

**Toward Comprehensive Perspectives on the Learning and Teaching of Proof<sup>1</sup>**

Guershon Harel  
Department of Mathematics  
University of California, San Diego  
San Diego, CA 92093-0112, USA  
harel@math.ucsd.edu  
(858)534-2650  
(858)534-5273 (Fax)

Larry Sowder  
Department of Mathematics and Statistics  
San Diego State University  
San Diego, CA 92182-7720, USA  
lsowder@sciences.sdsu.edu  
(619)594-6179  
(619)594-6746 (Fax)

To Appear in F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning*, National Council of Teachers of Mathematics.

One of the most remarkable gifts human civilization has inherited from ancient Greece is the notion of mathematical proof. The basic scheme of Euclid's *Elements* has proved astoundingly durable over the millennia and, in spite of numerous revolutionary innovations in mathematics, it still guides the patterns of mathematical communication (Babai, 1992).

### Organization

This chapter examines proof in mathematics, both informal justifications and the types of justification usually called mathematical proofs. The introductory section below calls for what we label a “comprehensive perspective” toward the examination of the learning and teaching of proof, and identifies the various elements of such a comprehensive perspective. Our viewpoint next centers on students' outlooks on proof, as described by the “proof schemes” evidenced in students' work; the second section elaborates on this proof-scheme notion and includes a description of various proof schemes as well as a listing of the various roles proof can play in mathematics. The third section then gives a brief overview of how the idea of proof in mathematics has evolved historically, and why historical considerations could be a part of educational research on the learning or teaching of proof. The fourth and fifth sections include a look at selected studies<sup>2</sup> dealing with proof, at both the precollege and the college levels, with an effort to show the value of the proof-schemes idea. The final section offers some questions prompted by the earlier sections.

### Comprehensive Perspective on Proof<sup>3</sup>

No one questions the importance of proof in mathematics, and in school mathematics (Ball & Bass, 2003; Haimo, 1995; Schoenfeld, 1994). Overall, the performance of students at the secondary and undergraduate levels in proof is weak, as the findings reported in this paper will show. Whether the cause lies in the curriculum, the textbooks, the instruction, the teachers' background, or the students themselves, it is clear that the status quo needs, and has needed, improvement. Earlier reviews of research on the teaching and learning of proof (e.g., Battista & Clements, 1992; Hart, 1994; Tall, 1991; Yackel & Hanna, 2003) have informed and inspired more recent studies of proof learning. This chapter argues for “comprehensive perspectives” on proof learning and teaching and provides an example of such a perspective. A comprehensive perspective on the learning and teaching of proofs is one that incorporates a broad range of factors: mathematical, historical-epistemological, cognitive, sociological, and instructional. A unifying and organizing element of our perspective is the construct of “proof scheme.” This construct emerged from a long sequence of studies into the concept of proof (Harel & Sowder, 1998), and, in turn, was utilized to address a cohort of foundational questions concerning these factors:

#### Mathematical and Historical-Epistemological Factors

1. What is proof and what are its functions?
2. How are proofs constructed, verified, and accepted in the mathematics community?
3. What are some of the critical phases in the development of proof in the history of mathematics?

#### Cognitive Factors

4. What are students' current conceptions of proof?
5. What are students' difficulties with proof?
6. What accounts for these difficulties?

#### Instructional-Socio-Cultural Factors

7. Why teach proof?
8. How should proof be taught?
9. How are proofs constructed, verified, and accepted in the classroom?
10. What are the critical phases in the development of proof with the individual student and within the classroom as a community of learners?
11. What classroom environment is conducive to the development of the concept of proof with students?
12. What form of interactions among the students and between the students and the teacher can foster students' conception of proof?
13. What mathematical activities—possibly with the use of technology—can enhance students' conceptions of proof?

This list does not purport to be exhaustive or unique; other researchers may choose a different list of questions or formulate those that appear here differently in their attempt to form their comprehensive perspective to proof. This list, however, represents those questions that we have confronted in a decade of investigations into the learning and teaching proof; hence, these questions constitute an essential part of the *content* of our comprehensive perspective on proof. Nor is the classification of these questions as it is outlined here invariable. Clearly, it is difficult—if not impossible—to classify these questions into pairwise disjoint categories of factors. Our use of the labels indicates a loose association rather than a firm classification. In particular, none of the questions should be viewed as a stand-alone question; rather they are all interrelated, constituting a cohort of questions that have guided the comprehensive perspective on proof offered here as an example. Thus, even though the headings of the third, fourth, and fifth sections of the chapter correspond to the three categories outlined in the above list, the content of each section pertains to the other two.

An important character of our perspective is that of *subjectivity*: While the term “proof” often connotes the relatively precise argumentation given by mathematicians, in our perspective “proof” is interpreted subjectively; a proof is what establishes truth for a person or a community. With this interpretation, “proof” connotes an activity that can permeate the whole mathematics curriculum, from kindergarten on as well as throughout the historical development of mathematics. This subjective notion of “proof” is the most central characteristic of the construct of “proof scheme,” which will be defined shortly. When we wish to make clear that we mean the mathematically institutionalized notion of proof, we will write “mathematical proof,” otherwise the terms “proof” or “justification” will be used in this subjective sense.

