CBMS Issues in Mathematics Education
Volume 7, 1998

Students' Proof Schemes: Results from Exploratory Studies

GUERSHON HAREL AND LARRY SOWDER

1. Introduction

1.1 Literature Review.

Traditionally, the mathematics taught in school was divided into two topics, algebra and geometry, where algebra was considered as a branch that deals with quantities, measurements, numerical variables, and numerical operations, and geometry as a branch that deals with space, spatial measurements, and properties of spatial objects. The idea of proof, as a deductive process where hypotheses lead to conclusions, has traditionally been stressed in the teaching of geometry but not in the teaching of algebra. Davis and Hersh (1981) pointed out that "as late as the 1950s one heard statements from secondary school teachers, reeling under the impact of the 'new math,' to the effect that they had always thought geometry had 'proof' while arithmetic and algebra did not" (p. 7). The death of the "new math" almost put an end to algebra proofs in school mathematics. This asymmetrical emphasis on proof in teaching algebra versus geometry has an historical origin: while geometry has been taught according to the Euclidean tradition established in 300 B.C., not until the 1800s did deductive aspects of arithmetic began to be stressed in the teaching or in the development of new mathematics (Davis & Hersh, 1981).

Nevertheless, the learning of proof has been a major goal of mathematics curricula in many countries and for many generations. *The Curriculum and Evaluation Standards for School Mathematics* (National Council of Teachers of Mathematics, 1989) has reemphasized this goal and has recommended that proof

This framework presented here is part of the PUPA (Proof Understanding, Production, and Appreciation) Project, supported, in part, by the National Science Foundation. Opinions expressed are those of the authors and not necessarily those of the Foundation. We thank Sonia Hristovich, Jack Tedeski, and Cara Masson for their assistance, and James McClure for his comments on this paper.

should be taught to all students and in all mathematics courses, not just geometry:

> In grades 9–12, the mathematics curriculum should include principles of inductive and deductive reasoning so that *all* students can: make and test conjectures; formulate counterexamples; follow logical arguments; judge the validity of arguments; construct simple, valid arguments; and so that, in addition, college-intending students can: construct formal proofs for mathematical assertions, including indirect proofs and proofs by mathematical induction. (p. 143, emphasis added)

Despite the efforts made by mathematics educators over decades to promote students' conception of proof, only the ablest students have achieved an understanding of it (Williams, 1980). Usiskin (1987) studied 99 high school geometry classes in five states in the U.S. and found that at the end of their geometry course, 28% of the students couldn't do a simple triangle congruence proof, and only 31% of the students were judged to be competent in constructing proofs. Senk (1985) learned that only 30% of the students in full-year geometry courses that teach proof reach a 75% mastery level in proof writing. Lovell (1971) found that a high percentage of 14- to 15-year-old students derive the truth of a general statement from a sequence of particular instances. This also was the case with as many as 80% of preservice elementary school teachers (Goetting, 1995; Martin & Harel, 1989). Galbraith (1981) reported a study in which one third of 13- to 15-year-old students did not understand the role of counterexamples in refuting general statements. Porteous (1986) pointed out that a very high percentage of 11- to 16-year-old students do not appreciate the significance of deductive proof in geometry, algebra, and general mathematical reasoning. Fischbein and Kedem (1982) showed that high school students do not understand that mathematical proof requires no further empirical verification. Vinner's (1983) study supported this result and added that high school students view the general proof as a method to examine and to verify a particular case. Martin and Harel (1989) showed that college students believe that a proof of a general statement concerning a geometric object does not guarantee that the statement is true for all instances of that object; a proof only guarantees that the statement is true for those instances where objects are spatially similar to the figure referred to in the proof. Schoenfeld (1985, p. 173) pointed out that, for the most part, college students' perspective on the role of proof is either to confirm something that is intuitively obvious or to verify something that is already known to be true.

Many mathematics educators have criticized the current approaches to teaching proof in school, arguing that findings such as those just described are inevitable consequences of these approaches. Some of these educators have suggested new ways of teaching proofs. Alibert and Thomas (1991) believe that mathematics in general, and proof in particular, are presented as a finished product; the student is not a partner in the knowledge construction, but rather a passive receiver of knowledge. They pointed out that "the conflict between the practice of mathematicians on the one hand, and their teaching methods on the other, produces problems amongst students. They exhibit a lack of concern for meaning, a lack of appreciation of proof as a functional tool and an inadequate epistemology" (p. 215). Alibert and Thomas suggest that we consider these problems in the light of research which has been carried out, paying "particular attention to studies emphasizing the nature of proof as an activity with a social character, a way of communicating the truth of mathematical statement to other people, helping them to understand why it is true" (p. 216).

This suggestion is consistent with the view expressed by Usiskin (1980), who argued that "we seem to have failed in our teaching of proof, because we too often ignore when and why mathematicians do proofs, the variety of possible types of proof, and how mathematicians write down proofs" (p. 419). The impression students get from geometry classes—the only place where they are exposed to the idea of proof—is distorted, according to Usiskin. First, most of the mathematician's work is spent on exploring and conjecturing, not on searching for proofs of well stated propositions, and certainly not on obvious propositions, as is often done in geometry classes. Second, there is a great variety of proof types, not just geometry proofs; for example, "proofs using mathematical induction are different from proofs of trigonometric identities, which are different from epsilon-delta limit proofs, which are different from the proofs found in abstract algebra, and so on, with differences in every branch of mathematics" (p. 420). Third, mathematicians do not write proofs in two columns. Usiskin's recommendations were to delete "rigorous" proofs of obvious statements and allow informal proofs with less demand on formal ways of writing.

Motivated by the van Hiele model of levels of geometric thinking, Shaughnessy and Burger (1985) arrived at a similar conclusion: "Students' introduction to geometry should be informal, without formal proofs or axiomatic treatment, for at least one-half year. . . . Activities that encourage inference and deduction should also be included, but the writing of carefully structured formal proofs should be omitted" (p. 426).

This approach was recommended by others: MacPherson (1985) suggested that to develop the vocabulary and concepts needed for the development of proofs, students must be engaged first with informal geometry which is based on problem solving and hands-on activities. Semadeni's (1980) idea of "action proof" is in this direction as well: Students first prove for a specific case through manipulations, such as drawing pictures, then use more specific cases, and finally visualize the actions in their minds and generalize. Other recommendations, in line with those described here, were discussed by Hanna (1990). She indicated that the works by Leron (1983), Volmink (1988), Movshovitz-Hadar (1988), and Alibert (1988) suggest teaching approaches that do away with rigor and formality and focus on aspects of communication and social processes.

Our own research on the concept of proof corroborates many of the observations described here. For example, one of the conclusions coming from this research is that a major reason that students have serious difficulties understanding, appreciating, and producing proofs is that we, their teachers, take for granted what constitutes evidence in their eyes. Rather than gradually refining students' conception of what constitutes evidence and justification in mathematics, we impose on them proof methods and implication rules that in many cases are utterly extraneous to what convinces them. This begins when the notion of proof is first introduced in high school geometry. We have demanded, for example, that proofs be written in a two-column format, with formal "justifications" whose need is not always understood by a beginning student (e.g., Statement: $AB = AB$. Reason: Reflexive Property). Also, we present proofs of well-stated, and in many cases obvious, propositions, rather than ask for explorations and conjectures. As a consequence, students do not learn that proofs are first and foremost *convincing* arguments, that proofs (and theorems) are a product of human activity, in which they can and should participate; that they are an essential part of doing mathematics. This is in essence the whole thrust of our teaching treatments. The goal is to help students refine their own conception of what constitutes justification in mathematics: from a conception that is largely dominated by surface perceptions, symbol manipulation, and proof rituals, to a conception that is based on intuition, internal conviction, and necessity.

1.2 Research Goals and Data.

The questions we are addressing revolve around the development of college students' proof understanding, production, and appreciation (PUPA). What are students' (particularly mathematics major students') conceptions of proof? What sorts of experiences seem effective in shaping students' conception of proof? Are there promising frameworks for teaching the concept of proof so that students appreciate the value of justifying, the role of proof as a *convincing* argument, the need for rigor, and the possible insights gained from proof? In answering these questions, we hope, through an extended series of studies, to be able to do the following:

(a) map students' cognitive schemes of mathematical proof,

(b) document the progress college mathematics students make in their conception of mathematical proof, in a typical undergraduate mathematics program,

(c) offer developmental models of the concept of proof among mathematics majors, in a teaching environment that is based on epistemological and pedagogical principles advocated by the current research in mathematics education, and

(d) offer principles for instructional treatments that facilitate proof understanding, production, and appreciation.

This paper is our first report on this broad investigation. The paper's main focus is item (a)—students' schemes of mathematical proof—but inevitably it touches upon the other three items as well. These schemes, or the lack thereof, were derived from the following sequence of teaching experiments:

NT: One-semester teaching experiment in elementary Number Theory taken by sophomore and junior mathematics majors ($N = 32$);

CG: One-semester teaching experiment in College Geometry taken by junior and senior mathematics majors ($N = 25$);

LA1,LA2: Two consecutive one-semester teaching experiments in elementary Linear Algebra taken by sophomore mathematics majors ($N = 23$, $N = 27$, respectively);

LA3: One-semester teaching experiment in advanced Linear Algebra taken by junior and senior mathematics majors ($N = 20$); and

EC: A case study of a precocious junior-high school student studying Euclidean Geometry and Calculus.

All these teaching experiments were taught by the first author of this paper. The data on NT, CG, LA3, and EC were collected from classroom observations in the form of field notes and retrospective notes, clinical interviews, team homework, individual homework, and written tests. The data in LA1 and LA2 were more extensive and included the following sources: (a) classroom sessions, which were all video-taped and transcribed; (b) classroom observations by graduate students, who recorded the classroom interactions, and small-group discussions; (c) 60-to-90-minute clinical interviews with students, which were all video-taped and recorded, and (d) students' homework, quizzes, and written tests.

The system of proof schemes reported in this paper has undergone numerous revisions dictated by the results from our qualitative analysis data, cross-checked through interviews of mathematics majors at a separate institution. The current version of this system's structure and components seems to have reached a stable stage. By this we mean in completing the analysis of about 50% of the data, we discovered no additional categories of proof schemes and none of the existent categories has been altered.

We characterize the results of this research as exploratory. The system of proof schemes described here must be validated by other researchers through multiple teaching experiments taught by various instructors in various institutions. Also, this report is not a full account containing all the empirical evidence, but one that sets the stage for such accounts by laying out the landscape and descriptive vocabulary. It is beyond the scope of this already lengthy paper to provide a detailed analysis of each category in this system of proof schemes. In this paper we restrict our report of the data to a collection of teaching experiment episodes, whose sole goal is to demonstrate the different categories of proof schemes we have observed.

We present a total of 24 episodes to describe a system of 16 subcategories of proof schemes which we observed with 128 students in 6 teaching experiments. These episodes are by no means isolated instances; each is an example of a phenomenon repeatedly observed in several teaching experiments. The detailed analyses of the exact quantitative accounts of each proof scheme are underway; we intend to provide them in a series of separate publications.

The reader has surely noticed the obvious paradox inherent to this research, namely, it applies empirical methods to validate the absence or presence of logico-mathematical reasoning.

2. The Process of Proving: A Definition

This section begins with a short sketch of the historical evolution of the concept of proof, followed by a cognitive analysis of this concept.

2.1 Proof: A Historical Perspective.

The view of what constitutes an acceptable mathematical proof has had many turning points. Babylonian mathematics is considered proof-free by current standards, because it does not deal with general statements, deduction, or explanations; rather, it prescribes specific solutions to specific problems.[1] The axiomatic method—that is, the notion of deductive proof from some accepted principles—was conceived by the Greeks. Its emergence is attributed to different factors (Kleiner, 1991): mathematical (e.g., resolving contradictory computational results obtained by earlier civilizations), social (e.g., Greek democracy required the skills of argumentation, which encouraged deductive reasoning), philosophical (Greeks' philosophical inquiries demanded the formulation of primitive assumptions and their logical implications), and pedagogical (the need to teach caused the Greeks to structure mathematics in a logical order).

The insistence of the Greeks on well-defined concepts and rigorous reasoning prevented them from using certain ideas, such as irrational numbers and infinity, which proved indispensable to the subsequent development of mathematics. This, in turn, was followed by a long period of mathematical activity with little attention to rigor. The most notable segment of this period is the sixteenth to eighteenth centuries, where symbolic notation and symbol manipulation were key methods of demonstration and discovery. These methods, despite their lack of rigor, led to major results which proved to be invaluable in solving physical problems and modeling physical phenomena. The application power of the results was a source of validity and legitimacy, and by implication the validity of the methods by which they were obtained (Kleiner, 1991).

[1] Bernal (1987) suggests that this characterization is indicative of the general "distancing" of Greek mathematics from Bablylonian mathematics. According to Bernal, this may in part be a reflection of mid-nineteenth and early twentieth-century historians' ethnic biases.

The end of this period may be marked by Fourier's "symbolic solution" to the Flow of Heat problem. Fourier reduced this problem to that of taking an even function and expressing it as an infinite sum of cosines, without attending to the meaning of infinite summation of functions. His solution led to observations which seemed at the time inconsistent with "regular" behavior of functions. This, in turn, fed by earlier doubts of the validity of ideas such as fluxions and infinitesimals (Boyer, 1989), led to thorough investigations into the assumptions of calculus and inspections of its structure, whereby the entire calculus was reconstructed into a new mathematical field, which now is called "analysis." With this reconstruction, the axiomatic method reemerged. Investigations into analysis led to questions on the foundations of real numbers. This in turn led to questions on the foundations of the positive integers, which were addressed in different ways in the late nineteenth century by Dedekind, Peano, and Frege (Boyer, 1989).

2.2 Proof: A Cognitive Perspective.

In what follows, we introduce the notions of "observation" and "certainty in an observation" to define "process of proving," "proof scheme," and a few other notions essential to our classification of proof schemes.

2.2.1 Observations.

Human beings live by the observations they make every moment they are conscious. An observation can be mere recognition ("It's raining now"), it may be based on a life-time of experience ("It always rains in the summer"), or it may take generation after generation to make ("Euclidean geometry is not an absolute truth"). For some observations, we still wonder, and may never know, whether they are true ("Every even number greater or equal to 4 is the sum of two prime numbers"). Despite any passive or sensory connotation of the term, observations are results of people's constructions, not mere preexisting, transmitted knowledge.

All observations—except possibly those which are simply a replay of past experience (i.e., recognitions)—are novel for the person who makes them. When an individual notices a relationship, it is a novel observation for her or him, because until then, he or she had never realized it. Observations do differ in three important characteristics: *originality*, *mode of thought*, and *certainty*. We briefly describe the first two and elaborate on the latter, which is the focus of our investigation.

