisticated statistical analysis, and the argument we present (in Section 3.1) is straightforward. Yet, our results suggest three underlying kinds of questions. First, what does the ability to unpack the logical structure of theorems have to do with understanding or learning mathematics, or with university mathematics education generally? A second kind of question is concerned with the importance of being able to determine the logical structure of mathematical statements. Will mid-level undergraduate students' inability to do so impede their progress? Can mathematicians do so? For that matter, can mathematicians reliably determine the correctness of proofs?

Our answers to such questions depend on explicitly stated views on the nature of mathematics and its relation to mathematics education. We believe these views are consistent with the practices of many mathematicians. This suggests a third question of a different kind. How do we know our views are consistent with the practices of mathematicians? How do we know other points we will bring up concerning the practices of mathematicians are accurate? In the sense of empirical studies, we don't know. There seems to be limited scientific information on the practices and beliefs of mathematicians, so we have had to depend primarily on our own informal observations as members of the mathematical community².

1.4. *Organization of the Paper*

We begin by pointing out the sense in which we use the terms “proof” and “validation”. We next introduce the notions of “formal” and “informal” versions of mathematical statements, as well as of “unpacking” an informal statement and of “proof frameworks”. Using these ideas, we argue that those who cannot reliably determine the logical structure of theorems also cannot determine the correctness of their proofs.

We next describe underlying positions on the nature of mathematics and mathematics education. This will be done briefly, but in sufficient detail to provide answers to the questions raised above. Our discussion of the role of proof validation in learning will lead us to introduce the idea of “statement image” in order to describe what is learned through validation. Subsequently, we describe our subjects, the data collection, and the data

itself. After an analysis of the data, there is a discussion which concludes with some possible implications for teaching.

2. RELEVANT CONCEPTS

2.1. *Proof and Validation*

There are several kinds of mathematical proofs and various ways to view them. For example, there are two-column U.S. secondary school geometry proofs, proofs in the style of Russell and Whitehead's *Principia Mathematica*, proofs without words (Nelson, 1993), and proofs that explain as contrasted with proofs that only prove (Hanna, 1989). In French, there is even a distinction between "preuve" and "démonstration" (Balacheff, 1982). However, when we speak of proof in this paper we refer only to the kind of proofs U.S. university students might find in typical undergraduate texts on, for example, abstract algebra or advanced calculus, or proofs which such students might be asked to construct. Examples of theorems we have asked beginning abstract algebra students to prove include: (i) A semigroup has at most one identity. (ii) A semigroup in which the equations $ax = b$ and $xa = b$ are solvable for x is a group. Appendix 1 also includes an example of a simple calculus proof.

Such proofs can convey, and can require for their construction, various kinds of mathematical understanding. However, despite their obvious importance, it is not the purpose of this paper to discuss proof from the perspective of students' deep understandings of connections between mathematical concepts. Rather, we restrict our attention to a consideration of the logical structure of proofs – we consider proofs only from the perspective of establishing the correctness of theorems and put aside all other considerations.

We know of no prescriptive definition of proof of a kind that would be useful here, or for that matter, for students. However, proofs can be partially described as a form of argument which depends on deduction as opposed to, for example, induction or analogy. In addition, we believe only a particular style of deductive argument is normally accepted by most mathematicians as a mathematical proof. For example, although redundancy can play a useful role in nonmathematical arguments, such as in legal or philosophical writings, it appears to be largely avoided in proofs found in mathematical texts or journals. It would be interesting to analyze the stylistic characteristics that distinguish mathematical proofs from other deductive arguments and the effect these characteristics have on understanding, but again this is beyond the scope of the present paper³.

We use the term *validation* to describe the process an individual carries out to determine whether a proof is correct and actually proves the particular theorem it claims to prove. This process involves much more than just passive reading – it is often quite complicated and includes making affirming assertions, asking and answering numerous questions of oneself, and perhaps even constructing subproofs. Although it is procedural in nature, it draws heavily on conceptual knowledge. This process can take only a few moments or extend across hours, days, or even longer. Since, in validating a proof an individual calls on domain knowledge, different persons will validate the same (written) proof differently. Furthermore, despite its complexity, this is often largely a mental process and, unlike the construction of a proof, there is no final material product which can be analyzed. Thus, the validation of proofs is exceedingly difficult to observe, even though it plays a major role in mathematicians' practice and is often – some would say must be – integrated into the proof creation process. See Appendix 1 for a detailed example of validation.

2.2. *Formal and Informal Statements*

University undergraduate students who specialize in mathematics, whether planning to teach, enter industry, or continue to graduate study, eventually need to determine the logical structure of informal mathematical statements, in particular, the statements of definitions and theorems. By an *informal statement* we mean one which departs from a natural language version of predicate calculus, i.e., departs from the use of “for all”, “there is”, “and”, “or”, “not”, “if-then”, “if-and-only-if” in a significant way. We ignore minor differences in wording such as those existing between “for all”, “for each”, and “for every”, or between “there is” and “there exists”. We also regard informality as a matter of degree, with some statements more informal than others.

We consider the statement, “differentiable functions are continuous”, to be informal because a universal quantifier is understood by convention, but is not explicitly indicated. We also see “a function is continuous whenever it is differentiable” as informal because it departs from the familiar “if-then” expression of the conditional. Such statements are commonplace in everyday mathematical conversations, lectures, and books. They are not generally considered ambiguous or ill-formed, apparently because widely understood, but rarely articulated, conventions permit their precise interpretation by mathematicians and, less reliably, by students.

2.3. *Unpacking Informal Statements*

The notion of unpacking has been applied variously to symbols, graphs, and diagrams. For example, when a group is denoted by G , realizing that a more expansive notation is $(G, +)$ and evoking the group axioms, is an unpacking of the symbol G .

However, in this paper, by *unpacking* (the logical structure of) an informal statement we will mean associating with it a logically equivalent formal statement including the logical features that are understood by convention, rather than explicitly expressed, in the original statement. For example, “a function is continuous whenever it is differentiable” can be unpacked to “for every function f , if f is differentiable then f is continuous”. The result of unpacking an informal statement is not unique since there can be different, but logically equivalent, formal statements. An unpacking of the above statement could just as well have resulted in “for every function f , f is not differentiable or it is continuous”. In this paper, unpacking will always refer to associating an informal statement with a logically equivalent formal one.

For another illustration of unpacking, consider the following informal statement of a theorem which, although unfamiliar to many mathematicians, would generally be regarded as expressed in an acceptable style:

(1) *In a compact semigroup every group is contained in its own maximal group which is closed.*

Drawing on widely shared background and conventions, but not on the particular field of topological semigroups, a mathematician or an advanced university mathematics student could unpack enough logical information to see this theorem as expressible in a more formal style, perhaps as:

(2) *For every semigroup S and every group G , if S is compact and G is a subgroup of S , then there is a group H such that H is a subgroup of S , G is a subgroup of H , and H is maximal and closed.*

Our experience suggests that for teaching purposes, informal statements such as (1) are often seen as more comprehensible, i.e., more easily understood, remembered, applied, and used in reasoning, than more formal ones such as (2). It would be interesting to know who holds this view and for whom, and to what degree, it is accurate. The very commonness of informal statements suggests they may be useful, perhaps in connection with “chunking” and the efficient use of memory⁴. In addition, mathematics has both its intuitive/conceptual and rigorous/logical (i.e., synthetic and analytic) sides (Tall, 1991, pp. 13–14; Hadamard, 1945). We suspect informal statements may more nearly reflect an intuitive grasp of relationships between concepts while their more formal equivalent versions may capture the precision needed for validating proofs. In this paper we will not further

examine or speculate on the importance and uses of informal statements. Instead we will concentrate on the ability to unpack them into equivalent formal statements, as these are closely related to proof frameworks and thus play an important role in the construction and validation of proofs. We will also not discuss the interesting question of how unpacking might be accomplished⁵.