Subjectivity equally applies to how the above questions should be understood. Some of these questions, particularly the first three, might be interpreted as calling for objective, clear-cut answers. Not so. The meaning of proof, its role, and the way it is created, verified, and accepted may vary from person to person and from community to community. One's answers to these questions are greatly influenced by her or his philosophical orientation to the processes of learning and teaching and would reflect her or his answers to questions such as: What bearing, if any, does the epistemology of proof

in the history of mathematics have on the conceptual development of proof with students? What bearing, if any, does the way mathematicians construct proofs have on instructional treatments of proof? What bearing, if any, does everyday justification and argumentation have on students' proving behaviors in mathematical contexts? Our emphasis on subjectivity—a motive we will repeat a few times in this paper—stems from the well-known recognition that students' construction of new knowledge is based on what they already know (e.g., Cobb, 1994; Piaget, 1952, 1973a, b, 1978; Vygotsky, 1962, 1978), and hence it is indispensable for teachers to identify students' current knowledge, regardless of its quality, so as to help them gradually refine it. To avoid unnecessary misunderstanding we note here, and we will repeat later, that subjectivity toward the meaning of proof does not imply ambiguous goals in the teaching of this concept. Ultimately, the goal is to help students gradually develop an understanding of proof that is consistent with that shared and practiced in contemporary mathematics.

### *Why are Comprehensive Perspectives on Proofs Needed?*

Comprehensive perspectives on proof are needed in an effort to understand students' difficulties, the roots of the difficulties, and the type of instructional interventions needed to advance students' conceptions of and attitudes toward proof. A single factor usually is not sufficient to account for students' behaviors with proof. For example, while some of the difficulties students have with proof can be accounted for by the cognitive factor, such as the students' lack of logical maturity and understanding of the need for proof, research studies conducted in the last two decades have given evidence that these sources are insufficient to provide a full picture of students' difficulties with mathematics in general and proof in particular. Balacheff's (1991) research, for example, has shown that students "have some awareness of the necessity to prove and some logic" (p. 176) and yet they experience difficulty with proof.

Official documents and research papers on the learning and teaching of proof almost never present their theoretical framework as a comprehensive perspective on proof, but stated curricular goals, instructional recommendations, research design, and so forth, may reveal some elements in the scholars' (potential) perspective on proof. For example, Balacheff (1991) points out that common to mathematics curricula in different parts of the world is the goal of training students in the construction and the formulation of "deductive reasoning," which is defined as "a careful sequence of steps with each step following logically from an assumed or previously proved statement and from previous steps" (National Council of Teachers of Mathematics [NCTM], 1989, p. 144). Balacheff points out that this characterization of proof, which is close to what a logician would formulate, is almost the same in mathematics curricula all over the world. Remarks in these documents on the conceptual development of proof with students stress that proof "has nothing to do with empirical or experimental verification" and "call attention to the move from concrete to abstract" (Balacheff, 1991, p. 177). This suggests a form-driven perspective, according to which the basis for one's answer to the question, how should proof be taught, is the form in which mathematics is organized and accordingly presented in scientific papers or books.

Approaching the concept of proof from an encompassing perspective such as the one suggested here is, in our view, a critical research need. Balacheff, a world-recognized pioneer in the area of the learning and teaching of proof, has addressed this

need by raising on several occasions (e.g., 2002) the question: “Is there a shared meaning of ‘mathematical proof’ among researchers in mathematics education?” Balacheff was not referring to just the standard, more or less formal definition of mathematical proof. Rather, his question, as he puts it, is “whether beyond the keywords, we had some common understanding” (Balacheff, 2002, p. 23). By “common understanding,” we believe he means agreed-upon parameters in terms of which one can formulate differences among perspectives into research questions. We agree with Balacheff that without such an understanding it is hard to envision real progress in our field. The comprehensive perspective on proof presented in this paper delineates a set of such parameters.

### The Concept of Proof Scheme

Consistent with our characterization of a comprehensive perspective on proof, several critical factors must be considered in addressing the question, What is proof? First, the construction of new knowledge does not take place in a vacuum but is shaped by existing knowledge. What a learner knows now constitutes a basis for what he or she will know in the future. This fundamental, well documented fact has far-reaching instructional implications. When applied to the concept of proof—our concern in this paper—this fact requires that an answer to the above question takes into account the student as a learner, that is, the cognitive aspects involved in the development of the concept of proof with the individual student. Second, one must maintain the integrity of the concept of proof as has been understood and practiced throughout the history of mathematics. Third, since the concept of proof is social—in that what is offered as a convincing argument by one person must be accepted by others—one must take into account the social nature of the proving process. We see, therefore, that one’s answer to the above question attends to a range of factors: cognitive, mathematical, epistemological-historical, and social. The approach we provide here takes into consideration these factors. The conceptual framework of our answer has been formed over a long period of extensive work on students’ conception of proof, and it incorporates findings reported in the literature as well as from our own studies. The latter included a range of teaching experiments and a three-year longitudinal study (see, for example, Harel & Sowder, 1998; Harel, 2001; Sowder & Harel, 2003) as well as historical, philosophical, and cultural analyses (e.g., Hanna, 1983, Hanna & Jahnke, 1996, Harel, 1999, Kleiner, 1991). The foundational element of this framework is the concept of “proof scheme.”