2.2.2 Originality and Mode of Thought.

Originality refers to the following distinction: When a student solves a certain problem, it is one thing if he or she produces the solution on her or his own, and it is another if the student reproduces a solution that was communicated to her or him by others. Accordingly, some observations may be called *innovative*, others *imitative*. An observation is innovative if it originated with the observer (the person who made the observation), and it is imitative if it was communicated to the observer by others. This distinction must be tempered with what we know about ways of knowing: First, people make observations through interactions with their environments, which include social interactions; thus, no observation can originate with the observer alone. Second, and consistent with Piaget's theory, imitative observations are not copies of observations made by others, but the results of reconstructions by the individual (Piaget & Inhelder, 1967).

Mode of thought refers to how observations are made—for example, an observation can be made by abstracting a phenomenon from several empirical observations or by thought experiments with no mediation of empirical observations (e.g., Einstein's special theory of relativity).

2.2.3 Certainty.

Significant to this paper is not so much the kind of observations (originality) or how observations are conceptually made (mode of thought) but rather how observations are evaluated. An observation can be conceived of by the individual either as a *conjecture* or as a *fact.*

> *A conjecture is an observation made by a person who has doubts about its truth. A person's observation ceases to be a conjecture and becomes a fact in her or his view once the person becomes certain of its truth.*

This is the basis for our definition of the *process of proving*:

> *By "proving" we mean the process employed by an individual to remove or create doubts about the truth of an observation.*

The process of proving includes two subprocesses: *ascertaining* and *persuading*.

> *Ascertaining is the process an individual employs to remove her or his own doubts about the truth of an observation.*
>
> *Persuading is the process an individual employs to remove others' doubts about the truth of an observation.*

Central to this paper is the question:

> *How are conjectures rejected or rendered into facts?*

Before we discuss this question, the following clarifications and qualifications to the above definitions are essential.

First, these definitions may imply that an observation remains a conjecture until the person reaches absolute certainty in its truth. It must be noted that conjectures are usually not viable without a certain degree of conviction in their truth. A person can be uncertain about, yet believe in, the truth or falsity of an observation. A student may be convinced about the truth of an observation, but not quite certain until, for example, her or his teacher has confirmed the observation, or until further evidence is provided. In some cases this evidence may look unnecessary to an outsider, but to the student it may be indispensable.

For example, Fischbein and Kedem (1982) noticed that students, even after producing a deductive proof of a given proposition, wished to check it in a few special cases, an indication that they continued to be uncertain about its truth.

Second, conjectures vary in the degrees of faith a person has in their potential truth. The amount of effort a person is willing to make in seeking evidence that would render the conjecture a fact (or refute it) is in some proportion to her or his faith in the truth (or falsity) of the conjecture. Also, if the person's faith in the truth of a conjecture is high, he or she would seek evidence to establish its certainty; conversely, if the person's faith is low he or she would seek evidence to refute the conjecture. This natural human behavior dictates the direction in which mathematicians choose to pursue their research. For example, one of the most famous, still open, problems of this century states: Does every bounded linear operator on a complex Hilbert space of dimension greater than 1 have a non-trivial closed invariant subspace? Halmos tells in his autobiography how his presumption about the solution of this problem affected his direction of research:

> I . . . conjectured that [quasitriangular] operators . . . all have invariant subspaces. . . . The Rumanian school . . . proved that the non-quasitriangular operators . . . [have] invariant subspaces; if my conjecture turns out to be right, then the invariant subspace problem is solved affirmatively. I admit that that rocked me. Since I am convinced that the solution of the invariant subspace problem is negative, I abandon my conjecture, and shall proceed instead to search for a quasitriangular counterexample. (Halmos, 1985, pp. 320-321)

Similarly, Lobachevsky wrote on his motives in doubting Euclidean geometry which until his time was a fact for the mathematics community for over two thousand years:

> The futile efforts since Euclid's time on throughout two thousand years have compelled me to suspect that the concepts themselves do not contain the truth which we have wished to prove. . . . (quoted in Alexandrov, 1963, p. 101)

Third, the proof schemes held by an individual are inseparable from her or his sense of what it means to do mathematics. If, for example, a student holds the view that mathematics is just a collection of truths, then he or she is likely to ascertain herself or himself and persuade others of the truth of an observation by an appeal to authority, such as a textbook or a teacher.

Last, and most importantly, as defined, ascertaining and persuading are entirely subjective and can vary from person to person, civilization to civilization, and generation to generation within the same civilization.

As was mentioned earlier, the methods of ascertaining and persuading employed in the Babylonian mathematics are strikingly different from those employed in Greek mathematics. While the former consists of specific prescriptions for specific problems, the latter is characterized by its attention to rigor. The European mathematics of the sixteenth and seventeenth centuries, on the other hand, employed methods that are far less rigorous than those employed by the Greeks. Within the Western culture, the methods of proof of the sixteenth to eighteenth centuries are markedly different from those of the nineteenth and twentieth centuries.

Methods of proof may differ even within the same culture during the same period. The emergence of formalism and constructivism—the two main competing mathematical philosophies of this century—attests to this fact. Constructivism, founded by Brouwer, views the natural numbers as the fundamental objects that can not be reduced to further basic notions, and so any meaningful mathematical proof must ultimately be based *constructively* on the natural numbers. An example of a corollary of this premise is that one cannot establish the truth of an argument by showing that its negation leads to a contradiction, for no construction that is based on the natural numbers is involved in such a demonstration. Formalism, founded by Hilbert, offers mathematical certainty by turning mathematics into a meaningless game. Any proof must start with undefined terms and accepted statements (axioms) about these terms. The meaning that one may assign to these terms and axioms is irrelevant to the subsequent logical deductions (theorems) from them. As the undefined terms and axioms, the theorems too, as far as mathematics is concerned, have no content until they are supplied with an interpretation.

In this paper we are interested mainly in how individuals prove or justify; more specifically, in how students ascertain for themselves or persuade others of the truth of a mathematical observation. In this regard, we assert that a person's ascertaining and persuading processes for an observation may be based on logical and deductive arguments, empirical evidence, intuitions, personal beliefs, an authority (e.g., an opinion is true because the teacher said so), social conventions, or any other knowledge the person considers as relevant to the truth of the observation. Further, a person can be certain about the truth of an observation in one situation, but seek additional or different evidence for the same observation in an another situation. For example, long before students learn geometry in school, they are convinced, based on personal experience and intuition, that the shortest way to get from one point to another is through the line segment connecting the two points. Later, as participants in an Euclidean geometry class, an instantiation of this observation—stated in the theorem "The sum of the lengths of two sides of a triangle is greater than the length of the third side" —may become a conjecture for the students until they find evidence that would be accepted by their class community or their teacher. The kinds of evidence the students may look for are based on whatever conventions are accepted in their class as evidence for a geometric argument. These conventions may differ from one class to another; for example, what might be accepted as evidence in a standard high school Euclidean geometry class is likely to be insufficient evidence for a college class studying axiomatic geometry.

In short, people in different times, cultures, and circumstances apply different methods to remove doubts in the processes of ascertaining and persuading. Accordingly:

> *A person's proof scheme consists of what constitutes ascertaining and persuading for that person.*

3. Classification of Proof Schemes

Our definitions of the process of proving and proof scheme are deliberately psychological and student-centered. In the remainder of this paper, *the verbs "to prove" and "conjecture" and the nouns "proof," "conjecture" and "fact" will always be used in this (subjective) sense and according to the definitions stated above.*[2] This interpretation is essential in order to understand what we intend to communicate to the reader.

Each of the categories of proof schemes in our classification represents a cognitive stage, an intellectual ability, in students' mathematical development, and all were derived from our observations of the actions taken by actual students in their process of proving.

Thus, this classification is *not* of proof content or proof method, such as the classification made by Usiskin (1980)—geometry proofs versus proofs by mathematical induction, versus proofs of trigonometric identities versus epsilon-delta limit proofs versus abstract algebra proofs, and so on. Nor is this classification a priori, such as that suggested by Hanna (1990), who distinguished between "proofs that explain versus proofs that prove," and according to whom proof by mathematical induction, for example, is a proof that explains, not a proof that proves. Such an a priori determination is not based on the individual's conception of what constitutes proofs and what constitutes refutations.[3] In contrast, it is the individual's scheme of doubts, truths, and convictions, in a given social context, that underlies our characterization of proof schemes. Nor is the classification about philosophical stands, though different proof schemes may be suggestive of various mathematical philosophies. The question of the alliance between our proof scheme system and the historical epistemology of the concept of proof (e.g., as in Kleiner, 1991) is beyond the scope of this paper, and it will not be addressed except for brief mentions in what follows.

Three categories—each with several subcategories—of proof schemes are the product of the work so far (Figure 1). It is important to note that these schemes are not mutually exclusive; people can simultaneously hold more than one kind of scheme.

---

[2]This restriction applies to other similar terms, such as "to justify," "to show," "justification," "verification," etc.

[3]We recognize, however, the importance of the pedagogical implication alluded to by Hanna's distinction. It conveys the message that a proof must be a convincing argument, not just a sequence of logically-inferred statements.

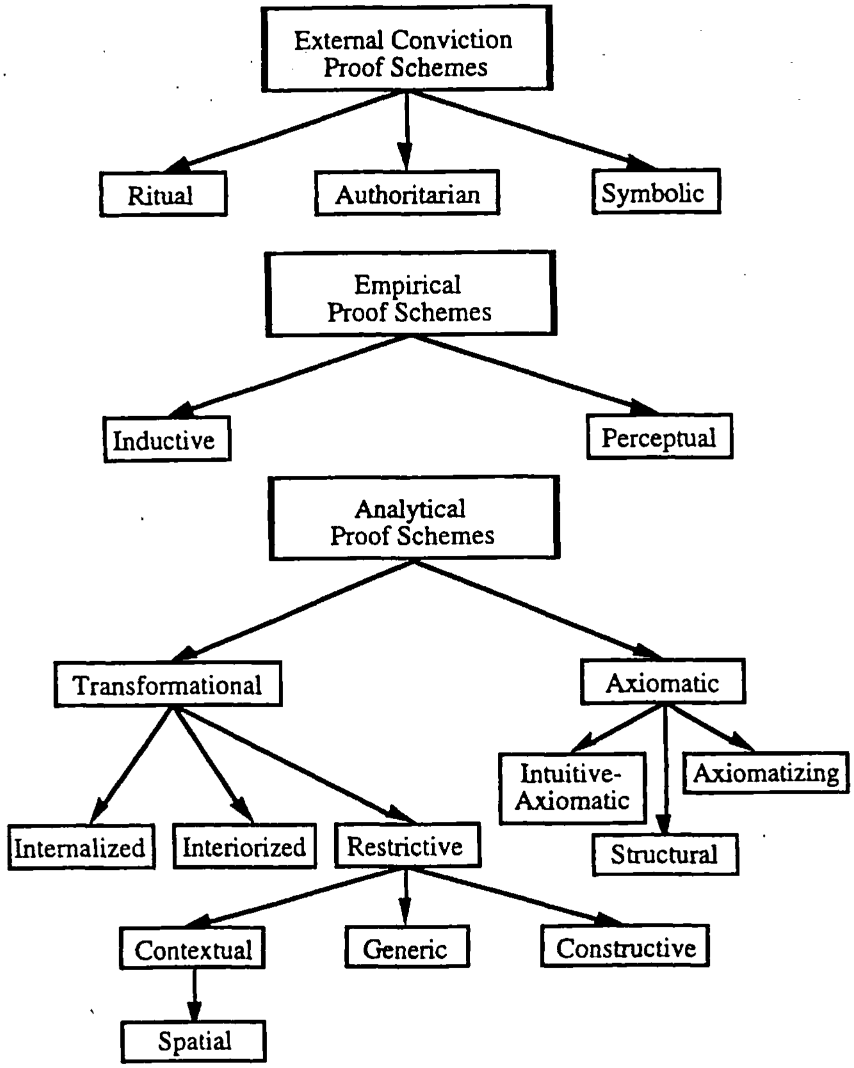

FIGURE 1

3.1 External Conviction Proof Schemes.

When formality in mathematics is emphasized prematurely, students come to believe that ritual and form constitute mathematical justification. When students merely follow formulas to solve problems, they learn that memorization of prescriptions, rather than creativity and discovery, guarantee success. And when the teacher is the sole source of knowledge, students are unlikely to gain confidence in their ability to create mathematics. These learning habits are believed to lead to the formation of *external conviction proof schemes*—schemes by which doubts are removed by (a) the ritual of the argument presentation—the ritual proof scheme, (b) the word of an authority—the authoritarian proof scheme, or (c) the symbolic form of the argument—the symbolic proof scheme.[4]

3.1.1 Ritual Proof Scheme.

Martin and Harel (1989) addressed the question: "Are students' judgments of an argument influenced by its appearance in the form of a mathematical proof—the ritualistic aspects of proof—rather than the correctness of the argument?" (p. 42). They addressed this question by examining students' responses to a false-proof "verification" that looked like a deductive proof of a certain mathematical statement. They found that "many students who correctly accepted a general-proof verification did not reject a false-proof verification; they were influenced by the appearance of the argument—the ritualistic aspects of the proof—rather than the correctness of the argument" (p. 49).

Accepting false-proof verifications on the basis of their appearance is a severe deficiency in one's mathematical education, which is possibly attributable to the over-emphasis in schools on proof writing prior to and even in place of proof understanding, production, and appreciation. The ritual proof scheme, however, does not need to manifest itself only in this severe behavior of judging mathematical arguments on the basis of their surface appearance. For example, often during the beginning period of a teaching experiment, either in a class discussion or in a personal exchange, students have asked whether a certain justification is considered a proof. When asked to explain the motivation for their question, the students indicate that although they are convinced by the justification, they have doubts whether it counts as a mathematical proof, for "it does not look like a proof." Typically, such doubts are raised when the justification is not communicated via mathematical notations and does not include symbolic expressions or computations, such as in the following episode from LA1, which took place shortly after the unit on linear independence was completed.

Episode 1. In response to a test question,

A, B, C, D, E are linearly dependent. If we add a new vector S to this set of vectors, would the vectors A, B, C, D, E, S be dependent or independent?

Andy wrote:

Adding a new vector S to a set that is currently linearly dependent will not change the dependency of the entire set. The set will remain linear[ly] dependent. This is because one vector can be expressed as a linear combination [of the other vectors] in the set (by definition), therefore the entire set remains linear[ly] dependent due to this fact.

Following this test, during an interview with this student, he expressed concern about the fact that his proof did not include any symbolic expressions and therefore he had doubts whether it is mathematically valid, for, according to him, it

[4]A tendency to form external conviction proof schemes may also reflect certain basic psychological tendencies of an individual of the sort discussed by Frid (1994), who found that some students tended to "learn" mathematics by collecting "isolated, relatively unconnected mathematical statements, rules, and procedures" (p. 77).

didn't look like a proof. The possibility that this student's concern was the level of formality of his proof is dismissed on the grounds that his other justifications which included symbolic expressions were acceptable to him as proofs despite their lack of formality and generality.