2.4. Proof Frameworks

By a *proof framework* we mean a representation of the “top-level” logical structure of a proof, which does not depend on detailed knowledge of the relevant mathematical concepts, but which is rich enough to allow the reconstruction of the statement being proved or one equivalent to it⁶. A written representation of a proof framework might be a sequence of statements, interspersed with blank spaces, with the potential for being expanded into a proof by additional argument.

To see more precisely what we mean by a proof framework, consider how one might prove that $\lim_{x \rightarrow 3} x^2 = 9$, or more formally, for every $\varepsilon > 0$, there is a $\delta > 0$, so that for each x , if $0 < |x - 3| < \delta$ then $|x^2 - 9| < \varepsilon$. Here is a sample written representation of a proof framework, where blank spaces have been left for the remaining mathematical argument needed to complete a proof.

Proof : Let ε be a number > 0 . Let $\delta =$ _____ . _____ so
 $\delta > 0$. Let x be a number. Suppose $0 < |x - 3| < \delta$. _____ .
 Therefore $|x^2 - 9| < \varepsilon$. \square

This example illustrates the close relationship between a statement, especially a formal statement, and a corresponding proof framework. Because the statement starts “For every $\varepsilon \dots$ ”, the proof framework begins by introducing ε in the first sentence. Some mathematicians might consider this step unnecessary in an actual proof and just start using ε because it appears in the formal version of what is to be proved. Because of the order of the quantifiers in the statement, ε would be understood to be independent of δ , and δ would be understood to depend on ε but not on x , determining their order in the proof framework. In addition, the “if-then” pattern in the statement leads to the “suppose-therefore” in the proof framework. There is no need to understand the meaning of absolute value to obtain the above proof framework. We mention this last point only because we are using independence from mathematical concepts as a way of indicating which portion of a proof we are calling the framework, not because such ignorance is useful. A proof framework and a corresponding formal statement

can be thought of as containing the same logical information in different linguistic forms, with each obtainable from the other.

Another example of a proof framework can be obtained from statement (2) concerning compact semigroups.

Proof Let S be a semigroup and let G be a group. Suppose S is compact and G is a subgroup of S . _____ . Let H be _____ . Then H is a subgroup of S because _____ . Also G is a subgroup of H because _____ . H is closed because _____ and H is maximal because _____ . \square

That the relationship between (2) and this proof framework is independent of knowledge of the relevant mathematical concepts can be seen by replacing “semigroup” with “set of real numbers”, “group” with “interval”, “a subgroup of ” with “a subinterval of”, and retaining the usual meaning of the other terms. This proof framework can then be expanded into a valid argument which proves an altered version of (2) – a true, but unremarkable, theorem.

Just as a theorem can have several proofs, it can have several proof frameworks, which we regard as an essential ingredient in the construction and validation of proofs – a part of Hadamard’s “precising” phase (1945, p. 56). For an example illustrating the ideas discussed thus far, see Appendix 1.

3. UNDERLYING ISSUES

3.1. *The Relationship between Unpacking and Validation*

In this section we link validation, which can be long, complex, and entirely mental, and hence very difficult to observe, with the ability to unpack logical structure, which is more easily studied. We will argue that students who cannot reliably unpack the logical structure of informally stated theorems, also cannot reliably validate their proofs. Here “reliably” is meant to indicate that a student can almost always do this, except for occasional inadvertent errors, due to such things as inattention or distraction. We are assuming that such students’ normal mathematical environment includes many informally stated theorems – this can be verified, at least for U.S. university students, by examining their textbooks.

We start our argument by considering the validation of proofs. This can be viewed as consisting of two parts, one (usually harder) is the determination that each portion of the argument follows from one’s domain knowledge and preceding portions, and another (usually easier) is the

judgment that the argument establishes the theorem, rather than some other theorem. The part of the argument which is used in making the latter judgment is what we are calling a proof framework. Being able to link a theorem with an appropriate proof framework is part of the validation process, so a student who can reliably validate the proofs of theorems must also be able to reliably recognize which proof frameworks are appropriate for which theorems. This is essentially equivalent to the ability to construct proof frameworks, since proof frameworks are relatively short and the construction process is amenable to trial and error. Thus those who can reliably validate proofs can also reliably construct proof frameworks.

We recall that a proof framework, by definition, contains adequate information to reconstruct the statement to be proved or one equivalent to it. Also, the process of finding a proof framework for a *formally* stated theorem is reversible, e.g., the “if-then” pattern in the statement of a theorem corresponds to the “suppose-therefore” pattern in a proof framework. Furthermore, in actual mathematical practice, the validation of a proof includes the judgment that the statement of the theorem could have been derived from the (perhaps partly implicit) proof framework associated with the argument. Thus those who can reliably construct proof frameworks can also build formal theorem statements from those frameworks.

Joining this conclusion with the assertion at the end of the second paragraph above, yields: students who can reliably validate proofs also can reliably unpack the logical structure of informally stated theorems. The contrapositive of this is our original claim.

We will illustrate the above argument using (1) of Section 2.3. Surely a student able to reliably validate proofs, in considering how (1) might be proved, could produce a proof framework similar to that of Section 2.4. Such a student might well view this as simply listing what is to be shown and adding a little notation. In considering what this proof framework proves, this same competent student could produce (2), which we are calling an unpacking of (1). The point is not that one would necessarily go through such a cumbersome two-step unpacking process, but that the ability to reliably validate proofs is sufficiently powerful to imply the ability to unpack logical structure. Thus, those whose unpacking abilities are deficient will also be unable to validate proofs reliably.

We note that the converse is probably not generally true. That is, a student who *can* reliably unpack informally stated theorems might still not be able to validate their proofs.

3.2. *The Importance of Unpacking and Validation*

Is it important for university mathematics majors to be able to reliably unpack the logic of mathematical statements or validate proofs? Would the continuing inability to do so impede their progress? Since the subjects of this study are, in varying degrees, in the process of becoming mathematicians, these questions raise a more fundamental issue. Is the ability to validate proofs important in the practice of mathematics?

We take the position that theorems and proofs are an essential part of mathematics and that mathematicians can construct and validate proofs. This is not an empirical claim about what mathematicians can do, but takes these abilities as part of the meaning of mathematician. We sharpen this general position by emphasizing that proving theorems and validating proofs provides a high degree of certainty and, although socially learned, can be carried out individually⁷. This does not mean that all mathematicians can prove all theorems or validate all proofs. We are simply saying that mathematicians have a cognitive ability which they can sometimes apply, and that they can normally tell when they have applied it properly.

Since we see the ability to validate proofs as central to the practice of mathematics, we also see it as important to students such as our subjects. Furthermore, applying the result of the previous section gives us that unpacking is also important.