Obviously, learners’ knowledge, in general, and that of proof, in particular, is not homogenous; most commonly, a high-school student’s conception of proof is different from that of a college mathematics major student, and a contemporary mathematician’s conception of proof is different from that of an ancient mathematician, which, in turn, was different from that of a Renaissance mathematician, etc. “Proof scheme” is a term we use to describe one’s (or a community’s) conception of proof; it will be defined precisely below. In Harel and Sowder (1998) we offer an elaborate taxonomy of students’ proof schemes based on seven teaching experiments with a total of 169 mathematics and engineering majors. The experiments were conducted in classes of linear algebra (elementary and advanced), discrete mathematics, geometry, and real analysis. Later, in Harel (in press), this taxonomy was refined and extended to capture

more observations of students' conceptions of proof as well as some developments of this concept in the history of mathematics. In this paper we use our taxonomy of proof schemes to describe or interpret

- (a) historical developments and philosophical matters concerning the concept of proof (third section),
- (b) results from other studies concerning students' conception of and performance with proof (fourth section), and
- (c) curricular and instructional issues concerning proof (fifth section).

#### *Definition of "Proof Scheme"*

The definition of "proof scheme" is based on three definitions.

1. ***Conjecture versus fact.*** An assertion can be conceived by an individual either as a *conjecture* or as a *fact*: (A conjecture is an assertion made by an individual who is uncertain of its truth.) The assertion ceases to be a conjecture and becomes a fact in the person's view once he or she becomes certain of its truth.

A critical question, to which we will return shortly, is the following: How in the context of mathematics do students render a conjecture into a fact? That is, how do they become certain about the truth of an assertion in mathematics?

The definitions of "conjecture" and "fact" are the basis for the notion of *proving*.

2. ***Proving.*** *Proving* is the process employed by an individual (or a community) to remove doubts about the truth of an assertion.

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

3. ***Ascertaining versus persuading.*** *Ascertaining* is the process an individual (or a community) employs to remove her or his (or its) own doubts about the truth of an assertion. *Persuading* is the process an individual or a community employs to remove others' doubts about the truth of an assertion.

Mathematics as sense-making means that one should not only ascertain oneself that the particular topic/procedure makes sense, but also that one should be able to convince others through explanation and justification of her or his conclusions. In particular, the convince-others, public side of proof is a social practice (e.g., Bell, 1976), not only for mathematicians but for all students of mathematics. The definition of "persuading" aims at capturing this essential feature of proving. As defined, the process of proving includes two processes: *ascertaining* and *persuading*. Seldom do these processes occur in separation. Among mathematically experienced people and in a classroom environment conducive to intellectual interactions among the students and between the students and the teacher, when one ascertains for oneself, it is most likely that one would consider how to convince others, and vice versa. Thus, proving emerges as a response to cognitive-social needs, rather than exclusively to cognitive needs or social needs—a view consistent with Cobb and Yackel's *emergent perspective* (1996).

As defined, ascertaining and persuading are entirely subjective, for one's proving can vary from context to context, from person to person, from civilization to civilization, and from generation to generation within the same civilization (cf. Harel & Sowder, 1998, Kleiner, 1991, Raman, 2003, Weber, 2001). Thus, we offer the following definition:

- 4. Proof scheme.** A person's (or a community's) proof scheme consists of what constitutes ascertaining and persuading for that person (or community).

We repeat here what we highlighted in Harel and Sowder (1998): Our definitions of the process of proving and proof scheme are deliberately student-centered. Terms such as “to prove,” “to conjecture,” “proof,” “conjecture,” and “fact” must be interpreted in this subjective sense throughout the paper. We emphasize again that despite this subjective definition the goal of instruction must be unambiguous—namely, to gradually refine current students' proof schemes toward the proof scheme shared and practiced by contemporary mathematicians. This claim is based on the premise that such a shared scheme exists and is part of the ground for scientific advances in mathematics.

### *Taxonomy of Proof Schemes*

The taxonomy of proof schemes consists of three classes: the *external conviction proof schemes class*, the *empirical proof schemes class*, and the *deductive proof schemes class*. Below is a brief description of each of these schemes and some of their subschemes. Further elaboration of the different schemes is provided, as needed, throughout the paper. Relations of our taxonomy to other taxonomies and to the functions of proof within mathematics are addressed at the end of this subsection. (For the complete taxonomy, see Harel & Sowder (1998) and Harel (in press).)

**External conviction proof schemes.** Proving within the *external conviction proof schemes class* depends (a) on an authority such as a teacher or a book, (b) on strictly the appearance of the argument (for example, proofs in geometry must have a two-column format), or (c) on symbol manipulations, with the symbols or the manipulations having no potential coherent system of referents (e.g., quantitative, spatial, etc.) in the eyes of the student (e.g.,  $(a + b)/(c + b) = (a + \cancel{b})/(c + \cancel{b}) = a/c$ ). Accordingly, we distinguish among three proof schemes within the *external conviction proof schemes class*:

- *External conviction proof schemes class*
  - *Authoritarian proof scheme*
  - *Ritual proof scheme*
  - *Non-referential symbolic proof scheme*

**Empirical proof schemes.** Schemes in the *empirical proof scheme class* are marked by their reliance on either (a) evidence from examples (sometimes just one example) of direct measurements of quantities, substitutions of specific numbers in algebraic expressions, and so forth, or (b) perceptions. Hence, we distinguish between two proof schemes within the *empirical proof scheme class*:

- *Empirical proof schemes class*
  - *Inductive proof schemes*
  - *Perceptual proof schemes*

**Deductive proof schemes.** The *deductive proof schemes class* consists of two subcategories, each consisting of various proof schemes:

- *Deductive proof schemes class*
  - *Transformational proof schemes*
  - *Axiomatic proof schemes*

All the *transformational proof schemes* share three essential characteristics: *generality*, *operational thought*, and *logical inference*.