3.1.2 Authoritarian Proof Scheme.

Why do so many students lack the intellectual curiosity to wonder why a theorem or a formula is true? We believe the answer to this question lies in the fact that current mathematics curricula emphasize truth rather than reasons for truth. This begins with elementary mathematics where children are rushed into using mathematical prescriptions to solve arithmetic problems (Harel, 1995) and continues with secondary and postsecondary mathematics where instrumental understanding rather than relational understanding is emphasized throughout the curriculum.[5] As a consequence, students build the view of mathematics as a subject that does not require intrinsic justification. Although students understand that the mathematics they do must be true, they are not concerned with the question of burden of proof; their main source for conviction is a statement appearing in a textbook or uttered by a teacher. Such a conception of proof we call *authoritarian*. We observed five different kinds of manifestations of the authoritarian proof scheme in students' mathematical behavior. The first and most common expression of this proof scheme is students' insistence on being told the procedure to solve their homework problems, and when proofs are emphasized, they expect to be told the proof rather than take part in its construction. *The underlying characteristic of this behavior is the view that mathematics is a collection of truths, with little or no concern and appreciation for the origin of the truths*. The following examples demonstrate that memorizing and applying ready-made formulas are what students expect to do in mathematics.

Episode 2. During the first week in the Teaching Experiment LA3, the students were assigned to prove "$Null(A) \subset Null(B)$ for any two matrices A and B for which the product is defined." In the class session that followed, three students who worked as a group on this assignment complained that they went through all their notes, including relevant material in their textbook and notebook from their previous linear algebra class, but found no theorem that could tell them how to prove this statement.

Episode 3. In one of the class sessions in LA3, toward the end of this course, the students were asked to complete the proof of a certain proposition by justifying some of its steps. In the class session that followed, Bob, one of the students in this class, asked the instructor to show how to complete these steps. The instructor responded by asking Bob to first share with the class his thoughts on the problem and his attempts to solve it. It turned out that neither Bob nor any

[5]The notions of "instrumental understanding" and "relational understanding" are from Skemp (1978). The former means knowing the "how" whereas the latter means knowing both the "how" and the "why."

one else in the class had worked on this assignment. The instructor, therefore, declared that the discussion of this assignment would be postponed to the next class session after the students have tried completing the proof on their own. Bob expressed his dissatisfaction with the instructor's decision by saying that he does not understand what difference it makes if he was told the proof or if he found it on his own; the end result in both cases is the same: he would know the proof.

The second expression (and consequence) of the authoritarian proof scheme is this: It is not uncommon that students ask for help on a certain problem without first making a serious effort to solve it on their own. Often in such cases, after a brief discussion of the problem the students realize that they in fact were capable of solving it on their own, but needed the presence and confirmation of an *authority* to arrive at their solutions, as in the following episode.

Episode 4. In the teaching experiment LA3, a student came to consult with the instructor of the class about a homework problem requiring her to generate examples of non-diagonal diagonalizable matrices of different sizes. The student complained that the method we had established in class did not work, for each time she took a diagonal matrix Λ and an invertible matrix S, the product $S\Lambda S^{-1}$ resulted in the same diagonal matrix Λ she began with. After a very short conversation, she realized on her own the reason for the outcome; namely, that she had been choosing matrices Λ with equal entries on the diagonal. This by is no means an exceptional case. If anything is exceptional about it, it is the fact that this student had traveled quite a distance from her home to campus, exclusively—as she had indicated—to consult with the instructor about this problem.

The third expression of the authoritarian proof scheme was first observed in LA1. It demonstrates the mystical power of the term "theorem" on students' process of proving.

Episode 5. Consistent with the instructional perspective on the concept of proof introduced earlier is the stand that a mathematical relationship should not be formulated as a "theorem" until it has been motivated, debated, and applied—a standard practice of the teaching experiments. In LA1 we first noticed that until a mathematical relationship is declared a theorem, the students continued—either voluntarily when they needed to use the relationship or upon request—to justify it. But once the relationship was stated as a theorem, there seemed a reduced effort, willingness, and even ability with some of the students to justify it. It seems as though, for these students, the label "theorem" renders the relationship into a formula—something to obey rather than to reason about. The fourth manifestation of the authoritarian proof scheme is an extension of the former. Often students prove a certain statement by rephrasing it into a statement that for them is a fact. Consider the following two episodes:

Episode 6. Lee, a student in LA1, was given the following problem:

M, N, K, R, V are linearly dependent.

(a) If we remove two of these vectors, say M and N, would the remaining vectors be dependent or independent?

(b) If we add a new vector S to this set of vectors, would $\{M, N, K, R, V, S\}$ be dependent or independent?

Lee responded:

(a) They would be dependent because you could express these vectors as linear combinations of each other.

(b) The vectors would be dependent because you could express S as a linear combination of the others.

The following additional episode of the authoritarian proof scheme is particularly interesting; it justifies a conjecture by simply saying its contrapositive.

Episode 7. Don, a student in LA2, responded to the question,

"Why must $rank(A) = rank([A|b])$ for any system $AX = b$ that has a solution?"

by saying:

$rank(A) = rank([A|b])$ because otherwise our system would be inconsistent. If a system is inconsistent it can't have a solution. Therefore, a system must be consistent to have a solution.

A fifth manifestation of the authoritarian proof scheme is the reluctance of students to ask questions about the instructor's motivation and the reason for his thinking, or to challenge his assertions even when they suspect them to be incorrect, as Episode 8 demonstrates. This occurred despite the major effort on the part of the instructor to establish a classroom atmosphere that was conducive to open discussion and questioning.

Episode 8. In one of the class sessions in teaching experiment LA2, while reviewing a set of problems in a preparation for the midterm exam, the following problem was discussed:

"Find a vector e that is not in the span of

$$a = \begin{bmatrix} 1 \\ -2 \\ 3 \\ 1 \\ -5 \end{bmatrix}, \quad b = \begin{bmatrix} -5 \\ -21 \\ 13 \\ -11 \\ 5 \end{bmatrix}, \quad c = \begin{bmatrix} -9 \\ -1 \\ -2 \\ -4 \\ 12 \end{bmatrix}, \quad d = \begin{bmatrix} -52 \\ -22 \\ 14 \\ 71 \\ 51 \end{bmatrix}.$$

Eric, one of the students in this class, suggested first row-reducing the matrix $A = [a\, b\, c\, d]$ into a row-reduced echelon form R, and then choosing a vector e that is *not* in the span of the columns of R. The instructor reiterated Eric's "solution" while computing R with MATLAB,[6] but gave no indication or gesture about its falsity. None of the students rejected Eric's solution or raised any question about its validity. At this point he asked the class to discuss Eric's solution in their working groups. In the class discussion that followed, Jon, one of the students in the class, said that he was not sure that Eric's solution was correct because he did not think that row operations preserve the span of the columns. He then went on and suggested checking his hypothesis on a specific example. The example he suggested (a 3×2 matrix, which he seemed to choose at random) showed that indeed row operations do not necessarily preserve the column space. Following this, another student, Mathew, indicated that he had suspected that Eric's solution was incorrect but he accepted it because the instructor did not reject it. At this point, the instructor asked the class, How many students doubted the correctness of Eric's solution prior to Jon's example? Over 50% of the class raised their hand. And when he asked the class, "Why didn't you express your doubts?" the answer, which all seemed to agree with, was: "Because you (the instructor) appeared to agree with Eric's solution."

The authoritarian proof scheme is both pervasive and difficult to relinquish. Students' most common questions during the first part of the teaching experiments have been about "how" rather than "why." Further, it is rather difficult to establish a new "didactical contract" (à la Brousseau, 1986) in which the concern is with both kinds of questions and in which they, the students, must participate in the search for answers to such questions. It is this latter demand that is so difficult for students to accept, for they view the instructor as the sole source for answers and believe that it is the instructor's responsibility to tell them the needed knowledge. To extinguish this belief, students' questions, both how and why, are seldom answered directly in our teaching experiments; rather, they are brought to reason together among themselves and with the instructor about the questions and the search for their answers. For further discussion, see the papers by Arcavi et al., Santos, and Schoenfeld (this volume).

3.1.3 Symbolic Proof Scheme.

Thinking of symbols as though they possess a life of their own without reference to their possible functional or quantitative reference, we call symbolic reasoning. As discussed below, symbolic reasoning can either be superficial and mathematically vacuous, or a very powerful technique. A symbolic proof scheme is a scheme by which mathematical observations are proven by means of such symbolic reasoning. Consider the following episode:

Episode 9. In his attempt to prove that any homogeneous system of linear equations $AX = 0$ is consistent, Hugh, a student in LA2, said:

[6] A computer display system was present in all the linear algebra teaching experiments. The computer package MATLAB was used during the class sessions to perform tedious computations (e.g., row reducing a matrix) and, in particular, to make and test conjectures, empirically verify proven results, etc.

Take, for example, $x_1A_1+\cdots+x_nA_n = 0$. If we want to find x_1 and we have some values $x_2, \ldots, x_n$ and $A_1, \ldots, A_n$, by moving the known values to the other side of the equation, we can solve x_1. This is the same for $x_2, \ldots, x_n$. Hence, this demonstrates that a homogeneous system has solution. [In giving this explanation, Hugh wrote $x_1 = \frac{x_2A_2 + \cdots + x_nA_n}{A_1}$.]

When he was asked what dividing by A_1 means, he replied:

> It is just this [pointing to the numerator of his fractional expression] is over this [pointing to the numerator of his fractional expression]. It is just like one over x.

The main characteristic of the symbolic reasoning is the behavior of approaching the solution of a problem without first comprehending its meaning, that is, without building a coherent image of the problem situation. Consider the following episode:

Episode 10. In teaching experiment NT, the instructor observed students' immediate actions when they were given a problem to solve individually (either in a class setting or an interview). Many of the students read the problem only once and haphazardly began manipulating the symbolic expressions involved in the problem, with little or no time spent on comprehending the problem statement. Hence, many approached the solution without knowing the meaning of some of the terms used in the problem statement, and many others were unable to articulate the exact task they were to accomplish.

Symbolic reasoning is a habit of mind students acquire during their school years—from elementary school to secondary and post secondary school—a habit that is very persistent and extremely difficult to relinquish. For an analysis of the consequences of symbolic reasoning on students' mathematical development, particularly on the development of multiplicative reasoning and algebraic reasoning see Harel (1995, in press).

Against this devastating reasoning, we shall point to a different, essential practice of symbolic reasoning. The definition of symbolic reasoning we gave earlier may have evoked with the reader a different image from the one revealed in Hugh's responses. For, relative to the reader's practice of mathematics, it is not uncommon that symbols are treated as if they possess a life of their own, and, accordingly manipulated without (necessarily) examining their meaning. Historically, this practice of symbolic reasoning played a significant role in the development of mathematics. For example, during the nineteenth century an enormous amount of work was done in differential and difference calculus using a technique called the "operational method," a method whose results are obtained by symbol manipulations without account of their possible meanings, and in many cases in violation of well-established mathematical rules. (See, for example, how the Euler-MacLaurin summation formula for approximating integrals by sums was derived [Friedman, 1991, pp. 176-178].) It was only with the aid of functional analysis, which emerged early in the twentieth century, that mathematicians were able to justify many of the techniques of the operational method.

3.2 Empirical Proof Schemes.

In an empirical proof scheme, conjectures are validated, impugned, or subverted by appeals to physical facts or sensory experiences. We distinguish between two kinds of this scheme: The *inductive* empirical proof scheme and the *perceptual* empirical proof scheme.

3.2.1 Inductive Proof Scheme.

When students ascertain for themselves and persuade others about the truth of a conjecture by *quantitatively evaluating*[7] their conjecture in *one or more* specific cases, they are said to possess an inductive proof scheme. Every teacher has likely observed the dominance of this scheme among students, and research into the concept of proof corroborates this observation (see, for example, Chazan, 1993). Martin and Harel (1989) studied how prospective elementary school teachers judged whether particular arguments were mathematical proofs. The arguments were either inductive (based on specific instances) or deductive (assertions via general statements). More than 80% of the 101 prospective teachers considered inductive arguments to be mathematical proofs. Their conviction about the truth of the conjecture became particularly strong when they observed a pattern suggesting that one can generate as many examples as wanted in support of the conjecture. In Goetting's study (1995), almost 40% of her advanced undergraduates used examples as a basis for judging the truth of a divisibility question (p. 43).

Since people's evaluation of hypotheses in everyday life is probabilistic in nature (Anderson, 1985), the use of inductive evidence is only natural. Moreover, the initial application of inductive reasoning in mathematical activities is in most cases essential. So the concern is not that college students think inductively; rather, the concern is that their proof schemes do not develop beyond the empirical proof schemes. This retarded development should not be a surprise when mathematics instruction in both the elementary and secondary levels is dominantly inductive in its best forms, and authoritarian, ritual, or symbolic in its worst forms.

Despite its dominance, the inductive proof scheme phenomenon is not entirely understood. Evidence exists to indicate that students—at least adult students—are aware of the limitations of an inductive scheme. Chazan (1993), for example, has shown that high school students who thought examples were sufficient proof "understood some of the limitations inherent in the use of examples, and had strategies for minimizing these limitations" (p. 370). Yerushalmy (1986), also with high school geometry students and in a computer environment inviting inductive work, noted that "throughout the (year-long) course the students'

[7]E.g., direct measurements of quantities, numerical computations, substitutions of specific numbers in algebraic expressions, etc.

appreciation of data as a source of ideas grew, while their appreciation of data as an argument declined" (p. vii). Our teaching experiments corroborate Chazan's and Yerushalmy's observations but reveal further complexity.

3.2.2 Proof By Example and Counterexample.

The issue of "proof by examples" was discussed with the students in our teaching experiments. In this discussion the instructor showed his students examples of mathematical statements that are true for numerous cases but untrue for *all* cases.[8] By doing this, he intended to convince the students that inductive verifications are insufficient to validate a conjecture. Although in these discussions the students seemed to understand the limitations inherent in the inductive method, their subsequent behavior was not consistent with this impression. Specifically, we observed the following behaviors:

a) students continued to prove mathematical statements by examples;

b) students did not protest when they were presented with an inductive proof;

c) students' inductive proofs mostly consisted of *one* example, rather than a multitude of examples.

When the instructor confronted his students with their contradictory behavior by directly asking them why they used proof by examples when they agreed that counterexamples may be found, they typically said something like the following:

d) The inductive method had always been used by their instructors in previous mathematical classes—typically they mentioned calculus classes; and

e) even if a counterexample to the statement is found, the statement stands, because the counterexample is just an exception.