3.3. *Understanding and Learning*

What do unpacking informal statements, constructing proof frameworks, and more generally, validating proofs have to do with understanding and learning – two fundamental concerns of mathematics education?

Understanding and learning of advanced mathematics are often viewed in terms of an individual's mental construction of concepts such as function or limit. By contrast, unpacking, constructing proof frameworks, and validation, which have a certain procedural flavor, can perhaps best be viewed as mental skills. On the one hand, they are not really part of conceptual understanding or learning per se. On the other hand, they are essential to the practice of advanced mathematics. Just as an artist must gain mastery of various techniques, so a mathematician must master these skills. One might regard the learning of them as contributing to an individual's procedural knowledge.

In addition, there is a situation in which the validation of proofs appears to actually enhance conceptual learning and understanding. Although proof validation is a way of determining whether a mathematical statement is true, it also appears to be applied by mathematicians in situations where a theorem and its proof are already believed. While we have not made

a detailed study of this practice, judging from our own experience, we believe it to be very common. Both individual mathematicians and groups of mathematicians in seminars often select a particular journal article or book and validate its proofs. For some mathematicians this may be their principal way of learning new mathematics, apart from that which they create themselves.

What is it that is learned when one validates the proof of a theorem one already believes? We suggest a partial answer – the process of validation greatly enhances what we will call one’s “statement image” of a theorem.

3.4. *Statement Images*

Most mathematics, whatever its origins, is eventually recorded and communicated using statements, including definitions, theorems, and conjectures. Here we are considering only statements having lasting significance to an individual and excluding statements such as “ n represents an integer”. Theorems, in particular, are often remembered because they state how mathematical concepts are related. To statements of lasting significance, a person may attach rich mental structures, which we call *statement images*. These are meant to include all of the alternative statements, examples, nonexamples, visualizations, properties, concepts, consequences, etc., that are associated with a statement. Such associations can arise from noticing relationships, such as seeing an example which illustrates a theorem; from repetition, such as using a theorem many times on one type of problem; or from affect, such as discovering a proof technique after many attempts. We view statement images as a unifying extension⁸ of the idea of concept images which we regard as statement images corresponding to definitions (Vinner and HersHKowitz, 1980; Tall and Vinner, 1981; Vinner, 1991).

When one validates a proof of a theorem, one’s knowledge base is altered by making or strengthening associations between mental representations of the theorem and various other features of one’s knowledge base. That is, one is constructing or enhancing one’s statement image of the theorem. Consider, for example, the calculus theorem whose validation is described in Appendix 1. It seems very likely that the validator would link mental representations of this theorem, its unpacking, its proof, and its proof framework. In addition, the triangle inequality might become associated with the theorem. Taken together, these associations become part of the validator’s statement image of the theorem. In this way, we see the enhancement of statement images as one result of proof validation. Appendix 2 illustrates another way to build up statement images.

As students progress toward becoming mathematicians their inclination and ability to validate proofs should increase. Since proof validation includes the production or recognition of proof frameworks, it is reasonable to expect statement images of theorems for more advanced students to contain proof frameworks or the associated formal statements.

Considerable emphasis has been placed on understanding the mental development and representation of mathematical concepts and one could view a person's knowledge base as consisting principally of concepts or concept images, with the relationships between them playing a secondary role. However, mathematical relationships can be quite complex and are precisely expressed in theorems. In addition, theorems appear to play a central role in mathematics – papers published in journals sometimes announce new theorems without introducing new concepts, whereas the reverse is rarely so. We find it hard to see either mathematics itself, or one's knowledge of it, as just a giant concept map with its relatively simple links (Novak and Gowin, 1984). Hence, we believe it will be useful to have a unit of analysis, such as statement image, that can correspond not only to concepts, but also to theorems, as carriers of the complex relationships between concepts.

3.5. *Summary*

We have now provided a setting, against which our data can be viewed, consisting of the ideas of informal statement, unpacking, formal statement, and proof framework. This setting, however, does not explain why the data are worth studying. To do so, we included the ideas of validation and statement image. We maintain that proofs, and hence, validations, are an important part of doing mathematics and that proof frameworks are an essential part of validations. Validations are used both to establish the correctness of theorems and for learning. Much of what is learned from validations consists of the construction or enhancement of statement images which, when they arise from validation, are likely to contain proof frameworks or their associated formal statements.

4. THE DATA

4.1. *The Question*

Can mid-level mathematics undergraduates reliably validate the proofs of theorems? Can they generate proof frameworks for informally stated theorems? Such students often have an inadequate conception of proof and may have difficulty comprehending and answering questions about proofs.

For example, given a familiar theorem, they might have trouble answering questions such as, “If you were to prove this theorem, what sort of notation would you introduce and how would you go about it?” Thus we used an indirect approach, studying undergraduates familiar with the language of predicate calculus and asking them to unpack informally written statements into that language. The results of this unpacking correspond to formal versions of the original informal statements. That is, “ $\forall x \exists y x < y$ ” is not very different from “for every x there is a y such that $x < y$ ”. If a student cannot reliably unpack the logical structure of informal statements, then, as we have argued above, he/she will lack adequate information for constructing proof frameworks or validating proofs. Our data suggest this is often the case.

4.2. *The Students and the Course*

This study was conducted at Tennessee Technological University, a comprehensive U.S. state university, with an engineering emphasis, then enrolling about 7300 students. During the 1989–1993 period of data collection, the average ACT composite score of entering freshman was about 21, well above average for the state.

Data were collected from 61 university students taking a “bridge” course designed to ease the transition from lower division, more computational, mathematics courses to upper division, more abstract, mathematics courses such as modern algebra and advanced calculus. As the intended clientele for this course are second-year students, the only prerequisite is one-variable university level calculus. However, the majority of students in our study (78.1%) were in their third- or fourth-year and had taken additional courses, such as multi-variable calculus, differential equations, and matrix algebra. A few had also taken discrete structures, probability, statistics, or complex variables. In addition, three students (4.7%) were secondary education masters students. The students were primarily mathematics (42.1%) and secondary education mathematics majors (43.8%), with an occasional engineering, accounting, computer science, or physics major. Their final grades were 23.4% A’s, 23.4% B’s, 31.3% C’s, 4.7% D’s, 9.4% F’s, 7.8% W’s or I’s.

Data were collected from three tests and five final examinations, while teaching the course six times (three times each for two instructors) from Fall Semester 1989 to Spring Semester 1993. The course was taught in a variety of ways including: whole group discussion with students presenting problem solutions and proofs at the board, lecture combined with students presenting problem solutions and proofs at the board, lecture combined with cooperative group work, and lecture combined with both cooperative

group work and students presenting problem solutions and proofs at the board. A variety of textual materials were used: one instructor's own notes, Fletcher and Patty's *Foundations of Higher Mathematics* (1988), and Kurtz's *Foundations of Abstract Mathematics* (1992), all of which were supplemented, as needed, with appropriate handouts and commentary. Classes were small and the 61 students represent the entire enrollment.

4.3. *Data on the Unpacking of Informal Statements*

We collected 94 responses on these students' ability to unpack four informal calculus statements ($n = 31$, $n = 19$, $n = 24$, $n = 20$). Just three of the 61 students repeated the course; only one responded twice (incorrectly) to the same statement. At the time the data were collected, the students had completed their coursework on propositional and predicate calculus and had done homework on similar unpacking problems.