The *generality* characteristic has to do with an individual understanding that the goal is to justify a “for all” argument, not isolated cases and no exception is accepted. Evidence that *operational thought* is taking place is shown when an individual forms goals and subgoals and attempts to anticipate their outcomes during the evidencing process. Finally, when an individual understands that justifying in mathematics must ultimately be based on logical inference rules, the *logical inference* characteristic is being employed (see also Harel, 2001).

Unlike the proof schemes in the previous two classes (the external conviction proof schemes class and the empirical proof schemes class), the transformation proof schemes require a more elaborate demonstration. Consider the following two responses (taken from Harel, 2001) to the problem:

Prove that for all positive integers  $n$ ,  
 $\log(a_1 \cdot a_2 \cdots a_n) = \log a_1 + \log a_2 + \cdots + \log a_n$ .

*Response 1*

$$\log(4 \cdot 3 \cdot 7) = \log 84 = 1.924$$

$$\log 4 + \log 3 + \log 7 = 1.924$$

$$\log(4 \cdot 3 \cdot 6) = \log 72 = 1.857$$

$$\log 4 + \log 3 + \log 6 = 1.857$$

Since these work, then  $\log(a_1 \cdot a_2 \cdots a_n) = \log a_1 + \log a_2 + \cdots + \log a_n$ .

A probe into the reasoning of the students who provide responses of this kind reveals that their conviction stems from the fact that the proposition is shown to be true in a few instances, each with numbers that are *randomly* chosen—a behavior that is a manifestation of the empirical proof scheme.

*Response 2*

(1)  $\log(a_1 a_2) = \log a_1 + \log a_2$  by definition

(2)  $\log(a_1 a_2 a_3) = \log a_1 + \log a_2 a_3$ . Similar to  $\log(ax)$  as in step (1), where this time  $x = a_2 a_3$ .

Then

$$\log(a_1 a_2 a_3) = \log a_1 + \log a_2 + \log a_3$$

(3) We can see from step (2) any  $\log(a_1 a_2 a_3 \cdots a_n)$  can be repeatedly broken down to

$$\log a_1 + \log a_2 + \cdots + \log a_n$$

It is important to point out that in Response 2 the student recognizes that the process employed in the first and second cases constitutes a pattern that recursively applies to the entire sequence of propositions,  $\log(a_1 a_2 \cdots a_n) = \log a_1 + \log a_2 + \cdots + \log a_n$ ,  $n = 1, 2, 3, \dots$

In both responses the generalizations are made from two cases. This may suggest, therefore, that both are empirical. As is explained in Harel (2001), this is not so: response 2, unlike response 1, is an expression of the transformational proof scheme. To see why, one needs to examine the two responses against the definitions of the two



schemes. While both responses share the first characteristic—i.e., in both the students respond to the “for all” condition in the log-identity problem statement—they differ in the latter two: whereas the mental operations in Response 1 are incapable of anticipating possible subsequent outcomes in the sequence and are devoid of general principles in the evidencing process, the mental operations in Response 2 correctly predict, on the basis of the general rule,  $\log(ax) = \log a + \log x$ , that the same outcome will be obtained in each step of the sequence. Further, in Response 1 the inference rule that governs the evidencing process is empirical; namely,  $(\exists r \in R)(P(r)) \Rightarrow (\forall r \in R)(P(r))$ . In Response 2, on the other hand, it is deductive; namely, it is based on the inference rule  $(\forall r \in R)(P(r)) \wedge (w \in R) \Rightarrow P(w)$ . (Here  $r$  is any pair of real numbers  $a$  and  $x$ ,  $R$  is the set of all pairs of real numbers,  $P(r)$  is the statement “ $\log(ax) = \log a + \log x$ ,” and  $w$  in step  $n$  is a pair of real numbers  $a_1 a_2 \cdots a_{n-1}$  and  $a_n$ .)

The axiomatic proof scheme too has the three characteristics that define the transformational proof scheme, but it includes others. For now, it is sufficient to define it as a transformational proof scheme by which one understands that in principle any proving process must start from accepted principles (axioms). The situation is more complex, however, as we will show in the section on historical and epistemological considerations (third section). For the purpose of this chapter we will introduce only the Greek axiomatic proof scheme and the modern axiomatic proof scheme—as manifested, for example, in Euclid’s *Elements* and Hilbert’s *Grundlagen*, respectively. The distinction between these two schemes is further discussed in the third section.

### *Relations to Other Taxonomies*

In broad terms, the empirical proof schemes and the deductive proof schemes correspond to what Bell (1976) calls “empirical justification” and “deductive justification,” and Balacheff (1988) calls “pragmatic” justifications and “conceptual” justifications, respectively. Balacheff further divides the pragmatic justification into three types of justifications: “naïve empiricism” (justification by a few random examples), “crucial experiment” (justification by carefully selected examples), and “generic example” (justification by an example representing salient characteristics of a whole class of cases). “Generic example” in our taxonomy belongs to the deductive proof scheme category, for an analysis similar to that we applied to Response 2 (above) will show that “generic example” satisfies the three characteristics of the transformational proof scheme. Balacheff further classifies conceptual justifications into two types: “thought experiment,” where the justification is disassociated from specific examples, and “symbolic calculation,” where the justification is based solely on transformation of symbols. In our taxonomy, the latter corresponds to the *referential symbolic* proof scheme. This scheme is the direct opposite of the *non-referential* proof scheme defined above. Recall this is a scheme where neither the symbols nor the operations one performs on them represent a coherent referential reality for the student. Rather, the student thinks and treats symbols and operations on them as if they possess a life of their own without reference to their functional or quantitative meaning. In the referential symbolic proof scheme, to prove or refute an assertion or to solve a problem, one learns to represent the situation algebraically and performs symbol manipulations on the

resulting expressions, with the intention to derive information relevant to the problem at hand. We return to this scheme in the next section.