In addition to these observations, here are some other related observations:

f) Students seldom used proof by counterexample and

g) they did not seem to be convinced by it;

h) nor were they convinced by proof by contradiction; further,

i) students confused the admissibility of proof by counterexample with the inadmissibility of proof by example(s); and

j) the use of inductive proofs diminished as students developed advanced proof schemes.

To explain these behaviors, we suggest that there are four combined cognitive forces influencing students' thinking with respect to inductive proofs. They interact in powerful ways, as discussed below.

---

[8]E.g., the conjecture, $\sqrt{1141y^2+1}$ is an integer, is false for $1 \le y \le 10^{25}$. The first value for which the statement is true is 30, 693, 385, 322, 765, 657, 197, 397, 208 (see Davis, 1981).

The first cognitive force is the psychologically natural tendency to evaluate conjectures probabilistically and hence the readiness with which this mode of evaluation is accepted as a mode of mathematical reasoning. This force accounts for students' use of specific examples to verify a general statement. But, against one's expectation, very often students use just *one* example, rather than a multitude of examples to prove their conjectures (observation *c*). As we have looked at this phenomenon more closely, we offer this as the reason: The students believed that since the single example was chosen randomly and it conformed to the general statement, the statement must be true. The single example is taken as representative of all cases. Even when a counterexample to the statement is found, the statement still stands in the students' eyes, because the counterexample is just one exception to the general rule (observation *e*). This argument is consistent with the fact that the first author's students seldom used proof by counterexample (observation *f*) and they did not seem to be convinced by it (observation *g*). The second author's students, on the other hand, often sought counterexamples first, as did Goetting's subjects (1995), although many of her interviewees were not certain whether a counterexample gave a proof. Whether this seeming difference from the first author's students is a fact or an happenstance of either the particular interviewees or the curricula at the different universities, we do not know. This, together with Balacheff's (1991) finding that younger students (junior-high school students in France) do react to counterexamples in various ways, requires a further look at college students' conception of proof by counterexample.

The second cognitive force fostering an inductive proof scheme is the effect of the authoritarian proof scheme discussed in the previous section. Combined with the first force, it explains why students do not protest when they are presented with an inductive proof (observation *b*), since other mathematics teachers, both in high school and college, have used inductive proofs regularly (observation *d*), as the students have explicitly indicated.

The third cognitive force is students' difficulty understanding proof by contradiction (Tall, 1979; Thompson, 1996), where they may believe the proof assumes what is to be proved (see Episode 22 below). This belief accounts for observation *h*. In a deeper sense this difficulty with proof by contradiction may also account for students continuing to prove mathematical statements by examples (observation *a*) despite instructor warnings, with which they seemed to concur. The instructor's attempt to convince the students of the invalidity of proof by example is based on proof by contradiction, which is a method of proof the students have never accepted.

The fourth cognitive force is in fact an absence of one, namely students' lack of advanced proof schemes other than the external conviction proof schemes and the empirical proof schemes. Students justify their conjectures by these proof schemes simply by default—these are the only schemes they possess. One of the most positive results of our teaching experiments is that the use of the empirical and external conviction proof schemes diminished as students developed more advanced proof schemes (observation j).

As students begin to acquire these alternative proof schemes, they begin to appreciate proof by contradiction and proof by counterexample. But the impact of the second force—the authoritarian proof scheme—forces them to accept these modes of proof before they have completely internalized them. The result is a confusion between the admissibility of proof by counterexample with the inadmissibility of proof by example(s) (observation i), as the following episode demonstrates.

Episode 11. Andy, a student in the teaching experiment LA1, protested the grade he received on one of the test problems, where he proved a certain statement inductively, by saying,

> I am confused. Sometimes you can prove things by examples and sometimes you can't.

When asked to indicate a case where a proof by an example was accepted, he mentioned a case where the instructor demonstrated the falsity of a certain statement by pointing to a single counterexample of the statement. The statement he referred to was: If two matrices are different from zero, then their product is different from zero. This statement was refuted by the counterexample:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} -2 & 1 \\ 1 & -0.5 \end{bmatrix}, \qquad AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

3.2.3 Perceptual Proof Scheme.

Perceptual observations are made by means of rudimentary mental images—images that consist of perceptions and a coordination of perceptions, but lack the ability to transform or to anticipate the results of a transformation. The important characteristic of rudimentary mental images is that they *ignore transformations on objects or are incapable of anticipating results of transformations completely or accurately*. "Such images constitute an imitation of actions that can be carried out in thought (e.g., rotations of objects) . . . [but they] cannot be adequately visualized all the way to [their] ultimate conclusion before [they have] actually been performed" (Piaget & Inhelder, 1967, p. 295). A full mental image (as opposed to a rudimentary mental image), on the other hand, is "a pictorial anticipation of an action not yet performed, a reaching forward from what is presently perceived to what may be, but is not perceived" (Piaget & Inhelder, 1967, p. 294).

One might look at the triangle ABC in Figure 2 and perceptually observe two equalities: that sides AB and AC are of the same length and that angles ACB and ABC are of the same size. If this observation is merely perceptual, the observer will not see a causality relationship between these two equalities (i.e., that one equality causes the other). On the other hand, in an observation that involves transformations, the same figure can be viewed as one out of many possible outcomes of transformations on the figure's parts. For example, the isosceles triangle ABC in Figure 2 can be seen as a special state—in which the angles opposite AC and AB are congruent—among all the possible states resulting from varying the point A on the half circle with center C and radius AC—in which the angles opposite AC and AB are not congruent. The following episodes demonstrate students' perceptual proof schemes.

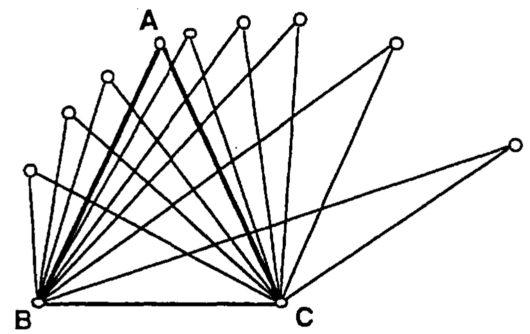

FIGURE 2

Episode 12. In the teaching experiment CG, while discussing the homework problem, "Prove that the midpoints of any isosceles trapezoid form a rhombus," the instructor drew a figure on the board similar to Figure 3.

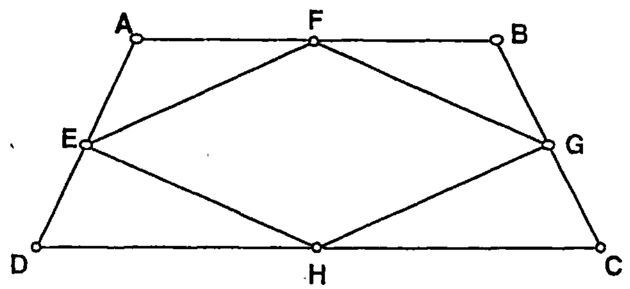

FIGURE 3

Melissa presented her proof of this proposition, saying that she first proved that FH is congruent to EG by showing that ΔFEH is congruent to ΔHEG. During her presentation, she kept looking at the figure she drew in her notebook, matching the letters in her figure to the letters in the figure the instructor drew on the board. Melissa's attempt to prove that the midsegments FH and EG are congruent was driven by her perceptual observation of the figure she drew, in which the midsegments indeed looked congruent. She was concerned with this specific case only, unaware of the variability of the trapezoid's shape.

Interestingly, the class discussion that followed did not address Melissa's proof per se; rather, it centered on the plausibility of such a "proof." The students argued that "there is no chance that Melissa can prove that the midsegments are congruent," for the trapezoid can vary, and among its variations it can be narrow and long, in which case the midsegments FH and EG would not be congruent. This is an example of how transformations are applied for evaluation and control purposes (Schoenfeld, 1985).

There is no doubt that Melissa was capable of varying the trapezoid (e.g., imagining how it can be stretched along the midsegment EG or FH to get different trapezoid shapes). Indeed, Melissa did understand the argument made by her classmates and fully agreed to it. The question, however, is whether Melissa was capable of initiating such transformational reasoning[9] on her own. An indication that Melissa did not possess such a cognitive ability can be found in her response to a similar problem on a written test that she took five days later. The problem was:

> A, B, C, D are the midpoints of the four sides of a parallelogram. The quadrilateral $ABCD$ is ALWAYS a: (a) parallelogram; (b) rectangle; (c) rhombus; (d) square; (e) none of the above. Justify your answer.

Melissa drew a parallelogram $TRPS$ which looked like a square (Figure 4) and argued:

> "by Pythagorean theorem $AB \cong BC$ in ΔARB and ΔCPB. By the same reason, $AD \cong DC$ in $\Delta TAD \cong \Delta DSC$"

and concluded that $ABCD$ must be a square.

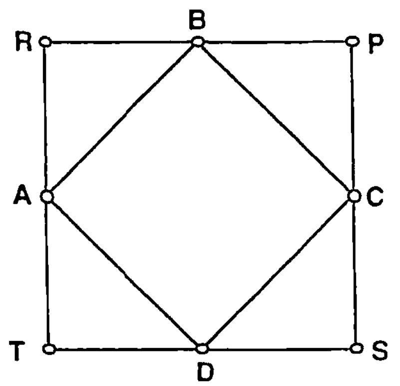

FIGURE 4

[9]Martin Simon (cf. 1996) suggested we use the term "transformational reasoning" instead of the term "transformational observation" that we initially used.

Again, Melissa's reasoning was based on the particular figure she drew, not the possible variation of the figure.

Students apply perceptual proof schemes in algebraic situations as well:

Episode 13. Don was asked to determine whether the following vectors are linearly dependent or independent:

$$u = \begin{bmatrix} \sqrt{3} \\ \sqrt{5} \end{bmatrix}, \qquad v = \begin{bmatrix} -\frac{7}{10} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad x = \begin{bmatrix} \sqrt{61} \\ 103 \end{bmatrix}.$$

He responded:

> I conclude u, v, and x are linearly independent by observation. My observation being one vector can not be written as a linear combination of the others.

Piaget's observations of different conservation tasks (Piaget, 1983), Steffe's research on children's counting schemes (Steffe, Cobb, & von Glasersfeld, 1988), and even Kitcher's philosophical outlook (1983) suggest that, epistemologically, the perceptual proof scheme is children's first source of internal conviction. This scheme continues to play an important role throughout students' mathematical education. In learning fractions, for example, children's early judgments of the order relation between fractions are merely perceptual (e.g., 2/3 is smaller than 3/4 because it looks so). The above two episodes demonstrate that even in advanced topics such as geometry and linear algebra, the perceptual proof scheme remains for many students a source of conviction and persuasion.

3.3 Analytical Proof Schemes.

Simply stated, an analytical proof scheme is one that validates conjectures by means of logical deductions. By this, however, we mean much more than what it is commonly referred to as the "method of mathematical demonstration"—a procedure involving a sequence of statements deduced progressively by certain logical rules from a set of statements accepted without proofs (i.e., a set of axioms). Our characterization of the analytical proof schemes category is best described by its subcategories: the transformational proof scheme and the axiomatic proof scheme.

3.3.1 The Transformational Proof Schemes.

Transformational observations involve operations on objects and anticipations of the operations' results. The operations are goal oriented. They may be carried out for the purpose of leaving certain relationships unchanged, but when a change occurs, the observer intends to anticipate it and, accordingly, intends to apply operations to compensate for the change. We call them "transformational" because they involve *transformations* of images—perhaps expressed in verbal or written statements—by means of deduction. By "images" in this context we mean the ones which Thompson (1994), based on Piaget (1967), has characterized as images that "support thought experiments and support reasoning by way of quantitative relationships" (p. 230). The following two episodes demonstrate this characterization of the transformational proof scheme.

Episode 14. Amy, one of the students in teaching experiment CG, demonstrated to the whole class how she imagines the theorem, "The sum of the measures of the interior angles in a triangle is 180°." Amy said something to the effect that she imagines the two sides AB and AC of a triangle ΔABC being rotated in opposite directions through the vertices B and C, respectively, until their angles with the segment BC are 90° (Figure 5a, b). This action transforms the triangle ABC into the figure $A'BCA''$, where $A'B$ and $A''C$ are perpendicular to the segment BC. To recreate the original triangle, the segments $A'B$ and $A''C$ are tilted toward each other until the points A' and A'' merge back into the point (Figure 5c). Amy indicated that in doing so she "lost two pieces" from the 90° angles B and C (i.e., angles $A'BA$ and $A''CA$) but at the same time "gained these pieces back" in creating the angle A. This can be better seen if we draw AO perpendicular to BC: angles $A'BA$ and $A''CA$ are congruent to angles BAO and AOC, respectively (Figure 5d).

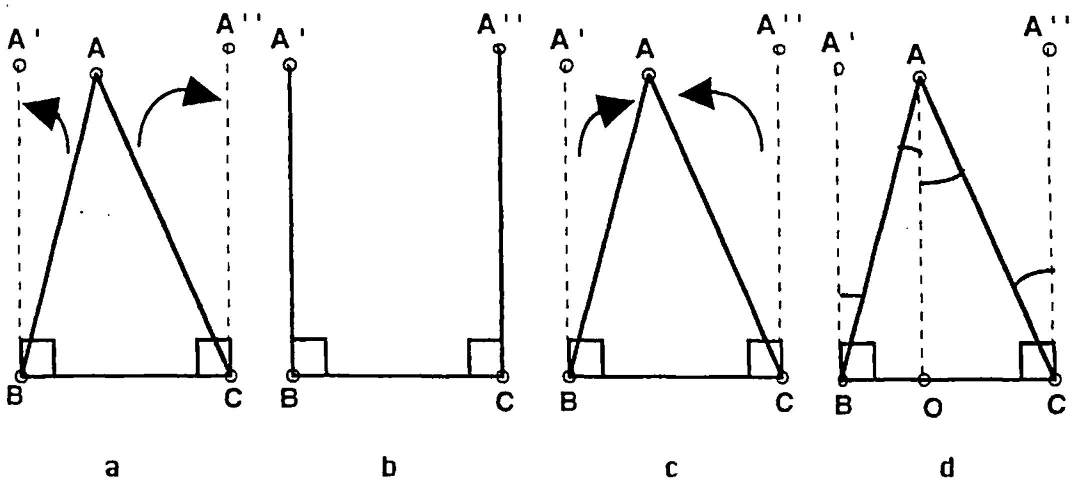

FIGURE 5

The focus of this episode is on Amy's way of thinking about the relationship stated in the theorem. Amy views a triangle as a dynamic entity; it is a product of her own imaginative construction, not of a passive perception. Her operations were *goal oriented* and intended the *generality* aspect of the conjecture (rather than a particular figure as in Episode 12). She *transformed* the triangle and was fully able to *anticipate* the results of the transformations, namely, that the change in the 90° angles B and C caused by the transformations is compensated for by the creation of the angle A. All this leads to her *deduction* that the sum of the measures of the angles of the triangle is 180°.