We judged the unpacking of actual statements from calculus texts to be too difficult for these students. We also thought they might consider such problems "unfair" on tests and final exams. We thus wrote four simplified informal statements regarding calculus topics – two true, two false, but all four syntactically correct and meaningful. When we subsequently gave some of our students actual statements photocopied from calculus texts on homework and on an unlimited-time, open-book, take-home test, these proved very difficult for them to unpack.

Directions for unpacking the simplified informal calculus statements varied slightly from year-to-year according to the way problems had been discussed in class. Parentheses were sometimes used, and sometimes omitted, around $\forall a$ or $x < y$. Students were initially asked to "analyze and write the following in our notation", but were not asked to provide sets, often called "universes" in textbooks, from which instances of quantified variables were to be taken. For example, x could be used without specifying its universe, which might be, say, the real numbers. In later years, students were asked to "translate into an equivalent statement" making sure all variables were quantified and had universes. They were also given contextual information, namely, that they should assume the statements had come from a calculus student or teacher.

Below, for each statement, we give a sample correct unpacking and a sample incorrect student unpacking.

4.3.1. *Statement 1 (n = 31)*

Data came from final exams in three courses taught by two instructors. One of 31 students (3.2%) was able to correctly unpack this statement. It was:

For $a < b$, there is a c so that $f(c) = y$ whenever $f(a) < y$ and $y < f(b)$.

If f were continuous, this statement would be the Intermediate Value Theorem as stated in many calculus texts.

A sample correct unpacking would be:

$$\forall a \forall b \forall f \forall y \exists c [(a < b \wedge f(a) < y \wedge y < f(b)) \rightarrow f(c) = y].$$

One incorrect student unpacking was:

$$\forall a < b \exists c \forall f \forall y [(f(a) < y \wedge y < f(b)) \rightarrow f(c) = y].$$

This innocent-looking interchange of universal and existential quantifiers radically alters the meaning of the original statement. It now asserts the rather bizarre idea that in every interval (a, b) there is a peculiar number c , at which each nonconstant f takes on all values y between $f(a)$ and $f(b)$.

4.3.2. Statement 2 ($n = 19$)

Data came from one test and one final exam in two courses taught by two instructors. Two of 19 students (10.5%) were able to correctly unpack this statement. It was:

There is a function g such that $g' = f$ whenever f is continuous at each x .

A sample correct unpacking would be:

$$\forall f [(\forall x \text{ } f \text{ is continuous at } x) \rightarrow (\exists g \text{ } g' = f)].$$

One incorrect student unpacking was:

$$\forall g \forall f [(g \text{ exists}) \rightarrow [\forall x (f \text{ cont on } x) \rightarrow (g' = f)]].$$

4.3.3. Statement 3 ($n = 24$)

Data came from one test and one final exam in two courses taught by one instructor. Five students (20.8%) were able to correctly unpack this statement. It was:

If f is defined at a , then $\lim_{x \rightarrow a} f(x)$ exists implies f is continuous at a .

A sample correct unpacking would be:

$(\forall f \in F) (\forall a \in R) \{(f \text{ is defined at } a) \rightarrow [(\lim_{x \rightarrow a} f(x) \text{ exists}) \rightarrow (f \text{ is continuous at } a)]\}$.

One incorrect student unpacking was:

$(\forall f \text{ (in some set of functions)}) (\exists a \in \mathbb{R}) (\forall x \in \mathbb{R}) (f \text{ is defined at } a) \rightarrow (\lim_{x \rightarrow a} f(x) \text{ exists} \rightarrow f \text{ is continuous at } a).$

4.3.4. *Statement 4* ($n = 20$)

Data came from one test and one final exam in two courses taught by one instructor. Not one of 20 students was able to correctly unpack this statement. It was:

A function f is increasing on an interval I means that for any numbers x_1 and x_2 in I , if $x_1 < x_2$, then $f(x_1) < f(x_2)$.

A sample correct unpacking would be:

$(\forall f \in F) (\forall I \in \mathcal{I}) [(f \text{ increasing on } I) \leftrightarrow (\forall x_1 \in I) (\forall x_2 \in I) \{(x_1 < x_2) \rightarrow (f(x_1) < f(x_2))\}]$.

One incorrect student unpacking was:

$(\exists f \in F) ((f \text{ is increasing}) \rightarrow (\forall x_1 \in I) (\forall x_2 \in I) ((x_1 < x_2) \rightarrow f(x_1) < f(x_2))).$

In evaluating student unpackings we counted all of the equivalent forms of a statement as equally correct and very simple syntax errors, such as a missing right parenthesis, were ignored. We also ignored distinctions such as that between $\exists f$ and $\exists f \in F$. We counted as errors anything that would significantly alter what one would expect to see in a proof. For example, confusing “if” with “if-and-only-if”, might in some contexts be seen as a minor oversight, but we counted it as an error. Similarly, omitting part of a hypothesis or a conclusion was counted as an error.

Student unpacking errors included: quantifying a variable twice, not quantifying actual variables, quantifying bound variables, incorrectly employing an existential in place of a universal quantifier or vice versa, giving the converse in place of the original statement, not replacing words like “means” in Statement 4 with the correct logical connective, and quantifying an additional variable such as x in Statement 4. Often students made more than one of these errors while unpacking a single statement. Altogether, only 8.5% of the unpacking attempts (8 of 94) were completely successful. These came from 13.1% of the students (8 of 61).

4.3.5. *Informal statements from calculus texts* ($n = 24$)

Nine statements of definitions and theorems were photocopied directly from a variety of U.S. calculus texts omitting the identifying words “definition” and “theorem”. These were given to 24 of these students for unpacking as a homework assignment or as an open-book, take-home test in two of the later classes. Included were the $\varepsilon - \delta$ definition of limit,

definitions of local extrema and critical points, and Rolle's theorem. Four others are given below:

(A) If $\lim_{x \rightarrow c} f(x)$ exists and is positive, then there is an open interval containing c such that $f(x) > 0$ for every $x \neq c$ in the interval.

(B) If $f'(x) = g'(x)$ for all x in an interval I then f and g differ by a constant on I , that is, there is a constant k such that $f(x) - g(x) = k$ for all x in I .

(C) If $f(c)$ is either a local maximum or minimum of f , then either (i) $f'(c) = 0$ or (ii) $f'(c)$ does not exist.

(D) If g is continuous at c and f is continuous at $g(c)$, then $f \circ g$ is continuous at c .

One class of fourteen students was given these nine unpacking tasks as homework. Of 126 responses, just three coming from one student were completely correct. However, this fourth-year student, who got C in the course, may have received outside help on this assignment. Another student gave an almost correct unpacking of (D).

Ten students in a subsequent class were given these nine unpacking tasks as an open-book, take-home test. Since the instructor had discussed the unpacking of the $\epsilon - \delta$ definition of limit in detail prior to the test, we did not tally either correct or incorrect unpackings of this statement. Just one individual, a secondary education masters student who got a final grade of A, successfully unpacked (D) and gave almost correct unpackings of three other statements, including (B) and (C). Thus, at most 5% of these unpacking tasks (4 of 80) could be considered successfully completed. Not one student from either class successfully unpacked (A).