The above definitions and taxonomy are not explicit enough about many critical functions of proof within mathematics. There is a need to point to these functions due to their importance in mathematics in general and to their instructional implications in particular. For this, we point to the work of other scholars in the field, particularly the work by Hanna (1990), Balacheff (1988), Bell (1976), Hersh (1993), and de Villiers (1999). de Villiers, who built on the work of the others mentioned here, raises two important questions about the role of proof: (a) What different functions does proof have within mathematics itself? and (b) How can these functions be effectively utilized in the classroom to make proof a more meaningful activity? (p. 1). According to de Villiers, mathematical proof has six not mutually exclusive roles: verification, explanation, discovery, systematization, intellectual challenge, and communication. At the end of the next section, after the relevant schemes are defined, we show that these functions are describable in terms of the proof scheme construct.

#### Mathematical and Historical-Epistemological Factors: Some Phases in the Development of Proof Schemes in the History of Mathematics<sup>4</sup>

Deductive reasoning is a mode of thought commonly characterized as a sequence of propositions where one must accept any of the propositions to be true if he or she has accepted the truth of those that preceded it in the sequence. This mode of thought was conceived by the Greeks more than 20 centuries ago and is still dominant in the mathematics of today. So remarkable is the Greeks' achievement that their mathematics became an historical benchmark to which other kinds of mathematics are compared. We are here particularly interested in analyzing the proof scheme construct across three periods of mathematics: Greek mathematics, post-Greek mathematics (approximately from the 16<sup>th</sup> to the 19<sup>th</sup> century), and modern mathematics (from the late 19<sup>th</sup> century to today). For the sake of completeness we say a few words about pre-Greek mathematics.

Proof schemes are applied to establish assertions in specific contexts. In this respect, therefore, in addressing proof schemes one must attend to the nature of the context about which the assertions are made. Also of importance is the motive—the intellectual need—that might have brought about the conceptual change from one period to the next. Thus, our discussion will center on three interrelated aspects of the historical-epistemological development: (a) the context of proving, (b) the means of proving (i.e., proof schemes), and (c) the motive for conceptual change. Understanding these elements can shed light on some critical aspects of the learning and teaching of proof, as we will see.

#### *Some Fundamental Differences*

##### *Practical world versus ideal existence: the emergence of structure.*

Generally speaking, pre-Greek mathematics was concerned merely with *actual* physical entities, particularly with quantitative measurement of different objects. Their “formulas” were in the form of prescriptions providing mostly approximations for measured quantities (area, volume, etc.). Ancient geometry developed in an empirical way through phases of trials and errors, where conjectures were proved by means of

empirical—inductive or perceptual—evidence. Even Babylonian mathematics, which, according to Kleiner (1991, p. 291), “is the most advanced and sophisticated of pre-Greek mathematics,” lacked the concept of proof as was understood and applied by the Greeks. Accordingly, it is safe to conclude that proving in pre-Greek mathematics is, by and large, governed by the empirical proof schemes.

The Greeks elevated mathematics from the status of practical science to a study of abstract entities. In their mathematics the particular entities under investigation are idealizations of experiential spatial realities and so also are the propositions on the relationships among these entities. The difference between pre-Greek mathematics and Greek mathematics, however, is not just in the nature of the entities considered—actual spatial entities versus their idealizations—but also in the reasoning applied to establish truth about the entities. In Greek mathematics, logical deduction came to be central in the reasoning process, and it alone necessitated and cemented the geometric edifice they created. To construct their geometric edifice the Greeks had to create primary terms—terms admitted without definition—and primary propositions—propositions admitted without proof—what the Greeks called axioms or postulates. In contrast, the mathematics of the civilizations that preceded them established their observations on the basis of empirical measurements, and so their mathematics lacked any apparent structure; it consisted merely of prescriptions of how to obtain measurements of certain spatial configurations.

*Constant versus varying referential reality.*

In constructing their geometry, as is depicted in Euclid’s *Elements*, the Greeks had only one model in mind—that of imageries of idealized physical reality. This is supported by the fact that the Greeks, in fact, strove to describe their primitive terms (e.g., “a point is that which has no parts”), which indicates that their sole imagery was that of physical space. Hartshorne (2000) points out that this way of thinking did not start with Euclid. About one hundred years before Euclid, Plato spoke of the geometers:

Although they make use of the visible forms and reason about them, they are not thinking of these, but of the ideals which they resemble; not the figures which they draw, but of the absolute square and the absolute diameter, and so on ... (*The Republic*, Book IV).