Episode 15. In teaching experiment CG, while discussing the relationship between an inscribed angle AHB and the angle between the chord AB and the tangent line AN (Figure 6a), Jeff argued that the two angles are congruent, because when H moves along the arc ATH (drawing a figure like Figure 6b on the board), a sequence of inscribed angles AH_1B, AH_2B, $AH_3B, \ldots$, is formed. All these angles, including all angles AXB where X is on the arc ATH, are congruent by a known theorem. The point A is one of these X's. When X coincides with A, the angle AXB becomes BAN. Therefore, the angles BAN and BHA are congruent.

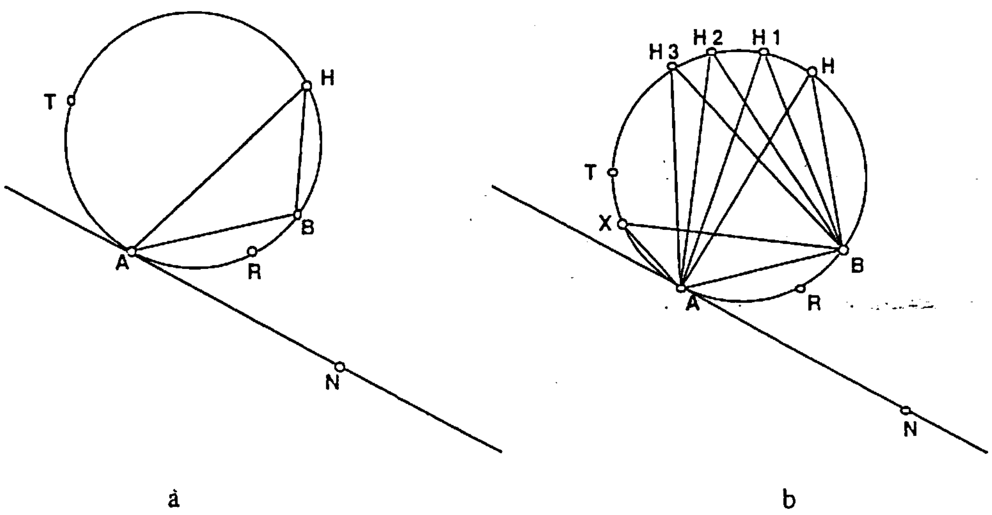

FIGURE 6

In this episode, we see how Jeff *varied* the position of point H for the *purpose* of evaluating the relationship between and the inscribed angles that result, and how he was able to *anticipate* the outcome of the variation: *deducing* the size of the angle. This ability is essential in viewing geometric objects as loci of points (e.g., a circle consists of all points whose distance from a given point is fixed; an angle bisector consists of all points whose distances from the angle rays are equal) rather than static figures.

Students apply perceptual proof schemes in algebraic situations as well:

Episode 16. Ed (in Case Study EC) explained his solution to the problem: Prove that for $x \geq 0$, $\log(x+1) \leq x$. He first converted this inequality into the form $x + 1 \leq e^x$, then he said:

> Both functions [$x + 1$ and e^x] are increasing but e^x goes faster. At zero they are equal, so e^x must be greater.

When Ed was asked to explain the mathematical meaning of his statement, "e^x goes faster [than $x + 1$]," he did so in terms of the concept of derivative.

Episode 17. When Ed was a high school student he took a college linear algebra course. The instructor asked him: If an $n \times m$ matrix transformation A is one-to-one, is A "tall" ($n > m$), "short" ($n < m$), or square ($n = m$)? Ed's answer was this:

> It can't be short because if you look at the system $AX = 0$, you would have free variables, which means the null-space of A contains a non-zero vector. This would destroy the definition of one-to-one, because you would have more than one vector going to zero.

In a follow-up discussion, shortly after this response, he added: "If the matrix is short, then it [the transformation] would run out of 'space'." The instructor learned that this latter metaphoric statement represented Ed's image of one-to-one matrix transformations. That is, when A is one-to-one, different vectors in R^m would correspond to different vectors in R^n. Since R^m is "larger" than R^n, there wouldn't be enough vectors, so to speak, in R^n to correspond to vectors in R^m. It must be mentioned that Ed understood the notion of R^m being "larger" than R^n in terms of dimensions, not cardinality. He brought up the concept of dimension when the instructor presented the following "counterexample" to his statement:

> The function $f(x) = \arctan(x)$ is one-to-one [a fact Ed agreed to by saying "Yes, it is an increasing function, so it must be one-to-one"] and maps a "larger" interval, $(-\infty, \infty)$, onto a "small" interval, $(-\pi/2, \pi/2)$, but it doesn't run out of space.

Ed even constructed his own "counterexample" that involves a linear transformation. He pointed out that one can map a "large" interval into a "small" interval by a linear function, by simply taking a straight line that goes through zero and with a slope less than 1.

3.3.1.1 Cognitive Levels of the Transformational Proof Scheme. To recapitulate, the transformational proof scheme is characterized by (a) consideration of the generality aspects of the conjecture, (b) application of mental operations that are goal oriented and anticipatory, and (c) transformations of images as part of a deduction process. These characteristics do not provide a complete picture of the transformational proof scheme, for they neither capture its inherent limitations, nor do they express its different cognitive levels. In the next two sections, we discuss two levels of the scheme: the internalized proof scheme and the interiorized proof scheme (cf. Steffe, Cobb, & Glasersfeld, 1988). Following this, we address the restrictions presumed by students on the transformation proof scheme.

3.3.1.1.1 Internalized Proof Scheme. An internalized proof scheme is a transformational proof scheme that has been encapsulated into a proof heuristic—a method (of proof) that renders conjectures into facts. For example, to prove two segments in a given figure are congruent, students commonly look for two congruent triangles that respectively include the two segments. This proof heuristic is abstracted by the students from repeated application of an approach they have often found to be successful. Here is another example of the internalized proof scheme pertaining to elementary mathematics. Students who were taught to think of fractions transformationally, not just symbolically, abstract different methods of comparing the order relation between fractions. For example, in many cases, when asked to compare the order relation between two given fractions, they respond by comparing the complements of the fractions to the whole (e.g. $7/8 < 8/9$ because $7/8$ is $1/8$ away from 1 whereas $8/9$ is only $1/9$ away from 1). An indication that the "comparing-to-the-whole" approach is an internalized transformational proof scheme rather than a ready-made formula is that the students apply it selectively, that is, only when they recognize its efficacy (for example, they may not apply it in comparing $2/5$ and $2/7$ or $5/11$ and $6/11$).

In teaching experiments LA1 and LA2 most of the students internalized the matrix row reduction process as a method of proof. That is to say, row reduction became for these students a spontaneous conceptual tool for conjecturing and evaluating conjectures. This way of thinking is demonstrated in the following episode.

Episode 18. Lee, a student in LA2, was asked,

Let $A_1, A_2, \ldots, A_m$ be different vectors in R^n. If $m > n$, are these vectors linearly independent?

Lee began by writing

$$\# \text{ equations} < \# \text{ unknowns},$$

then she wrote:

> The vectors are dependent because there are more unknowns than equations. So you will have some free variables. These free variables will be linear combinations of the independent vectors so the vectors are dependent.

Implicit in Lee's solution is the translation of the problem at hand into a system of linear equations. This was evident in her solution to the problem that immediately preceded the previous one:

Determine whether the following vectors are linearly dependent or independent:

$$A_1 = \begin{bmatrix} 0.2 \\ 0.4 \\ 1.6 \\ 4 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.9 \\ 3 \\ 12 \\ 0.8 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0.03 \\ 535 \\ 97 \\ 1 \end{bmatrix},$$

$$A_4 = \begin{bmatrix} 7 \\ 4 \\ 68 \\ 175 \end{bmatrix}, \quad A_5 = \begin{bmatrix} 68 \\ 890 \\ 4 \\ 2 \end{bmatrix}, \quad A_6 = \begin{bmatrix} 69 \\ 20 \\ 9 \\ 0 \end{bmatrix}.$$

Lee wrote:

> If we consider the dependency equation $xA_1 + yA_2 + zA_3 + wA_4 + vA_5 + uA_6 = 0$. If we wrote this out as a system of equations we would get 4 equations with 6 unknowns. 4 equations is not enough to [uniquely] determine 6 unknowns since the system is homogeneous and hence consistent, we know there are an infinite number of solution because we will have 2 free variables. Hence the vectors are linearly dependent.

Also implicit in Lee's solution is the image she had of the reduced row echelon form of a matrix and its implications to solution existence. Her next solution to another problem gives some hints of the structure of this image: Lee was asked to solve a specific non-homogeneous linear system $AX = b$ of four equations and five unknowns, then was asked to look at the relation between $rank(A)$, the number of unknowns, and the number of spanning vectors in the solution set (that is, the basis of the solution space of the corresponding homogeneous system), and to prove that that this relation holds for any system that has a solution. Lee responded:

> Want to show that: number of unknowns— $rank(A) = \#$ spanning vectors. The number of unknowns will determine how many columns are in the matrix. The $rank(A)$ (reduced row echelon form of A) corresponds to the number of non-zero rows which in reduced form correspond to the pivot variables [paraphrasing: $rank(A)$ corresponds to the number of non-zero rows in $rref(A)$, which, in turn, corresponds to the number of pivot columns]. So if we subtract the pivot variables from the number of unknowns the resulting columns are the free variables which form our spanning vectors [paraphrasing: If we subtract the number of pivot columns from the total number of columns, we get the number of the non-pivot columns, or the number of free variables, which is the number of spanning vectors in solution form of the equation]. Thus it follows: number of unknowns— $rank(A) = \#$ spanning vectors.

The development of this way of reasoning was a result of our effort to eradicate students' external conviction proof scheme by helping them develop alternative, transformational proof schemes. We engaged students in numerous analyses of linear algebra questions in terms of systems of linear equations and, in turn, in terms of the meaning of row operations on the system's equations. As a result, students built images of the structure of the rref (reduced row echelon form) of a matrix and the meaning and implications of row reduction in questions of existence and uniqueness of solutions of linear systems and in questions of linear independence and span.

Further, some students have even discovered important theorems by analyzing the situation in terms of row operations, the most notable of which were the Rank Theorem (the row rank of a matrix equals its column rank) and the Fundamental Theorem of Linear Algebra ("Nullity($A_{m\times n}$) + Rank($A_{m\times n}$) = n"). The students' thinking was along these lines: Let A be a matrix and let r and c be its row-rank and column-rank, respectively. By definition, r and c are the maximum number of linearly independent rows of and the maximum number of linearly independent columns of A, respectively. Both r and c are invariant under row operations. Thus r is the number of non-zero rows in $rref(A)$ and c is the number of pivot columns in $rref(A)$. Based on their image of what the structure $rref(A)$ must be, they concluded that $r = c$, for each non-zero row in $rref(A)$ corresponds to exactly one pivot column in $rref(A)$.

A particularly important instance of the internalized proof scheme lies in the transformational use of symbols. To prove or refute a certain conjecture, the conjecture may be represented algebraically and symbol manipulations on the resulted expressions are performed, with the intention to derive relevant information that deepens one's understanding of the conjecture, and that can potentially lead to its proof or refutation. In such an activity, the individual does not necessarily form specific images for some or all of the algebraic expressions and relations that result in the process. It is only in critical stages in this process—viewed as such by the individual—that he or she intends to form such images.

Students should develop this scheme during their early experiences with algebra word-problems and especially through activities in analytic geometry. Despite this, we found that the scheme was largely absent with among students, at least during the first part of the teaching experiments. It was not that these students were unable to understand a symbolic translation of a verbal conjecture when it was presented to them, but that such an approach was seldom initiated by them. For example, in discussing the conjecture $Null(A)$ is orthogonal to $Rowspace(A)$, none of the students in LA3 suggested analyzing the conjecture by first representing it symbolically (e.g., $A_{m\times n}x = 0 \Leftrightarrow \sum_{i=1}^{n} x_i A_{j,i} = 0, \quad j = 1, \ldots, m$). This observation is worth mentioning because mathematics instructors usually take this scheme of transforming verbal statements into symbolic statements for granted, perhaps believing students' mathematical education in high school guarantees its acquisition.

3.3.1.1.2 Interiorized Proof Scheme: The Case of Mathematical Induction. An interiorized proof scheme is an internalized proof scheme that has been reflected upon by the person possessing it so that he or she becomes aware of it. A person's awareness of the proof scheme is usually observed when the person describes it to others, compares it to other proof schemes, specifies when it can or cannot be used, etc. As an example of this scheme, we illustrate with the case of mathematical induction. By definition, the interiorization process cannot occur unless the internalization process has taken place. In what follows we will describe briefly how students internalized and then interiorized the principle of mathematical induction. This process is divided into three stages: The construction stage, the internalization stage, and the interiorization stage.

3.3.1.1.2.1 The Construction Stage. We investigated the concept of mathematical induction in the teaching experiment NT. As a result of previous pilot experiments we hypothesized that students' difficulties with the principle of mathematical induction can largely be attributed to two factors: (a) the formal expression of this principle is hastily introduced to students, and (b) the kinds of problems typically introduced to students in their first experience with mathematical induction are cognitively inadequate in the sense we will explain below. Accordingly, students in the NT teaching experiment were engaged for a relatively long period of time—before the principle of mathematical induction was explicitly mentioned—in working on problems typified by:

1. Find an upper bound to the sequence $\sqrt{2}$, $\sqrt{2+\sqrt{2}}$, $\sqrt{2+\sqrt{2+\sqrt{2}}}$, ...
3. Three pegs are stuck in a board. On one of these pegs is a pile of disks graduated in size, the smallest being on top. The object of this puzzle is to transfer the pile to one of the other two pegs by moving the disks one at a time from one peg to another in such a way that a disk is never placed on top of smaller disk. Prove that this can be done for 31 disks in 2,147,483,647 moves. Generalize the problem and your solution accordingly.

This kind of problem, which we call construction-stage problems, led students to focus on the relationship between consecutive items in the sequence. Thus, for example, in working on Problem 1 in small groups, students first conjectured that 2 is an upper bound—typically by finding the calculator values of several items in the sequence. Then they explained that the third item is less than 2 because it is the square root of a number that is smaller than 4 (this number being the sum of 2 and a number that is smaller than 2). They repeated this argument several times by applying it to the next few items in the sequence, and concluded that all the items in the sequence must be less than 2 because the same relationship exists between any two consecutive items in the sequence. The students' conviction through the latter argument was in the ascertainment level rather than the persuasion level, so the instructor focused their attention on proving—persuading him, that is—that indeed the relationship they had observed holds for any consecutive items of the sequence. This stage of development may be viewed as students' construction of a transformational proof scheme for mathematical induction.[10]

[10]For a different view, see Dubinsky, 1986, 1989.