4.4. *Summary and Analysis*

No student in this study, the majority of whom were third- or fourth-year university students specializing in mathematics or secondary mathematics education, could consistently unpack informally written mathematical statements into equivalent formal versions using the symbols of predicate calculus with which they were familiar. At best, a few students could do so occasionally. Thus when such a student considers an informally written theorem, his/her image of it would be very unlikely to contain a formal version of it, and consequently, probably would not contain a proof framework for it. The knowledge that a written proof framework embedded in a particular argument is appropriate for establishing the truth of a particular theorem is an essential part of knowing that the argument is a proof of that theorem, i.e., of knowing that the "proof" actually proves the theo-

rem. Thus, for many informally stated theorems, these students could not have known that a given argument proved a theorem, however much they might have thought so and however well they understood the relationship between the mathematical concepts involved at an intuitive level.

5. DISCUSSION

5.1. *Beginning Undergraduates and Validation*

The subjects in this study had a number of advantages over typical beginning university students, most of whom study mathematics in service courses. They were third- or fourth-year mathematics or mathematics education majors, who had taken a number of university-level mathematics courses beyond calculus. Consequently, it is reasonable to conclude that beginning university mathematics students, lacking these advantages, would also be unable to unpack the logical structure of informal statements. This means they would neither have nor be able to generate proof frameworks in their images of theorems and would also not be able to validate proofs. In general, anything requiring logical unpacking or validation would probably be less transparent for them than might be expected. We see lack of validation skills as linked to beginning university students' well-documented inadequate conceptions of proof and discuss this in the following section. Next we discuss undergraduate students' lack of an appropriate experiential base of argumentation in school, which may contribute to their poor validation skills and inadequate conceptions of proof.

5.2. *Poor Validation Skills and Inadequate Conceptions of Proof*

Students, who lack unpacking and validation skills, would be unlikely to develop an appreciation for one of the major purposes of conveying information in the theorem-proof format, namely, that a reader can independently check a theorem's correctness. They may also have difficulty distinguishing supplementary and explanatory remarks from the proof itself, and thus lack prototypical examples for constructing the concept of proof. Because of this, it should not be surprising that they have inadequate conceptions of proof, and this expectation is supported by a number of observations. Fischbein and Kedem (1982) and Martin and Harel (1989) found that many students at this level are confused about the nature and use of mathematical proofs. Owens (1992), in a study of preservice secondary mathematics teachers, found that they saw proofs as remote and irrelevant, that this view seemed to originate with high school geometry, and that it tended to be reinforced by college work. In addition, Schoenfeld

(1985, pp. 355–357) has described undergraduates who, having just proved a geometry theorem, failed to apply the results of their own reasoning and reverted to naive empiricism when attempting a geometric construction problem essentially contained in the theorem they had just proved. Indeed such students do not seem to regard deductive reasoning as useful for solving problems or linked to common sense reasoning.

5.3. *Building an Early Experiential Base for the Concept of Proof*

We believe that both weak validation skills and viewing proofs as ritualistic, and unrelated to common sense reasoning, may be partially traceable to the absence of arguments, especially student-produced arguments, in school mathematics. School mathematics need not be devoid of such arguments. In a study from a social perspective, Yackel, Cobb, and Wood (1991) encouraged elementary students to reflect on their work. They induced them to solve problems cooperatively and to make arguments supporting their conjectures. In another study, Davis (1992) described four fourth-grade students who were given two colors of Unifix cubes and challenged to build as many three-cube towers as possible. Each was to convince the others he/she had found every possible tower with no duplications. The children found three different methods for doing this and in the process invented the idea of a situated “mathematical proof”. For these students, mathematical arguments arose naturally without explicit instruction and were probably similar to their arguments in other contexts.

Encouraging elementary school children to reflect on their own actions and give reasoned arguments may not only enhance their construction of concepts such as number and exponent, the arguments might also act as precursors to the concept of proof – precursors linked to common sense arguments. However, the continued development of student-made arguments appears to depend heavily on the culture of the mathematics classroom, the influencing of which is well illustrated in Lampert (1990).

Such a rich experiential base of argumentation could be, but usually is not, present prior to U.S. undergraduates’ first encountering proofs. What would happen if it were? Would such undergraduates’ conceptions of proof include that proofs are refinements of everyday arguments and really do assure the truth of the associated theorems? Perhaps such undergraduates would regard theorems and proofs as a natural, useful part of mathematics. Perhaps their validation skills might be better developed and their statement images of theorems might more readily contain proof frameworks.

5.4. *Implications for Teaching*

We are not the only ones to have noticed difficulties with the unpacking of informal statements. Epp observed that undergraduate students have great difficulty writing formal versions of statements involving quantifiers, such as “everyone is older than someone”. However, if she first asked students for completions of the statement “ \forall people $x \exists \dots$ ”, they were subsequently able to tackle the original statement (Epp, 1994). Her teaching strategy seems to draw on students’ existing contextual knowledge to help them enrich emerging statement images.

Moore (1991), in studying a “transition” course at the University of Georgia, found that knowing a definition and being able to give examples and nonexamples did not necessarily provide the language and logical structure to write proofs. However, learning to translate a definition into symbolic form in which quantifiers are explicit, e.g., f is one-to-one if and only if $\forall x \forall y (f(x) = f(y) \rightarrow x = y)$, helped students to see the logical structure of proofs based on that definition.

Radford (1990) gave first-year engineering students at the University of San Carlos in Guatemala geometrical and numerical theorems, together with scrambled, seven or eight-line proofs thereof. The students reassembled these proofs in a variety of mostly incorrect ways, indicating they had little appreciation for the logical structure of an argument.

The analysis presented in this paper leads us to three suggestions for teaching. First, since many undergraduate students have difficulty unpacking informally written statements, it might be useful to present theorems and definitions both in a more informal way and in a more formal way⁹. This should accommodate the needs both for intuitive understanding and for the careful validation of proofs or even validation of less formal arguments.

Second, it might be useful to offer university students some explicit instruction or advice on validation, an area currently more or less neglected. There is no reason for such instruction to be restricted to proofs. Much of the reasoning employed in determining the correctness of solutions to nonroutine problems is similar to the reasoning used in validating proofs. Asking students to explain why their proposed solutions to novel problems should be accepted or discussing why an argument proves a theorem, rather than some other theorem, might be helpful. This suggestion is similar to part of the “method of scientific debate” (Alibert and Thomas, 1991).

Finally, because university students who still have weak validation skills may not be able to distinguish proofs from supplementary or explanatory comments, it might be good to present material in a way that makes this distinction clear. For example, explanations of definitions, illustrations of

relevant concepts, cautionary remarks, and even remarks on the structure of proofs, except those needed to organize their linear presentation and help readers with validation, might best be treated as annotations. This might simultaneously provide prototypical examples for enhancing students' conceptions of proof and also encourage them to validate proofs carefully.

APPENDIX 1

A SAMPLE VALIDATION

The following example illustrates some of the ideas discussed in this paper, namely, the kind of proof we are considering, validation, informal and formal statements, the unpacking of the former into the latter, and proof frameworks. For this purpose, we have selected a familiar calculus theorem, but are not suggesting what follows should be used in teaching calculus. Nor is the example meant to illustrate how a theorem or its proof might be invented since it does not consider such things as intuition, insight, visualization, false starts, or cognitive strategies.