From the vantage point of modern mathematics, neither the primitive terms nor the primary propositions in Greek mathematics were variables, but constants referring to a single spatial model (Klein, 1968; Wilder, 1967)—as is expressed in the ideal world of Plato’s philosophy. This ultimate bond to a real-world context had an impact on the Greeks’ proof scheme. Specifically, in his proofs Euclid often uses arguments that are not logical consequences of his initial assumptions but are rooted in humans’ intuitive physical experience (the method of superposition, which allows one to move one triangle so that it lies on top of another triangle, is an example). In this respect, while proving in Greek mathematics was governed by the deductive proof scheme and can be characterized as an axiomatic proof scheme, it is different from that of contemporary mathematics where every assertion must be, in principle, derivable from clearly stated assumptions. Wilder (1967) points to a crucial difference between Greek mathematics and modern mathematics: Namely, in modern mathematics the primitive terms are treated as variables, not just undefined; they are free of any referent—real or imagined.

In Greek mathematics, on the other hand, they are undefined terms referring to humans' idealized physical reality. Treating primitive terms as undefined is fundamentally different from treating them as variables. Wilder quotes Boole (from 1847) to stress this difference:

The validity of the processes of analysis does not depend upon the interpretation of the symbols which are employed, but solely upon the laws of their combination. Every system of interpretations which does not affect the truth of the relations supposed, is equally admissible... (in Wilder, 1967, p. 116).

Despite the monumental conceptual difference in the referential imageries between Greek mathematics and modern mathematics, the essential condition in applying deductive reasoning in both is the existence of primary terms and primary propositions (axioms).

#### *Content versus form.*

The deductive reasoning—or the axiomatic method, as it often is called—of the Greeks dominated the mathematics of the western world until the late 19<sup>th</sup> century and in essence is still intact today. The difference in referential imageries between Greek mathematics and modern mathematics entails another critical difference: In Greek mathematics proof is valid by virtue of its *content*, not its *form*. Since the Greeks' concern was about relationships among entities in their physical space, the *form* of the proof cannot be completely detached from the *content* of that space. In this conception, one would not consider, for example, the question of whether there exists a consistent geometric model in which the parallel postulate does not hold. In modern mathematics, on the other hand, proof is valid by virtue of its form alone.

In *Greek Mathematical Thought and the Origin of Algebra*, Klein (1968) adds clarity to this distinction. In Greek science concepts are formed in continual dependence on their “natural” foundations, and their scientific meaning is abstracted from “natural,” pre-scientific experience. In modern science, on the other hand, what is intended by the concept is not an object of immediate insight. Rather, it is an object whose scientific meaning can be determined only by its connection to other concepts, by the total edifice to which it belongs, and by its function within this edifice. This difference accounts, for example, for the fact that the Greeks, while interested in idealized entities, were highly selective in their choice of these entities. For instance, what we now call transcendental numbers, like the number  $\pi$ , disturbed the Greeks. “They were unhappy that their ideas of perfection in geometry and arithmetic seemed to be challenged by the existence of such a number” (Ginsburg & Oppen, 1969, p. 214-215). In modern mathematics, on the other hand, entities can be quite arbitrary. Upon encountering transcendental numbers modern mathematics incorporated them into a mathematical structure, that of the real numbers, determined by a set of axioms.

It should be highlighted that the idea that the objects are determined by a set of axioms was a revolutionary way of thinking in the development of mathematics. An important manifestation of this revolution is the distinction between Euclid's *Elements* and Hilbert's *Grundlagen*. While the *Elements* is restricted to a single interpretation—namely that its content is a presumed description of human spatial realization—the *Grundlagen* is open to different possible realizations, such as Euclidean space, the surface of a half-sphere, ordered pairs and triples of real numbers, and so forth, including the interpretation that the axioms are meaningless formulas. In other words, the

*Grundlagen* characterizes a structure that fits different models. 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. To reflect this fundamental conceptual difference, we refer to the Greeks' proving means as the *Greek axiomatic proof scheme* and to the proving means of modern mathematics as the *modern axiomatic proof scheme*.

*Operation versus results of an operation.*

Klein (1968) argues that the revival and assimilation of Greek mathematics during the 16<sup>th</sup> century resulted in fundamental conceptual changes that ultimately defined modern mathematics. This conceptual transformation culminated in Vieta's work on symbolic algebra, where the distinction between modern mathematics and Greek mathematics began to crystallize.

The use of symbols in the modern sense is due mainly to Vieta, who was followed in this effort by Descartes and Leibniz. Until then, mathematics had evolved for at least three millennia with hardly any symbols. Symbolic notation was the key to a method of demonstration, and Leibniz—more than anyone else in his generation—advanced the role of symbolism in the process of mathematical proof. He was the first to conceive of proof as a sequence of sentences beginning with identities and proceeding by a finite number of steps of logic and rules of *definitional substitution*—by virtue of symbolic notation—until the theorem was proved.