3.3.1.1.2.2 The Internalized Stage. After repeated applications of this way of reasoning in solving problems such as Problems 1 and 2, and of the formulation and proving of the general relationship between consecutive statements, and the establishment of the first statement in the sequence, we introduced a different kind of problem typified by this one:

3. Prove that for any positive integer $n \geq 1$,

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{n \cdot (n+1)} = \frac{n}{n+1}.$$

At first, the students did not appear to see any relation between this kind of problem, which we call internalized-stage problems, and construction-stage problems; their main attempts were to calculate a closed form for the summation on the left-hand side of the equality. It turned out that the students did not interpret the problem in terms of a sequence of statements (that is, as a proposition-valued function), so we worked on this way of interpretation. Problem 3, for example, was reformulated into:

3′. Find the general pattern in

$$\begin{aligned} \frac{1}{1 \cdot 2} &= \frac{1}{1+1} \\ \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} &= \frac{2}{2+1} \\ \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} &= \frac{3}{3+1} \\ &\vdots \end{aligned}$$

At this stage, many students realized that for this new breed of problems they could use reasoning similar to that used for construction-stage problems. Our interpretation is that this is a stage where the related transformational proof scheme built in the construction stage has been internalized and become a method of proof for these students, but they were not aware of it as a method of proof. This took place in a subsequent stage.

3.3.1.1.2.3 The Interiorized Stage. At this stage we introduced interiorized-stage problems, which are typified by these:

4. Prove that for any positive integer n,

$$\log(a_1 a_2 \cdots a_n) = \log a_1 + \log a_2 + \cdots + \log a_n$$

5. Prove that for any positive integer $n \geq 4$, $2^n \leq n!$

Before we continue, a brief analysis of these three stages of problems is in order. Textbooks typically present their problems on mathematical induction in an order that is almost opposite to ours: They usually present internalized-stage problems and interiorized-stage problems first, leaving construction-stage problems to the end. Presumably, their order is based on the view that problems of the latter kind are harder than the former kind. In our teaching experiments, we were concerned about which problems fit what stage of students' mathematical development rather than about the (subjective) order of difficulty of the problems. Based on previous experience, we found that interiorized-stage problems were inadequate to introduce the principle of mathematical induction. Students felt no need to search for a relationship between consecutive items of the given sequence because they could easily find "alternative" ways of solving them. Problem 5, for example, was viewed by the students as completely trivial (i.e., for any integer $n \geq 4$, all but one factor in $n!$ is greater than or equal to the corresponding factor in 2^n); similarly, they solved Problem 4 by a repeated application of the formula $\log xy = \log x + \log y$, and saw no need to look for an alternative proof, certainly not a proof by mathematical induction.

It was this simplistic aspect of interiorized-stage problems that posed a challenge to some of the students in the NT teaching experiment. They sensed the underlying structure of the three sets of problems and began to see how the same method of proof could be applied to all. They had to deny their natural inclination to solve these problems with a simple method and look for a different way to apply the method they used in the construction and internalized stages. This forced them to be explicit about the method of proof they had internalized in the second stage and to examine its applicability to the new set of problems. Indeed, we found that most of the students who had internalized the principle of induction in the internalized stage were able to formulate it explicitly and apply it to solve interiorized-stage problems. Thus, these students had interiorized the principle of mathematical induction as a method of proof.

It should be clearly stated that the interiorization stage does not (and it did not in the case of our students) imply the understanding that mathematical induction is essential in proving proposition-valued statements whose domain is the positive integers. Indeed those students who successfully proved interiorized-stage problems by mathematical induction continued to believe that mathematical induction is not necessary in the case of problems such as Problems 4 and 5, for "easier" proofs do exist. We did not expect our students to reach this level of understanding based on the instructional program of the teaching experiment. We believe that a full appreciation of mathematical induction—specifically this level of understanding—can only be reached when the *axiomatic* proof scheme (to be defined below) is built.

3.3.1.2 Restrictive Characteristics of the Transformational Proof Scheme. We found that many of our students who were able to think transformationally presumed certain restrictions on either the context of the conjecture, the generality of the justification, or the mode of the justification. Accordingly, a transformational proof scheme that involves any one of these restrictions is called a *restrictive* proof scheme, and depending on the kind of the restriction, it may be called a *contextual*, *generic*, or *constructive* proof scheme, respectively.

3.3.1.2.1 Contextual Proof Scheme. Researchers (e.g., Lave & Wenger, 1991) have addressed the importance of the broad context in which a particular activity takes place. As can be seen from the discussion below, our interest is more narrow than this broad level of contextualization. We are interested in the specific interpretations students give to mathematical assertions in the process of their justification in collegiate mathematics settings.

In a contextual proof scheme conjectures are interpreted, and therefore proved, in terms of a specific context. For example, a student might interpret and prove the *general* statement "$n+1$ vectors in an n-dimensional vector space are linearly dependent" in the *specific* context of R^n. One reason for this is that the only linear algebra "world" the student has ever "lived" in is R^n, and so he or she has not yet abstracted the concept of "linear independence" beyond this specific context.

Another example of the contextual proof scheme comes from students' inability to deal with any geometric structure but the one that deals with their spatial imageries. One of the questions we asked in the CG teaching experiment was: Can students (in this teaching experiment) think in terms of abstract structures, namely, that the axioms in geometry require no specific interpretation? In particular, can they consider their own imaginative space as a specific system which may or may not satisfy the structure at hand?

Students' responses in an extensive set of interviews conducted at the end of the instructional unit that dealt with axiom systems revealed three categories of proof schemes: The *perceptual* proof scheme (discussed above, where students are unable to reason about the problem situation unless it is present perceptually, in which case their justification is solely based on their particular perception), the *spatial* proof scheme, which is a particular case of the contextual proof scheme (where students are able to think of the problem situation only in terms of their imaginative space), and the *axiomatic* proof scheme (to be discussed below, where students can think of the axioms in general structures, see below). The decisive majority of the students resided in the first two categories. Examples of the perceptual proof scheme were presented earlier; the following interview episode demonstrates the spatial proof scheme.

Episode 19. In the Teaching Experiment CG we introduced a hypothetical participant, who was called Mr. or Ms. Smart. This participant is an intellectually able creature with whom the students can communicate in their own natural language, including the language of basic set theory, but possesses none of the physical senses, such as visual and tactile perceptions. The class is presented with the task of communicating to Mr. or Ms. Smart geometric concepts, conjectures, propositions, and justifications they have formed intuitively and transformationally. The idea is to bring the students to realize that in order to facilitate such a communication, they must formulate certain "agreements" with Mr. or Ms. Smart. These agreements amount to a system of axioms. In the beginning, the instructor played the role of Mr. or Ms. Smart, but gradually the students took on her or his persona.

Prior to the following interview, the class discussed finite geometries, where the notions of "line" and "parallel" were defined for Mr. or Ms. Smart set theoretically.

I[11]: Consider the following model: The points are: A, B, C, D, E, F, G, and the lines are: $\{A, B, F\}$, $\{A, C, E\}$, $\{A, D, G\}$, $\{B, C, D\}$, $\{B, E, G\}$, $\{C, F, G\}$, $\{D, E, F\}$. Determine whether the following property holds in this model. "Given a line and a point NOT on the line, there is a line which contains the given point and is parallel to the given line."

Duane: If I were to take line ABF which of course contains those points ... no because one of those points [pointing to A, B, F] appears ... in each of the other lines so there won't be a line parallel to that.

I: You said the line ABF contains those points. Do you mean consists of these points?

Duane: I guess I don't understand the difference between contain and consist.

(The interviewer explains the difference between these two terms.)

Duane: There are other points on that line [line ABF]. Even if A, B, and F are consecutive points if it is defined as a line and that line will continue in each direction, so there has to be an infinite number of points on one line.

I: In this problem each line is defined to be consisting of exactly three points. There are no points on this line [pointing to $\{A, B, F\}$] other than the points A, B, and F.

Duane: To have a line that only contains A, B, F?

I: Yes.

Duane: It would be a segment with three points.

I: But a segment too contains infinitely many points.

Duane: ... To say a line consists of only three points is impossible because there should be an infinite number of points.

I: How did we define a line for Mr. Smart?

Duane: We defined the line for Mr. Smart as a set of consecutive points and that's all I can remember.

I: In terms of Mr. Smart's definition of a line, is there a model where the following property does not hold: "Between two points there exists a line, and such a line is unique," and (b) "Given a line and a point NOT on the line, there is a line containing the given point and is parallel to the given line."

Duane: No, because that defines the line on a plane. Between two points there exists a line and such a line is unique. . . . We defined a line as uh a line can be drawn, a unique line can be drawn between any two points in space. . . . And then for the second one we also defined the plane as ... The first one will have, would have to hold. And the second one, ... yeah, that would have to hold. I can't think of any alternatives.

(At this point the interviewer explained again to Duane that by a line it is not meant a Euclidean line, that is, his spatial imagery line, but a line as a set of objects, such as in Problem 1.)

I: Do you now understand the question?

Duane: Yeah I understand the question. I'm trying to think about different situations you can encounter. I think it's gonna go. OK. If I was given two points and I want to construct a situation where the two points cannot be joined by a line, is that what you're saying? ... No for number one I cannot think of anything ...

I: How about the possibility that you have more than one line going between two points? Duane: Uh that's not possible is it? . . . No you cannot have ... you can label it as two different lines but it's the same line.

I: OK. What if Mr. Smart took two lines *ABC* and *ABD*. Line *ABC* is a set consisting of three objects labeled *A*, *B*, and *C*, and a line *ABD* is also a set of objects labeled *A*, *B*, *D*. These objects Mr. Smart calls points. According to Mr. Smart, he has two lines that go through different points, *A* and *B*.

Duane: Well, I would have to, I would try to show him [Mr. Smart] that if we have two lines *ABC* and *ABD* ... and I would have to try to explain to him by the fact that uh *A*, *B*, and *C* are first line and *A*, *B*, and *D* are second line contain[ing] two common points. Therefore, since they contain two common points they have to lie on the same line or the same, ... it would be easier probably here to demonstrate it, to demonstrate them as two different rays. The line, a singly line ...

I: Let's say I am Mr. Smart, and I am telling you that what I mean by a line is just a set of three objects, no more and no less. This is what I call a line. The objects I call points.

Duane: Only a set of three points.

I: Yes. Would you then agree with Mr. Smart that between two points there is more than one line.

Duane: ... Why would he, Mr. Smart, define the lines as only three points. It has to be all the points contained [between] those three points.

I: What can't he?

Duane: Because if you have three points there, it can be named by those three points but not defined by those three points. Three points, you can only define three points as three points. If they're three non-collinear points then we can define a plane. If they're three collinear points you can define a line. But there has to be an infinite number of points between them.

This excerpt demonstrates Duane's inability to represent the geometric properties stated, in any context but his own imaginative space. Textbooks in axiomatic geometry, on the other hand, usually begin with finite geometries as a preparation for non-Euclidean geometries, not taking into account the impact of the spatial proof scheme. The idea that geometric properties are not supposed to evoke spatial imagery properties is a relatively new concept in mathematics; it was born at the turn of this century with the publication of Hilbert's *Grundlagen*. Poincaré, in his review of the *Grundlagen*, saw a need to point to this seemingly self-evident feature of the *Grundlagen*:

> ... the expressions "lie on," "pass through," etc., [in the *Grundlagen*] are not intended to evoke images; they are simply synonyms of the word "determine." The words "point," "straight line," and "plane" should not produce any sensible representation in the mind. They could with indifference designate objects of any nature whatever, provided that one can establish a correspondence among these objects so that there corresponds to each system of two of the objects called points one and only one object called a line [and so on]. (quoted in Mueller, 1981, p. 5)

So judging from the historical development of geometry, we believe that the spatial proof scheme is epistemologically inevitable. What surprises us is its robust influence on students in an advanced stage in their mathematics education.

3.3.1.2.2 Generic Proof Scheme. In a generic proof scheme,[12] conjectures are interpreted in general terms but their proof is expressed in a particular context. This scheme reflects students' inability to express their justification in general terms, as is demonstrated in the following episode.

Episode 20. In teaching experiment NT, several students proved the statement, "If a whole number is divisible by 9 then the sum of its digits is divisible by 9," by taking a specific whole number, say 867, and saying something to the effect: This number can be represented as $8 \times 100 + 6 \times 10 + 7$, which is $(8 \times 99 + 6 \times 9) + (8 + 6 + 7)$. Since the first addend, $8 \times 99 + 6 \times 9$, definitely is divisible by 9, the second addend, $8 + 6 + 7$, which is the sum of the number's digits, must be divisible by 9. Some of these students indicated, in addition, that this process can be applied to any whole number. In so doing, the students were utilizing a generic proof scheme.

The generic proof scheme was particularly apparent in students' justifications of statements involving the reduced echelon form of a matrix. Here is an example:

Episode 21. In response to the problem,

> Is the following statement true? "If $rank(A)$ is smaller than the number of unknowns in the system $AX = b$, then the system has infinitely many solutions."

Adam, a student in teaching experiment LA2, wrote:

> For this to be true the system must be consistent. Assuming consistency this [is] true because if the rank of A is less than # of unknowns one or

[12] Consistent with Tall's (1979), following Steiner (1976), notion of a generic proof. See also Harel and Tall (1991).

> more free variables will exist. These free variables can be any constants and are part of the spanning vectors in the solution. Therefore the solution will either [be] a line, plane, etc. and have an infinite # of solutions.

Apparently Adam sensed that his answer was not completely convincing because his main argument, "if the rank of A is less than # of unknowns one or more free variables will exist," was not justified, and so he continued:

$$\left[\begin{array}{ccc|c} 1 & 0 & a & b \\ 0 & 1 & c & d \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

> For this general system, $rank(A) = 2$, there are three unknowns.
> The solution is: $[x \quad y \quad z] = [b \quad d \quad 0] + z[-a \quad -c \quad 1]$.

We consider this latter part of Adam's response as an indication of a generic proof scheme, because his concept of rank—as we knew it from other instances—included the idea of rank of a matrix being the number of non-zero rows in the reduced echelon form of the matrix. So the matrix he presented is a representation of a "general system" as he had explicitly indicated.

3.3.1.2.3 Constructive Proof Scheme. In the constructive proof scheme, students' doubts are removed by actual construction of objects—as opposed to mere justification of the existence of objects. (This scheme is reminiscent of the constructivist mathematics philosophy founded by Brouwer at the turn of the 20th century.) Just as was the case historically, a manifestation of this proof scheme is students' dislike of proof by contradiction, as it is seen in the following episode.

Episode 22. In teaching experiment EC, Ed's proof of the theorem,

> A segment k emanating from the midpoint of AB in a triangle ABC and parallel to BC bisects AC.

was this:

> Let D and E be the midpoints of AB and AC, respectively, and assume DE is different from k. By previous result, DE must be parallel to AB. So DE and k have the point D in common and both are parallel to BC—a contradiction to the parallel postulate.