Although the theorem we have selected is similar to Leron's Theorem 2.2 (1983), our purpose is different – in fact, it is complementary. Leron presents the argument in a nontraditional, top-down, structured way in order to facilitate grasping its basic structure, as well as the relationship between relevant mathematical concepts. We, on the other hand, give a traditional linear presentation which facilitates the validation process. Both points of view are important in the practice of mathematics.

Neither the statement of the theorem, its unpacking, its proof framework, its proof, nor its validation are unique. These could be done in other ways. Validation, in particular, depends heavily upon the individual validator, his/her domain knowledge, and his/her knowledge of validation procedure.

We present a plausible, but imaginary, transcript detailing the main questions, answers, and comments of our (hypothetical) validator. The transcript is written entirely as internalized speech, but an actual validator might also draw some pictures, think out loud, or visualize portions of the proof. For reference, we have numbered the sentences in the proof, using brackets to enclose both reference numbers and our comments.

Theorem. $f + g$ is continuous at a point, provided f and g are.

Proof [1] Let a be a number and let f and g be functions continuous at a . [2] Let ε be a number greater than 0. [3] Note that $\varepsilon/2$ is greater than

0. [4] Now because f is continuous at a , there is a δ_1 greater than 0, such that for any x_1 , if $|x_1 - a| < \delta_1$, then $|f(x_1) - f(a)| < \varepsilon/2$. [5] Also there is a δ_2 greater than 0, such that for any x_2 , if $|x_2 - a| < \delta_2$, then $|g(x_2) - g(a)| < \varepsilon/2$. [6] Let δ be the smaller of δ_1 and δ_2 . [7] Note that δ is greater than 0. [8] Let x be a number. [9] Suppose $|x - a| < \delta$. [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] Now $|(f(x) + g(x)) - (f(a) + g(a))| = |(f(x) - f(a)) + (g(x) - g(a))| \leq |f(x) - f(a)| + |g(x) - g(a)| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. [13] Thus $|(f(x) + g(x)) - (f(a) + g(a))| < \varepsilon$. [14] Therefore $f + g$ is continuous at a . \square

Our (hypothetical) transcript follows.

[1] “OK, a , f , and g can be introduced this way because they haven’t been used before – this is the first line of the proof. Why is the argument starting with something about f and g when the theorem starts with $f + g$? Does the argument really prove the theorem? The theorem really means: For every number a and every function f and every function g , if f and g are continuous at a , then $f + g$ is continuous at a . [This is an unpacking of the original informal statement into a more formal one.] So the proof should start with an arbitrary a , f , and g and assume f and g are continuous. That amounts to [1]. After some argument, the proof should end with $f + g$ is continuous at a , which turns out to be the last line, [14]. OK, if the argument checks out, the theorem is true.” [That is, a proof framework obtained from the statement of the theorem agrees with the one embedded in this argument.]

[2] “OK, ε hasn’t come up before, so it’s permissible to let it represent any number.”

[3] [This comes from [2] and a bit of domain knowledge.] “Half a positive number is positive. So [3] is OK, but why say it?” [The answer to this would be helpful but is not essential to the validation.]

[4] “This is supposed to come from ‘ f continuous at a ’ which means: For every ε greater than 0 there is a δ greater than 0 so that for all x , if $|x - a| < \delta$, then $|f(x) - f(a)| < \varepsilon$. Here the symbols may mean something different from those in the proof. [This comes from domain knowledge, perhaps with some unpacking.] This says ‘for every ε ’ so this part can be dropped with ε replaced by any positive number. So $\varepsilon/2$ can replace ε . Also, the x can be renamed as x_1 and the δ as δ_1 throughout. That gives me [4].”

[5] “There is no ‘because’ here. But, it looks like [4]. So I can start with the definition of ‘ f is continuous at a ’ and change ‘ f ’ to ‘ g ’. Just as before in ‘ ε greater than 0’, I can replace ε by $\varepsilon/2$. Also the x_2 can replace x and δ_2 can replace δ . So [5] is OK.”

[6] “Now ‘ δ ’ has not been used before in the proof, so it can represent any number made out of numbers already there.” [It will turn out later that δ has to be treated this way, and cannot be arbitrary like ε .]

[7] [This comes from some domain knowledge.] “The smaller of two positive numbers is positive.”

[8] “OK, ‘ x ’ hasn’t been used yet so it can represent any number.”

[9] “This inequality is a statement because x , a , and δ already represent numbers and one can suppose any statement. But why?” [It turns out this supposition is part of proving $f + g$ is continuous at a . Knowing that would help the validator keep track of the argument, but it is not essential to the validation.]

[10] “The first part of this assertion follows from [9] and [6] [and simple domain knowledge – if a quantity is less than the smaller of two numbers, then it is less than either number separately]. The second part comes from the first part and the implication in [4], replacing ‘ x_1 ’ by ‘ x ’. This is permissible because [4] said ‘for any x_1 ’.” [In addition, [10] is using another kind of domain knowledge, namely, that Q can be inferred from P and P implies Q .]

[11] “The first part comes from [9] and [6], just as in [10]. The second part comes from the first part and [5], just as in [10].”

[12] “The first equality comes from $(a + b) - (c + d) = (a - c) + (b - d)$. [Some domain knowledge.] The inequality comes from the triangle inequality, $|p + q| \leq |p| + |q|$, where ‘ $f(x) - f(a)$ ’ replaces ‘ p ’ and ‘ $g(x) - g(a)$ ’ replaces ‘ q ’. Next the strict inequality comes from the last parts of [10] and [11] [and some domain knowledge, namely, $a + b < c + d$ can be inferred from $a < c$ and $b < d$]. The final equality is just arithmetic.”

[13] “This follows from [12] [and several instances of the transitive property for real numbers, e.g., $a \leq b$ and $b < c$ implies $a < c$].”

[14] “This must depend on the definition of continuous at a , applied to $f + g$. That is, for every ε , there is a δ greater than 0, such that for every x , if $|x - a| < \delta$, then $|(f + g)(x) - (f + g)(a)| < \varepsilon$.”

“Now I need to check that the definition of continuity is satisfied and that amounts to proving a theorem. [The validator, in effect, next constructs a proof framework.] Let ε be an arbitrary real number greater than 0. That’s [2] of the proof. Let δ be a real number greater than 0, which can depend on ε but not on x . That’s [6] and [7] of the proof. Let x be an arbitrary real number. That’s [8], which comes after [6], so δ doesn’t depend on x . Suppose $|x - a| < \delta$. That’s [9]. Then, after some argument has been given, $|(f + g)(x) - (f + g)(a)| < \varepsilon$. This doesn’t look like [13], but it means the same thing because $(f + g)(x)$ means $f(x) + g(x)$ and $(f + g)(a)$ means $f(a) + g(a)$.” [Here the validator does some notational unpacking.]

“So [14] is OK.”

“Now the entire argument is OK, and at the beginning, I checked that if the argument was OK, then it would prove the theorem.” [The validator decided, as part of checking [1], that the proof framework of the argument was appropriate for the statement of the theorem.]

This completes the validation.

The proof in this example is fairly detailed – some might say overly so. Many proofs omit lines like [1] which serve to “introduce” a , f , and g . An expert validator would not need [1]. The importance of [1] however is that it emphasizes the need to check that a , f , and g are arbitrary. If this were not the case – say f somehow depended on g or a – then an expert validator would notice the error even though [1] were missing. That is, he/she would act as if [1] were there should arbitrariness fail, but might not even notice its absence otherwise.