The works of Vieta that led to the creation of algebra and that of Descartes that led the creation of analytic geometry illustrate another important difference between Greek mathematics and modern mathematics: the focus on “operation” versus “results of an operation.” Ginsburg and Oppen (1969) discuss the distinction as described by Boutroux, a mathematician, who analyzed the evolution of mathematical thought. According to Boutroux, the Greeks restricted their attention to attributes of spatial configurations but paid no attention to the operations underlying them. As an example, Ginsburg and Oppen mention the problems of bisecting and trisecting an angle by means of straightedge-and-compass only. As is well known, the Greeks offered a simple way to solve the bisecting problem but the solution of the trisecting problem had to wait two millennia, until Galois, a French mathematician, solved it in the 19<sup>th</sup> century. The Greeks attended to geometrical objects (rectilinear angles in this case) by investigating their attributes—whether an angle can be bisected or trisected, for example. The 19<sup>th</sup> century mathematics, on the other hand, investigated the operation themselves—their algebraic representations and structures. Specifically, the standard Euclidean constructions using only a compass and straightedge were translated in terms of constructibility of the real numbers (i.e., a real number is constructible if its absolute value is the distance between two constructible points). This translation leads to an important observation about the structure of constructible numbers; namely, the constructible numbers form a subfield of the field of the real numbers. A deeper investigation into the theory of fields led to a proof of the impossibility of certain geometric constructions, including the impossibility of finding a single method of construction for trisecting any given angle with the classical tools. More importantly, with this theory one can understand why certain constructions are possible whereas others are not. The Greeks had no means to build such an understanding. They did not attend to

the nature of the operations underlying the Euclidean construction, and hence were unable to understand the difference between bisecting an angle and trisecting an angle and why they were able perform the former but not the latter.

Thus, not until the 17<sup>th</sup> century with the invention of analytic geometry and algebra did mathematicians begin to shift their attention from the result of mathematical operations to the operations themselves. By means of analytic geometry, mathematicians realized that all Euclidean geometry problems can be solved by a single approach, that of reducing the problems to equations and applying algebraic techniques to solve them. Euclidean straightedge-and-compass constructions were understood to be equivalent to equations, and hence the solvability of a Euclidean problem became equivalent to the solvability of its corresponding equation(s).

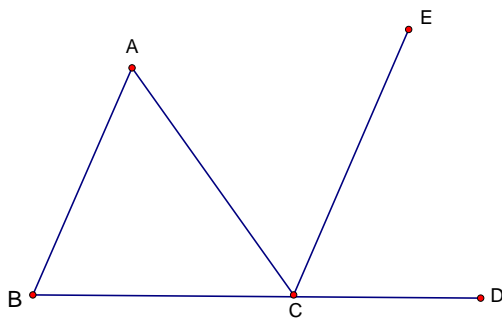
*Cause versus reason.*

Consider Proposition I.32 and its proof in Euclid's *Elements* (slightly adapted from Heath, 1956, pp. 316-317):

In any triangle, if one of the sides is produced, then the exterior angle equals the sum of the two interior and opposite angles, and the sum of the three interior angles of the triangle equals two right angles.

*Proof:*

Let  $ABC$  be a triangle, and let one side of it  $BC$  be produced to  $D$ .



I say that the exterior angle  $ACD$  equals the sum of the two interior and opposite angles  $CAB$  and  $ABC$ , and the sum of the three interior angles of the triangle  $ABC$ ,  $BCA$ , and  $CAB$  equals two right angles.

Draw  $CE$  through the point  $C$  parallel to the straight line  $AB$  (by Proposition I.31).

Since  $AB$  is parallel to  $CE$ , and  $AC$  falls upon them, therefore the alternate angles  $BAC$  and  $ACE$  equal one another (by Proposition I.29).

Again, since  $AB$  is parallel to  $CE$ , and the straight line  $BD$  falls upon them, therefore the exterior angle  $ECD$  equals the interior and opposite angle  $ABC$  (by Proposition I.29).

But the angle  $ACE$  was also proved equal to the angle  $BAC$ . Therefore the whole angle  $ACD$  equals the sum of the two interior and opposite angles  $BAC$  and  $ABC$ . Add the angle  $ACB$  to each. Then the sum of the angles  $ACD$  and  $ACB$  equals the sum of the three angles  $ABC$ ,  $BCA$ , and  $CAB$  (by Common Notion 2).

But the sum of the angles  $ACD$  and  $ACB$  equals two right angles. Therefore the sum of the angles  $ABC$ ,  $BCA$ , and  $CAB$  also equals two right angles (by Proposition I.13 and Common Notion 1).

This proof appeals to two facts, one about the auxiliary segment  $CE$  and the other about the external angle  $ACD$ . Note that the property holds whether or not the segment  $CE$  is produced and the angle  $ACD$  considered. One might then raise the question, what is the true cause of the property proved? This question was a center of debate during the 16<sup>th</sup>-17<sup>th</sup> centuries about whether mathematics is a *science*. Philosophers of this period, according to Mancosu (1996), used this proof to demonstrate their argument that mathematics is not a perfect science because “implication” in mathematics is a mere logical consequence rather than a demonstration of the *cause* of the conclusion. Their argument was based on the Aristotelian definition of science, according to which one does not understand something until he or she has grasped the why of it. “We suppose ourselves to possess unqualified scientific knowledge of a thing, . . . , when we think that we know the cause on which the fact depends as the cause of the fact and of no other . . .” (Aristotle, p. 111-112).

Mathematical statements of the form “A if and only if B” provided an additional argument against the scientificness (i.e., causal nature) of mathematics, for—these philosophers claimed—if mathematical proof were scientific (i.e., causal), then such a statement would entail that A is the cause of B and B is the cause of A, which implies A is the cause of itself—an absurdity.

This position entailed rejection of proof by contradiction, for such a proof does not demonstrate the cause of the property that is being argued. When a statement “A implies B” is proved by showing how not B (and A) leads logically to an absurdity, one does not learn anything about the causality relationship between A and B. Nor does one gain any insight into how the result was obtained. Consequently, proof by exhaustion (e.g., Archimedes’ known method of proof for calculating volume, area, and parameter of different objects), which is necessarily based on proof by contradiction, also was unsatisfactory to many mathematicians of the 16th and 17th centuries. They argued that the ancients, who broadly used proof by exhaustion to avoid explicit use of infinity, failed to convey their methods of discovery.