The instructor presented this proof in class during the CG teaching experiment. Dean, a student in this class, responded by saying,

> I really don't like proof by contradiction. I have never understood proofs by contradiction, they never made sense. You are assuming what was required to prove. Could you please give us a different proof?

3.3.2 Axiomatic Proof Schemes: Intuitive, Structural, and Axiomatizing.

When a person understands that at least in principle a mathematical justification must have started originally from undefined terms and axioms (facts, or statements accepted without proof), we say that person possesses an *axiomatic* proof scheme.

Such a person is necessarily aware of the distinction between the undefined terms, such as "point" and "line," and defined terms, such as "square" and "circle," and between statements accepted without proof and ones that are deducible from other statements.[13] He or she, however, may be able to handle only axioms that correspond to her or his intuition, or ideas of self-evidence, such as "for any a and b in F, $a + b = b + a$" in relation to her or his experience with real numbers, or "one and only one line goes through two points" in relation to her or his imaginative space. Such an axiomatic proof scheme we call an *intuitive-axiomatic* proof scheme.

Historically, one of the philosophical distinctions between Euclid's *Elements* and Hilbert's *Grundlagen* is that while the former is restricted to a single interpretation—that its content is a presumed description of human spatial realization—the latter is open to different possible realizations—such as the Euclidean space, the surface of a half sphere, the ordered pairs (x, y) and triples (x, y, z), where x, y, z are real numbers, etc.—including the interpretation that the axioms are meaningless formulas. In both cases special attention is paid to the formulation of undefined terms and accepted statements as the basis for any justification in mathematics, and therefore both are consistent with the axiomatic proof scheme. However, while the axioms in the *Elements* describe intuitively grasped truth—and therefore are consistent with the intuitive-axiomatic proof scheme[14]—the axioms in the *Grundlagen* characterize a *structure* that fits different models (Mueller, 1981). This obviously is not unique to geometry. In algebra, a group or a vector space is defined to be any system of objects satisfying certain axioms that specify the structure under consideration. Accordingly, we define a *structural* proof scheme as an axiomatic proof scheme by which one thinks of conjectures and theorems as representations of situations from *different* realizations that are understood to share a common structure characterized by a collection of axioms.

In the structural proof scheme, the axioms that define the structure are permanent, and the focus of the study is on the structure itself, not on the axiom system; so, for example, one studies real analysis on the basis of the axioms of

[13]Evidence exists to indicate that the distinction among the concepts of "definition," "axiom," and "theorem" is far from simple even for college students (see Vinner, 1977).

[14]The *Elements* is consistent with the intuitive proof scheme but does not entirely satisfy its definition, for proofs of some propositions are not grounded in the stated axioms. For example, the proposition that triangles which agree in two sides and the enclosed angle are congruent is proved by using the idea of motion and superimposition of figures, not the stated axioms.

a complete ordered field, or studies the theory of vector spaces on the basis of the vector space axioms, etc. We speculate that the structural proof scheme is a cognitive prerequisite to the *axiomatizing* proof scheme—a scheme by which a person is able to investigate the implications of varying a set of axioms, or to axiomatize a certain field.

The structural proof scheme is essential in studying many undergraduate mathematics topics, such as the theory of vector spaces, group theory, real analysis, etc. Without it, one cannot understand, for example, the need for the completeness axiom in real analysis or the need for the vector space axioms in linear algebra. In teaching experiment LA1 we found that most of the students viewed the vector-space axioms—as well as the basic properties that they imply—as a collection of trivial statements, which deserve no attention, as it is seen in the following episode:

Episode 23.

Interviewer: Consider the statement, "If x be an element of a vector-space V, then $(-1)x = -x$." Your friend complains to you that he does not understand the point made by this statement. Explain to your friend the argument made by this statement.

Student: To tell you the truth, I too think this is a ridiculous statement. Of course, negative one times x is negative x!

Further conversation with this student revealed that he restricted the context of the conjecture ("If x be an element of a vector-space V, then $(-1)x = -x$") to the context of R^n and was unable to think of it in general terms. Further, at the time of the interview, he was familiar with the operations in R^n and had experienced the fact that $(-1)x = -x$ for x in R^n and so he viewed this statement as self evident in this context. This phenomenon, which is another manifestation of the contextual proof scheme, hindered students appreciating the non-triviality of the statement, namely, that the product of the scalar -1 by any vector x produces the additive inverse of x. Evidently, students' past mathematical experiences were insufficient to build the structural proof scheme, which we believe is needed to cope with the theory of vector space.

An important distinction between the structural proof scheme and the intuitive proof scheme is the ability to separate the abstract statements of mathematics (e.g., $1+1=2$) from their corresponding quantitative observations (e.g., 1 apple + 1 apple = 2 apples), or the axiomatically-based observations from their corresponding visual phenomena. As an example of the latter, consider many students' difficulties in appreciating the need to prove the Intermediate Value Theorem in real analysis once they have formed its pictorial meaning. Finally, in our view the axiomatic proof scheme is epistemologically an extension of the transformational proof scheme. One might mistakenly think of the axiomatic proof scheme as the ability to reason formally, that is, to apply rules of inference to meaningless formulas. In fact, no philosopher of mathematics of the twentieth century seems to have maintained this position about formalism. Even Hilbert, the founder of the formalist movement, "looked on formalization as a means of solving certain mathematical questions, notably the question of consistency, but he regarded mathematics itself as the study of ideal objects created by the intellect to simplify treatment of the empirically and intuitively given" (Mueller, 1981, p. 7).

4. Recapitulation and Concluding Remarks

4.1. Recapitulation.

Studies of university students' understanding of mathematical proof have shown serious deficits. These findings are not surprising if collegiate mathematics instruction assumes on the part of entering students a general understanding of proof and its roles in mathematics, since studies of high school students have shown that only a limited number of them acquire a respectable degree of proof understanding and proof-writing ability during their high school mathematics. Thus for most university students, including even mathematics majors, university coursework must give conscious and perhaps overt attention to proof understanding, proof production, and proof appreciation as goals of instruction.

Accordingly, knowing what is guiding students' thinking about mathematical proof is essential. The notion of a psychological *proof scheme* has been very valuable to us in thinking about students' reactions in mathematical settings. The phrase "proof scheme" refers to what convinces a person, and to what the person offers to convince others. "Proof scheme" as used here refers to justifications in general, so it should not be interpreted narrowly in terms of mathematical proof in its conventional sense. One's proof scheme is idiosyncratic and may vary from field to field, and even within mathematics itself.

The classes of proof schemes offered here is based largely on extensive work with college mathematics majors and includes three major categories: *External conviction proof schemes*, *empirical proof schemes*, and *analytical proof schemes* (see Figure 1). Justifications under an external conviction proof scheme might depend on an authority such as a teacher or a book (the *authoritarian proof scheme*), on strictly the appearance of the argument (the *ritual proof scheme*), or on symbol manipulations, with the symbols and/or the manipulations having no meaningful basis in the context (the *symbolic proof scheme*). The latter two proof schemes are direct results of faulty instruction and have little to recommend them, since they lead to the endorsement or rejection of an argument solely on its appearance, or to presenting an argument based on the plausible-looking symbol manipulations, as in $\log(x+y) = \log x + \log y$. A dash of the authoritarian proof scheme, on the other hand, is not completely harmful and perhaps unavoidable, and people may use this scheme to some extent when they are sampling an area outside their specialties. In two of its worst forms, however, either the student is helpless without an authority at hand, or the student regards a justification of a result as valueless and unnecessary.

Empirical proof schemes are marked by their reliance on evidence from examples (sometimes just one example) of direct measurements of quantities, substitutions of specific numbers in algebraic expressions, etc. (the *inductive proof scheme*), or perceptions (the *perceptual proof scheme*). These proof schemes are commonly observed with mathematics students, perhaps partly because natural, everyday thinking utilizes examples so much. As with the authoritarian proof scheme, despite its grave limitations in mathematical proof,[15] the empirical proof scheme does have value. Examples and nonexamples enrich one's images and can help to generate ideas or to give insights, for instance. The problem arises in contexts in which a mathematical proof is expected, and yet all that is necessary or desirable in the eyes of the student is a verification by one or more examples.

The final categories—the *analytical proof schemes*—encompass mathematical proof, although again the emphasis is on the student's thinking rather than on what he or she writes. Key to the analytical proof schemes is the *transformational proof scheme*: the creation and transformations of general mental images for a context, with the transformations directed toward explanations, always with an element of deduction. At some point, a collection of cues and transformations may be chunked into a heuristic, yielding an *internalized (analytical) proof scheme*. Ideally an internalized proof scheme becomes itself an object of reflection (an *interiorized (analytical) proof scheme*), allowing comparison and contrast with other heuristics and usually giving greater insight into the context and/or the heuristic. Transformational proof schemes may be unnecessarily but unconsciously limited by the owner, either because of her or his restriction of the images to a familiar but narrow setting (e.g., an argument in R^3 for an R^n situation), or because of her or his use of a specific case in a generic way (e.g., a justification based on a specific or even a general 3×3 matrix in an $n \times n$ setting), or because of a need for an envisioned step-by-step creation of a result (e.g., a distrust of a proof by contradiction).

Axiomatic proof schemes, the second category of analytical proof schemes and an epistemological extension of the transformational proof scheme, involve an awareness of an underlying formal development. Although many undergraduates reach a point at which they can utilize an axiomatic proof scheme, their thinking may be grounded solely in familiar examples for very general ideas—e.g., the real numbers for a field, or $(Z_5, +)$ for a group, or a Euclidean line for a line, or a continuous real-valued function for a function—giving an *intuitive-axiomatic proof scheme*. With the realization that definitions and axioms about "fields," "groups," "lines," and "functions" can be applied very generally and can assume the central roles, one's proof scheme becomes a *structural axiomatic proof scheme*. An occasional undergraduate reaches this last stage and may then even reach a higher stage where he or she appreciates the possibility of alternate axiomatic developments for the same body of results—giving an *axiomatizing proof scheme*.

[15] "Mathematical proof" means a proof at the level of the analytical proof scheme.

4.2 Concluding Remarks.

By their natures, teaching experiments and interview studies do not give definitive conclusions. They can, however, offer indications of the state of affairs and a framework in which to interpret other work. The following comments and speculations should be read in that spirit.

The current analysis of proof schemes is no doubt not the last word on the typology. The analysis has been based mainly on work with students at one site, so it naturally describes their behaviors well. It has, however, also been tested to a degree by its adequacy in interpreting the work of mathematics majors at another university. Further validations should be done and will likely lead to a refinement of the categories.

It is important not to regard the taxonomy in a hierarchical, single-niche sense. A given person may exhibit various proof schemes during one short time span, perhaps reflecting her or his familiarity for, and relative expertise in, the contexts, along with her or his sense of what sort of justification is appropriate in the setting of the work. Although the authoritarian and empirical proof schemes have value, we feel that mathematics majors in particular should also eventually show evidence of the analytical proof schemes.

Although we prefer to speak of an individual's proof schemes in the plural so as to recognize that one might operate differently in different contexts, within a particular context there is often at least a partially hierarchical nature implicit in the categories. For example, in the analytic proof schemes, it is plausible if not definitional that the structural proof scheme is a cognitive prerequisite to the axiomatizing proof scheme, and we hypothesize that the transformational proof scheme is a necessary prerequisite to building an axiomatic proof scheme. On the other hand, we do not believe that the external proof schemes are essential in the development of the analytic proof schemes. The external authoritarian scheme and the empirical schemes should for mathematics students at some stage fill only confirming and conjecturing roles. van Dormolen (1977) hypothesized that the development of proof ideas in a given domain might follow a developmental sequence. With our terminology, this sequence might begin with a dependence on external schemes (e.g., the professor's assertions) and/or empirical schemes (e.g., studying examples) but grow gradually into analytic schemes; students in the teaching experiments have followed such paths. But it would be sobering if one must follow this crawl-before-walk-before-run in each domain, sobering because instruction may begin at the run stage, so to speak. Some description of student understanding of proof other than ours, however, might better describe a developmental sequence. All this conjecturing is provoked by the work of the van Hieles (cf. Fuys, Geddes, & Tischler, 1988), who identified different levels of understanding of geometric topics by precollege students. Their claim that students necessarily have to go through all the levels has some research support (summarized in Clements & Battista, 1992). Hence, instruction that begins at an advanced level might well not register with any students who are at earlier levels. In the proof schemes context, this hypothesis could mean, for example, that proofs at an axiomatic level for some topic might indeed be "over the head" of students whose proof schemes are at less sophisticated levels. Experienced students do recognize that there is an expected type of justification even though they might not be able to give it; we have often heard in the interviews statements like "I know this isn't a proof, but here is what I would do."

As is noted in section 2.1, the empirical proof scheme is not well understood. For example, our work at two different sites has given different impressions of students' comfort with the use of counterexamples. Is this an artifact of the limits of teaching experiments and interviews, or might some difference in the college (or precollege) curricula experienced at the two locations account for the contrast?

The importance of the setting in which the justification takes place naturally leads to implications for instruction. An awareness that there are different proof schemes, some unsophisticated, is important for the instructor. The dependence on the authoritarian and empirical proof schemes by most students must be a consideration in planning our teaching. We may, for example, be fostering the empirical proof schemes *through* our teaching: During instruction, empirical justifications themselves serve as examples of arguments given by mathematicians, and may inadvertently sanction the empirical proof scheme as a mode of justification fully acceptable in the mathematical context. Psychologists have found that some examples of natural concepts are psychologically more representative of the concepts than others are (cf. Mervis & Rosch, 1981). It may be that mathematics students who hear the word "proof" only in the presence of a certain form of argument may then generate their representative exemplar of proof on that basis and be guided by these aspects, perhaps ritualistic ones, in judging whether other arguments are proofs.

One role of our instruction is to exemplify or educe more powerful and logically sound proof schemes, and to make the case for their necessity and their value. The usual lecture-textbook method does appear to suffice for some students, but other students apparently only grudgingly "play the game" without a genuine conviction that mathematical proof is really necessary or personally goal-worthy. Students might mouth, for example, that examples do not give a "proof," but then offer only examples as justifications. To be fair, some explain that they can offer only examples because they do not know how even to start a mathematical proof, as Moore (1994) noted with his majors.

Writing a mathematical argument appears to require instruction; Wilf is convinced that there is a pay-off in devoting the "first few class hours" of his junior level analysis course to students' writing and critiquing proofs of familiar results from outside analysis (1996). In our teaching experiments, we adopted the proof-writing refinement-through-reflection approach, where students refine their own writing after they have reevaluated the meaning of their own proof from the perspective of another reader. Often students' written proofs include the right idea but are very difficult to understand because of inadequate wording and poor self expression. In many of these cases the instructor meets with the student individually to discuss the written proof. The goal is to bring the student to realize that her or his writing did not convey to an outsider the meaning he or she originally meant to convey, either because it is incomprehensible or because it conveys different meanings. While this approach appeared to be effective, it demanded from the instructors, as was expected, enormous amounts of time outside the regular classroom hours.