All or part of [12] might also be omitted from the proof. In that case, the validator might add the omitted portions to his/her own expanded version of the proof. In general, the need to produce additions to written proofs cannot be avoided. The writer of a proof must depend on his/her own domain knowledge and writes, in a sense, for an idealized reader. The actual validator of a proof may depend on a very different base of domain knowledge. In some cases, a validator may treat a proof as just an outline within which his/her own version of the proof is to be constructed during the validation process. In such cases validation can be creative and very long.

An actual validator might well carry out the validation process in this example much more quickly than the transcript can be read. On the other hand, if [12] were omitted from the proof, a relatively inexperienced student might take some time to realize that the key to proving [13] was the triangle inequality. Furthermore, if the need to insert an argument to justify [13] were not noticed until [14] was checked, another problem might arise. The checking of [14] is sufficiently complex that stopping for some time to justify [13] might overload working memory and cause part of the checking of [14] to be neglected. To guard against this sort of potential error, some validators make multiple passes through a proof.

The proof we have discussed here is short and “structurally simple”, consisting only of a brief proof within a proof. Many proofs are much longer and structurally more complex, and hence, require more complicated validations. As mathematics is now taught at university, advanced undergraduate students are apparently expected to acquire this complex ability to validate proofs without much direct instruction. We do not know

whether they do acquire it, but our data suggest that our mid-level undergraduate students have not yet done so.

APPENDIX 2

BUILDING A STATEMENT IMAGE

Both the idea of statement images and their construction can be illustrated by considering a (hypothetical) mathematician or advanced mathematics student, whom we call M . We assume M knows some abstract algebra, a little topology, and nothing of statement (1) concerning compact semigroups. We suggest a few plausible steps in M 's (not necessarily intentional) construction of a statement image of (1).

Suppose M first encounters (1) during a particularly long and boring colloquium and casually considers why it might be true and how it might be proved. M might first unpack (1) into (2).

M may next wonder how an apparently algebraic hypothesis could lead to a topological conclusion, i.e., that the maximal group is closed. This may cause M to think of topological groups and conclude by analogy that a compact semigroup must mean a topological semigroup, that is, that multiplication is assumed to be continuous, thereby linking the topology and the algebra. Now M has associated (1) with his/her concepts of topological algebra and topological group.

Next M might look for an example, even though he/she knows so little about topological semigroups that nothing comes to mind at first. M might nevertheless persist because finding examples helps enhance M 's feeling of understanding and often leads to ideas for a proof. Now suppose M is familiar with complex numbers and favors visualization. Eventually, after some initial period of frustration, M might think of the complex unit disk with multiplication, realize it "fits" (1), and see that the unit circle is a maximal group contained in it. This sudden visualization of an example, after a period of frustration, makes quite an impression on M and forms a strong association between (1) and the complex unit disk. In effect, the disk becomes, for M , a prototypical example – so much so that, whenever M thinks of (1) in the future, shortly thereafter he/she is very likely to visualize the disk. M might next try to construct a proof of (1).

Although this scenario illustrates what we mean by a statement image and suggests events which might lead to the formation of such an image, a discussion of the mechanisms by which the required associations are made and remembered is beyond the scope of this paper.

NOTES

1. This study was supported in part by a Faculty Research Grant from Tennessee Technological University. The authors would like to thank the referees for their many helpful suggestions.
2. We have, however, taken into consideration Livingston's (1986) ethnomethodological study of mathematicians' work, as well as Michener's (1978) study of mathematicians' understandings.
3. Toulmin (1964) has considered how the soundness of practical arguments is established, treating logic as "generalized jurisprudence" and introducing the idea of warrants which allow one to establish claims from data.
4. Such informal statements may play a role similar to that of conceptual entities and their associated mathematical notations in reducing working memory load, facilitating comprehension of complex concepts, and focusing one's attention during problem solving (Harel and Kaput, 1991).
5. One impediment to unpacking can be ambiguity. Our experience suggests that an ambiguity producing two inequivalent "unpackings" can sometimes be resolved by domain knowledge indicating that one of the two is "clearly" false. Thus, the same informal statement may be ambiguous for a student, while not being seen as so by a teacher with greater domain knowledge. For example, (1) is similar to (1') *In a compact semigroup every group is contained in a maximal group which is closed*. However, (1') is ambiguous and might be "unpacked" to either (2) or (2') *For every semigroup S , if S is compact then there is a group H such that (i) H is a subgroup of S , (ii) H is maximal and closed, and (iii) for every group G , if G is a subgroup of S then G is a subgroup of H* . As it turns out, there are many compact semigroups containing several disjoint groups. However, the identities of these could not be distinct were they contained in a single group H , showing that (2') is false. Thus a semigroup expert would fail to see the ambiguity of (1') which might arise from an alternative "unpacking" – one that could not have been intended as it is clearly false (for the expert). While we do not know the extent or importance of this kind of unpacking impediment, we do know it occurs. An early draft of this paper contained (1') instead of (1). This ambiguity was pointed out by our referees.
6. Although proof frameworks do not depend on detailed knowledge of the mathematical ideas in the statement of a theorem, they do depend, like proofs themselves, on the style in which mathematics is customarily written. Sometimes it is said that proofs are simply everyday arguments made more precise. In a sense this is true, but not every style of everyday argument is generally favored by the mathematical community. For example, we have noticed that undergraduate students first attempting proofs will commonly unnecessarily include universal quantifiers – for example, arguing about all ϵ 's simultaneously, instead of about an arbitrary, but fixed, ϵ . They also quote definitions and theorems in their proofs, in contrast to mathematicians who tend to avoid this except in lectures and explanations. Furthermore, students may not avoid redundancy and may vary the meaning of symbols within an argument. These are not necessarily errors for everyday arguments, but are usually avoided in proofs. We conjecture that the style in which proofs are customarily written by mathematicians is related to reducing cognitive load during the validation process, thereby increasing its reliability.
7. Our experience suggests that this position, which is not meant to be a complete description of mathematics, is consistent with the practice and view of a large portion of the

mathematics community that actually produces or applies mathematics, as well as of those who teach mathematics to advanced university students. There is an alternate position (Ernest, 1991) which emphasizes social aspects of the mathematics community and, to some extent, appears to deny the kind of individual ability and certainty we describe, and which might well see these features of practice as illusory. Despite enjoying considerable support amongst mathematics education researchers, this position is, we suspect, more or less unknown to most practicing mathematicians.

8. Theorems-in-action (Vergnaud, 1990) resemble 'statementless' statement images, like concept images without concept definitions. An historical example of a 'statementless' statement image is the homotopy axiom, whose main idea was implicit in "homotopic cycles are homologous" long before there was a structure in which to state the axiom (Eilenberg and Steenrod, 1952, p. 47).

In addition, it would be interesting to know if impoverished statement images could account for the lack of access to their knowledge bases which we observed in calculus students' inability to solve nonroutine problems (Selden, Selden and Mason, 1994).

9. In courses taught by the R. L. Moore, or Texas, Method, where students prove all theorems themselves, definitions and theorems tend to be stated in a more formal way than in typical textbooks. For more information on the Moore Method, see Jones (1977).

REFERENCES

- Alibert, D. and Thomas, M.: 1991, 'Research on mathematical proof', in D. Tall (ed.), *Advanced Mathematical Thinking*, Mathematics Education Library, Kluwer Academic Publishers, Dordrecht, pp. 215–230.
- Balacheff, N.: 1982, 'Preuve et démonstration', *Recherches en Didactiques des Mathématiques* 3(2), 261–304.
- Davis, R. B.: 1992, 'Understanding "understanding"', *Journal of Mathematical Behavior* 11(3), 225–241.
- Eilenberg, S. and Steenrod, N.: 1952, *Foundations of Algebraic Topology*, Princeton University Press, Princeton.
- Epp, S.: 1994, 'The role of proof in problem solving', in A. H. Schoenfeld (ed.), *Mathematical Thinking and Problem Solving*, Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 257–269.
- Ernest, P.: 1991, *The Philosophy of Mathematics Education*, Falmer Press, London.
- Fischbein, E. and Kedem, I.: 1982, 'Proof and certitude in the development of mathematical thinking', in A. Vermandel (ed.), *Proceedings of the Sixth International Conference of the Psychology of Mathematics Education*, Universitaire Instelling Antwerpen, Antwerp, Belgium, pp. 128–131.
- Fletcher, P. and Patty C. W.: 1988, *Foundations of Higher Mathematics*, PWS-KENT Publishing Company, Boston MA.
- Hadamard, J.: 1945, *An Essay on the Psychology of Invention in the Mathematical Field*, Princeton University Press, Princeton (page references are to the Dover edition, New York 1954).
- Hanna, G.: 1990, 'Some pedagogical aspects of proof', in G. Hanna and I. Winchester (eds.), *Creativity, Thought and Mathematical Proof*, Interchange 21(1), Ontario Institute for Studies in Education, Canada, pp. 6–13.

- Hanna, G.: 1989, 'Proofs that prove and proofs that explain', *Proceedings of the Thirteenth International Conference for the Psychology of Mathematics Education*, Vol. 2, Paris, France, 45–51.
- Harel, G. and Kaput, J.: 1991, 'The role of conceptual entities and their symbols in building advanced mathematical concepts', in D. Tall (ed.), *Advanced Mathematical Thinking*, Mathematics Education Library, Kluwer Academic Publishers, Dordrecht, pp. 82–94.
- Hersh, R.: 1993, 'Proving is convincing and explaining', *Educational Studies in Mathematics* 24(4), 389–399.
- Jones, F. B.: 1977, 'The Moore Method', *American Mathematical Monthly* 84(4), 273–278.
- Kleiner, I.: 1991, 'Rigor and proof in mathematics: A historical perspective', *Mathematics Magazine* 64(5), 291–314.
- Kurtz, D. C.: 1992, *Foundations of Abstract Mathematics*, McGraw-Hill, New York, NY.
- Lakatos, I.: 1976, *Proofs and Refutations: The Logic of Mathematical Discovery*, Cambridge University Press, Cambridge.
- Lampert, M.: 1990, 'When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching', *American Educational Research Journal* 27(1), 29–63.
- Leron, U.: 1983, 'Structuring mathematical proofs', *American Mathematical Monthly* 90(3), 174–184.
- Leron, U.: 1985, 'Heuristic presentations: The role of structuring', *For the Learning of Mathematics* 5(3), 7–13.
- Livingston, E.: 1986, *The Ethnomethodological Foundations of Mathematics*, Routledge, London, UK.
- Martin, W. G. and Harel, G.: 1989, 'Proof frames of preservice elementary teachers', *Journal for Research in Mathematics Education* 20(1), 41–51.
- Michener, E. R.: 1978, 'Understanding understanding mathematics', *Cognitive Science* 2, 361–383.
- Moore, R. C.: 1991, 'A conceptual-understanding scheme and the learning of proof', *Proceedings of the Thirteenth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, Vol. 2, Blacksburg, VA, 210–216.
- Movshovitz-Hadar, N.: 1988, 'Stimulating presentation of theorems followed by responsive proofs', *For the Learning of Mathematics* 8(2), 12–19.
- Nelson, R. B.: 1993, *Proofs Without Words: Exercises in Visual Thinking*, Mathematical Association of America, Washington, DC.
- Novak, J. and Gowin, D. B.: 1984, *Learning How To Learn*, Cambridge University Press, Cambridge.
- Owens, J. E.: 1992, 'Preservice secondary mathematics teachers' constructs of mathematics and mathematics teaching: Four case studies', manuscript prepared for the Contributed Paper Session, Research in Undergraduate Education, Joint Annual Meeting of the American Mathematical Society and the Mathematical Association of America, Baltimore, MD.
- Radford, L.: 1990, 'Organisations deductives et demonstration', *Proceedings of the Fourteenth PME Conference*, Vol. I, Mexico, 85–92.
- Schoenfeld, A. H.: 1985, *Mathematical Problem Solving*, Academic Press, Orlando.
- Selden, J., Selden, A., and Mason, A.: 1994 'Even good calculus students can't solve nonroutine problems', in Kaput, J. and Dubinsky, E. (eds.), *Undergraduate Mathematics Learning: Preliminary Analyses and Results*, MAA Notes, Mathematical Association of America, Washington, DC, pp. 17–26.

- Tall, D.: 1979, 'Cognitive aspects of proof, with special reference to $\sqrt{2}$ ', *Proceedings of the Third International Conference for the Psychology of Mathematics Education*, Warwick, UK, 203–205.
- Tall, D.: 1991, 'The psychology of advanced mathematical thinking', in D. Tall (ed.), *Advanced Mathematical Thinking*, Mathematics Education Library, Kluwer Academic Publishers, Dordrecht, pp. 3–21.
- Tall, D. and Vinner, S.: 1981, 'Concept image and concept definition with particular reference to limits and continuity', *Educational Studies in Mathematics* 12(2), 151–169.
- Toulmin, S. E.: 1964, *The Uses of Argument*, Cambridge University Press, Cambridge, UK.
- Vergnaud G.: 1990, 'Epistemology and psychology of mathematics education', in P. Nesher and J. Kilpatrick (eds.), *Mathematics and Cognition: A Research Synthesis by the International Group for the Psychology of Mathematics Education*, ICMI Study Series, Cambridge University Press, Cambridge, pp. 14–30.
- Vinner, S.: 1991, 'The role of definitions in the teaching and learning of mathematics', in D. Tall (ed.), *Advanced Mathematical Thinking*, Mathematics Education Library, Kluwer Academic Publishers, Dordrecht, pp. 65–81.
- Vinner, S. and Hershkowitz, R.: 1980, 'Concept images and common cognitive paths in the development of some simple geometrical concepts', *Proceedings of the Fourth International Conference for the Psychology of Mathematics Education*, Berkeley, CA, 177–184.
- Yackel, E., Cobb, P., and Wood, T.: 1991, 'Small-group interactions as a source of learning opportunities in second-grade mathematics', *Journal for Research in Mathematics Education* 22(5), 390–408.

JOHN SELDEN

Mathematics Education Resources Company

P.O. Box 2781

Cookeville, Tennessee 38502

U.S.A.

ANNIE SELDEN

Mathematics Department

Tennessee Technological University

Cookeville, Tennessee 38505

U.S.A.