In the hope of fostering the transformational proof schemes, the teaching experiments have focused on the following technique. A carefully chosen problem, which might be called a "proof-eliciting problem" (a suggestive description due to R. Lesh, 1995), is presented to the students. They work on the problem, sometimes in groups, sometimes as individual homework or team homework. The students are then encouraged to express in their own words their justification for any conjecture. This ascertainment phase of the proof scheme—convincing oneself—is followed by the persuasion phase—convincing others—where the justification is examined by the whole class. Once the justification is accepted, the teacher usually gives a further assignment whose goal is to bring students to reflect upon the justification for further learning. Here is an example of a proof-eliciting problem.

Episode 24. At the beginning of the first lesson on modular arithmetic in teaching experiment NT, the class worked in groups on the problem: "Find the remainder when $(107)^4 + (107)^3 \times 2346 - 2376 \times 3475$ is divided by 5." One of the students, Laura, suggested her solution to her group-mates, who seemed to be satisfied with the solution. The instructor asked Laura to present her solution to the whole class. She presented her solution as follows: Laura first pointed to the addend, 2376×3475, arguing that its remainder, when divided by 5, is 0 because this product results in a number with unit digit 0. Then she looked at 2346 and argued that its remainder is 1 because its unit digit is 6. Next, Laura argued that the remainder for $(107)^3$ is 3. To defend her argument she wrote $(100 + 7)^3$ below the expression $(107)^3$, and said that she will be looking at 100^3 and 7^3 [sic]. She continued by arguing that the remainder for $(107)^3$ is 3 since the remainder for 100 is 0 and the remainder for 7^3 which is 343 is 3. Laura then indicated that in a similar way she computed the remainder for $(107)^4$ which she argued was 1 because the remainder for 100 is 0 and the remainder for 7^4, which is 2401, is 1. Lastly, Laura concluded the remainder for the expression when divided by 5 is $1 + 3 \times 1 + 0$, which is 4.

In the class discussion that followed, there were some students who agreed with Laura's actions and there were others who raised questions about their mathematical legitimacy. The instructor saw this as an opportunity for a possibly fruitful debate about properties of modular arithmetic operations. He asked the whole class to express the general statements that they believed constituted Laura's steps, in order to evaluate their correctness. After some discussion within each of the working groups, many students stated the general relationships believed to be implicitly used in Laura's actions. These statements then stood for evaluation by the whole class. Those which were later found to be true were declared as theorems. *Since these theorems originated from the students' own actions, they were better appreciated by them. More importantly, because of the uncertainty the students initially had about the truth of their statements, they saw a real need to seek justifications.* With shaping by the instructor in the choice of initial problems and sometimes in the provision of further test cases for conjectures or steps like Laura's erroneous $(a+b)^3 = a^3 + b^3$, students are led away from both an authoritarian proof scheme and a complete reliance on examples. When the classroom authority remains silent (see the account of the Schoenfeld course, this volume) and when the shortcomings of examples are accepted, the classwork leads naturally toward a search for a method of justification based on transformations of images and companion deductions.

The education of students toward transformational reasoning must not start in college. Years of instruction which focus on the results in mathematics, rather than the reasons behind those results, can leave the impression that *only* the results are important in mathematics, an opinion sometimes voiced even by mathematics majors. We argue that instructional activities that educate students to reason about situations in terms of the transformational proof schemes are crucial to students' mathematical development, and they must begin in an early age (see Harel, in press, for more details). Exciting projects in which children in even the primary grades routinely expect to give explanations for their thinking (e.g., Ball, 1993; Cobb et al., 1991; Lampert, 1990), show that the precollege mathematics curriculum can indeed promote proof schemes beyond the authoritarian and empirical ones.

Until such background is common among our college students, we can think about what proof schemes we are cultivating in our courses. When we go beyond authoritarian or example-based stances, will it be appreciated why the mathematical proofs we call for or give are needed? Should our lectures be a recitation of what we are writing on the chalkboard, or a "thinking out loud" which exemplifies transformational reasoning? With students' likely proof schemes in the back of our minds, is a transition course with a central concern the development of more sophisticated thinking about proof a necessity? If we are relying on a transition course, or on a discrete mathematics course, to bring students' proof production, understanding, and appreciation to a higher state, is the course accomplishing that goal? Finally, the difficulty of students in reaching a structural axiomatic proof scheme (cf. Episode 19) suggests that a capstone course including some attention to metamathematics as a topic might be of value to mathematics majors. As Babai notes, "One of the most remarkable gifts human civilization has inherited from ancient Greece is the notion of mathematical proof" (1992). Our students should share in this gift.

REFERENCES

1. Alexandrov, A. D. (1963). Non-Euclidean geometry. In A. D. Aleksandrov, A. N. Kolmogorov, & M. A. Lavrent'ev (Eds.), *Mathematics, its content, methods, and meaning, vol. III* (pp. 97–189) (K. Hirsch, Trans.). Cambridge, MA: M.I.T. Press.
2. Alibert, D. (1988). Towards new customs in the classroom. *For the Learning of Mathematics, 8*, 31–35.
3. Alibert, D., & Thomas, M. (1991). Research on mathematical proof. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 215–230). Dordrecht, The Netherlands: Kluwer.
4. Anderson, J. R. (1985). *Cognitive psychology* (2nd ed.). New York: W. H. Freeman.
5. Babai, L. (1992). Transparent proofs. *Focus. 12*(3), 1–2.
6. Balacheff, N. (1991). Treatment of refutations: Aspects of the complexity of a constructivist approach to mathematics learning. In E. von Glasersfeld (Ed.), *Radical constructivism in mathematics education* (pp. 89–110). Dordrecht, The Netherlands: Kluwer.
7. Ball, D. L. (1993). Halves, pieces, and twoths: Constructing and using representational contexts in teaching fractions. In T. P. Carpenter, E. Fennema, & T. A. Romberg (Eds.), *Rational numbers: An integration of research* (pp. 157–195). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
8. Bernal, M. (1987). *Black Athena: The Afroasiatic roots of classical civilization*. New Brunswick, NJ: Rutgers University Press.
9. Boyer, C. B. (1989). *A history of mathematics* (2nd edition). New York: Wiley.
10. Brousseau, G. (1986). Basic theory and methods in the didactics of mathematics. In P. F. L. Verstapen (Ed.), *Proceedings of the Second Conference on Systematic Co-operation Between Theory and Practice in Mathematics Education* (pp. 109–161). Enschede, The Netherlands: NICD.
11. Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics, 24*, 359–387.
12. Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 420–464). Reston, VA: National Council of Teachers of Mathematics.
13. Cobb, P., Wood, T., Yackel, E., Nicholls, J., Wheatley, G., Trigatti, B., & Perlwitz, M. (1991). Assessment of a problem-centered second-grade mathematics project. *Journal for Research in Mathematics Education, 22*, 3–29.
14. Davis, P. J. (1981). Are there coincidences in mathematics? *American Mathematical Monthly, 88*, 311–320.
15. Davis, P., & Hersh, R. (1981). *The mathematical experience*. Boston: Birkhauser.
16. Dubinsky, E. (1986). On the teaching of mathematical induction. *Journal of Mathematical Behavior, 5*(3), 305–17.
17. Dubinsky, E. (1989). On the teaching of mathematical induction II. *Journal of Mathematical Behavior, 8*, 285–304.
18. Fischbein, E., & 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* (pp. 128–131). Antwerp, Belgium: Universitaire Instelling Antwerpen.
19. Frid, S. (1994). Three approaches to undergraduate calculus instruction: Their nature and potential impact on students' language use and sources of conviction. In E. Dubinsky, A. H. Schoenfeld, & J. Kaput (Eds.), *Research in collegiate mathematics education I* (pp. 69–100). Washington, DC: Conference Board of the Mathematical Sciences.
20. Friedman, B. (1991). *Lectures on applications-oriented mathematics*. New York: Wiley.
21. Fuys, D., Geddes, D., & Tischler, R. (1988). *The van Hiele model of thinking in geometry among adolescents* (JRME Monograph No. 3.). Reston, VA: National Council of Teachers of Mathematics.
22. Galbraith, P. (1981). Aspects of proving: a clinical investigation of process. *Educational Studies in Mathematics, 12*, 1–29.
23. Goetting, M. M. (1995). The college student's understanding of mathematical proof (Doctoral dissertation, University of Maryland, 1995). *Dissertations Abstracts International,*

56-A, 3016, Feb., 1996.
24. Halmos, P. (1985). *I want to be a mathematician*. New York: Springer-Verlag.
25. Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange, 21*, 6–13.
26. Harel, G. (1985). *Teaching linear algebra in high school*. Unpublished Doctoral Dissertation, Ben-Gurion University of the Negev, Israel.
27. Harel, G. (1995). From naive interpretist to operation conserver. In J. Sowder & B. Schappelle (Eds.), *Providing a foundation for teaching mathematics in the middle grades* (pp. 143–165). New York: SUNY Press.
28. Harel, G. (in preparation). Symbolic reasoning and transformational reasoning and their effect on algebraic reasoning. In James J. Kaput (Ed.), *Employing children's natural powers to build algebraic reasoning in the content of elementary mathematics*.
29. Harel, G., & Tall, D. (1991). The general, the abstract, and the generic. *For the Learning of Mathematics, 11*, 38–42.
30. Kitcher, P. (1983). *The nature of mathematical knowledge*. New York: Oxford University Press.
31. Kleiner, I. (1991). Rigor and proof in mathematics: A historical perspective. *Mathematics Magazine, 64*(5), 291–314.
32. 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.
33. Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.
34. Leron, U. (1983). Structuring mathematical proofs. *American Mathematical Monthly, 90*, 174–185.
35. Lesh, R. (1995). Personal communication.
36. Lovell, K. (1971). The development of the concept of mathematical proof in abler pupils. In M. Rosskopf, L. Steffe, & S. Taback (Eds.), *Piagetian cognitive-development research and mathematical education* (pp. 66–80). Washington, DC: National Council of Teachers of Mathematics.
37. MacPherson, E. D. (1985). The themes of geometry: design of the nonformal geometry curriculum. In C. Hirsch & M. Zweng (Eds.), *The secondary school mathematics curriculum*, 1985 Yearbook (pp. 65–80). Reston, VA: National Council of Teachers of Mathematics.
38. Martin, G., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education, 20*(1), 41–51.
39. Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*, 89–115.
40. Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics, 27*, 249–266.
41. Movshovits-Hadar, N. (1988). Stimulating presentation of theorems followed by responsive proofs. *For the Learning of Mathematics, 8*, 12–19.
42. Mueller, I. (1981). *Philosophy of Mathematics and Deductive Structure in Euclid's Elements*. Cambridge, MA: MIT Press.
43. National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.
44. Piaget, J. (1983). Piaget's theory. In P. H. Mussen (Ed.), *Handbook of Child Psychology* (4th ed., pp. 103–128). New York: John Wiley & Sons.
45. Piaget, J., & Inhelder, B. (1967). *The child's conception of space*. New York: W.W. Norton & Company.
46. Porteous, K. (1986). Children's appreciation of the significance of proof. *Proceedings of the Tenth International Conference of the Psychology of Mathematics Education* (pp. 392–397). London, England.
47. Schoenfeld, A. (1985). *Mathematical problem solving*. Orlando, FL: Academic Press.
48. Semadeni, Z. (1980). Action proofs in primary mathematics teaching and in teacher training. *For the Learning of Mathematics, 4*, 32–34.
49. Senk, S. L. (1985). How well do students write geometry proofs? *Mathematics Teacher, 78*(6), 448–456.

50. Shaughnessy, J. M., & Burger, W. F. (1985). Spadework prior to deduction in geometry. *Mathematics Teacher, 78*(6), 419–428.
51. Simon, M. A. (1996). Beyond inductive and deductive reasoning: the search for a sense of knowing. *Educational Studies in Mathematics, 30*, 197–210.
52. Skemp, R. R. (1978). Relational understanding and instrumental understanding. *Arithmetic Teacher, 26*, 9–15.
53. Steiner, M. (1976). Mathematical explanation. Mimeographed notes, Columbia University.
54. Steffe, L., Cobb, P., & von Glasersfeld, E. (1988). *Construction of arithmetical meanings and strategies*. New York: Springer-Verlag.
55. Tall, D. O. (1979). Cognitive aspects of proof, with special reference to the irrationality of $\sqrt{2}$. *Proceedings of the Third Conference of the Psycology of Mathematics Education* (pp. 203–205). Warwick, England.
56. Thompson, D. R. (1991, April). *Reasoning and proof in precalculus and discrete mathematics*. Paper presented at the meeting of the American Educational Research Association, Chicago.
57. Thompson, D. R. (1996). Learning and teaching indirect proof. *Mathematics Teacher, 89*, 474–482.
58. Thompson, P. W. (1994). Images of rate and operational understanding of the Fundamental Theorem of Calculus. *Educational Studies in Mathematics, 26*(2–3), 229–274.
59. Usiskin, Z. (1980). What should not be in the algebra and geometry curricula of average college-bound students? *Mathematics Teacher, 73*, 413–424.
60. Usiskin, Z. (1987). Resolving the continuing dilemmas in school geometry. In M. M. Lindquist & A. P. Shulte (Eds.), *Learning and Teaching Geometry, K–12, 1987 Yearbook* (pp. 17–31). Reston, VA: National Council of Teachers of Mathematics.
61. van Dormolen, J. (1977). Learning to understand what giving a proof really means. *Educational Studies in Mathematics, 8*, 27–34.
62. Vinner, S. (1977). The concept of exponentiation at the undergraduate level and the definitional approach. *Educational Studies in Mathematics, 8*, 17–26.
63. Vinner, S. (1983). The notion of proof: Some aspects of students' view at the senior high level. *Proceedings of the Seventh International Conference of the Psychology of Mathematics Education*. Rehovot, Israel: Weizmann Institute of Science.
64. Volmink, J. (1988). The role of proof in students' understanding of geometry. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.
65. Wilf, H. S. (1996). Epsilon sandwiches. *Focus. 16*(2) 24–25.
66. Williams, E. (1980). An investigation of senior high school students' understanding of the nature of mathematical proof. *Journal for Research in Mathematics Education, 11*(3), 165–166.
67. Yerushalmy, M. (1986). Induction and generalization: An experiment in teaching and learning high school geometry. Doctoral dissertation, Harvard University.

PURDUE UNIVERSITY, WEST LAFAYETTE, IN

SAN DIEGO STATE UNIVERSITY, SAN DIEGO, CA