Not all philosophers of the time held this position. According to Mancosu (1996), Barozzi, for example, argued that some parts of mathematics are more scientific (causal) than others; but that a proof by contradiction is not a causal proof, and therefore it should be eliminated from mathematics. Others, like Barrow, argued that all mathematics proofs are causal, including proof by contradiction:

It seems to me . . . that Demonstrations, though some do outdo others in Brevity, Elegance, Proximity to their first Principles, and the like Excellencies, yet are all alike in Evidence, Certitude, Necessity, and the essential Connection and mutual Dependence of the Terms one with another. Lastly, that Mathematical Ratiocinations are the most perfect Demonstrations. (Quoted in Mancosu, 1996, p. 23)

Of particular interest is the position held by Rivaltus on the issue of causal proof in mathematics. Mancosu (1996) illustrates this position by Rivaltus’ commentary on Archimedes proof for the theorem that the area of the surface of the sphere is four times

the area of a great circle of the sphere. In this proof, Archimedes inscribes and circumscribes the sphere with auxiliary solids to show that the surface of the sphere can be neither smaller nor greater than four times the great circle—a typical Archimedean proof by exhaustion, which necessarily involves proof by contradiction. So, there are two issues here: the use of proof by contradiction and the use of the auxiliary solids (as in the case of the proof of Proposition I.32). Each of these two features render, in the eyes of some philosophers of the time, Archimedes' proof non-causal. Rivaltus rejects this possibility on the basis of a distinction between “cause” and “reason”:

Ostensive demonstrations in mathematics are not considered more perfect than the ones by contradiction, since in these disciplines it is not made use of the cause of the thing, but of the cause of the knowledge of the thing. ... The figures drawn are not truly the *cause* [italics added] of that equality but are *reasons* [italics added] from which we know it. From whence it follows that whatever is more fit to knowledge is more appropriate to the mathematician. But we know more easily that absurdities are impossible, false and repugnant by reason that we know the true things. Indeed the truths are concealed and conversely the errors are obvious everywhere. ... Again it is to be observed that the Geometers do not make use of the cause of a thing, but of the cause from which the thing is known. Indeed it is sufficient to them to show the thing to be so and they do not enquire by which means it is so. (Rivaltus, 1615; quoted in Mancosu, 1996, pp. 26-27)

The Aristotelian theory of science and particularly its appeal to cause and effect manifested itself in another aspect of mathematical practice during the 17<sup>th</sup> century—that of the use of “genetic definitions.” These definitions appeal to the generation of mathematical magnitude by motion; Euclid's definition of a sphere as an object generated by the rotation of a semicircle around a segment taken as axis is an example. An example of a non-genetic definition is that of a circle defined as a set of all points in the plane that are equidistant from a given point in that plane. According to Mancosu (1996), although the insistence on the use of genetic definitions was not universal, important mathematicians of the time emphasized their importance because they were viewed as demonstrating cause, and, hence, conform to Aristotle's epistemological position of what constitutes a science. For example, Barrow stated that of all the possible ways of generating a magnitude, the most important is the method of local movements, and he uses motion as a fundamental concept in his work in geometry” (Mancosu, 1996, p. 96), and Hobbes and Spinoza, “emphasize the role of genetic definitions as the only causal definitions, thereby excluding the nongenetic definitions from the realm of science” (p. 99).

### *Motives*

The Greeks' motive for constructing the remarkable geometric edifice that we now call Euclidean Geometry, was their desire to create a consistent system that is free from paradoxes. Avoiding paradoxes constituted, in part, an intellectual need for the transition from Greek to modern mathematics. According to Wilder (1967, based, in part, on Freudenthal, 1962), against the customary view that attributes the change to the introduction of non-Euclidean geometry,

there was little evidence of excitement or even interest in the mathematical community regarding the work of Bolyai and Lobachevski, for at least 30 years



after its publication. ... The ... impact of the admissibility of non-Euclidean geometry into mathematics only promoted an evolution already under way (p. 116).

Hacking (1980) argues that the change in perspective occurred much before the late 19<sup>th</sup> century as is commonly held, but was pioneered by Leibniz. In fact, Hacking insists that Leibniz was aware of the conceptual change and even explained why it did not develop earlier:

Leibniz himself has a plausible explanation of why the concept of proof emerged at this time. [The idea that proof is independent of its content] is not to be expected when geometry is the standard of rigor. Geometrical demonstrations can appear to rely on their content. Their validity may seem to depend on facts about the very shapes under study, and whose actual construction is the aim of the traditional Euclidean theorems. (p. 170)

Here, too, Descartes' contribution was crucial: What brought about Leibniz' new way of understanding the concept of proof was Descartes' algebrization of geometry. By its nature algebra is a way of thinking where one can dispose of or reduce overt attention to content. Together with the recognition that the non-Euclidean geometries are just as consistent as the Euclidean, there seems to have been a general feeling that both the Euclidean and the non-Euclidean systems were still only candidates as sciences of space.

The change in proof scheme from Greek mathematics to modern mathematics was necessitated again by the attempt to establish a consistent foundation for mathematics—just as the Greeks did to construct a consistent system—one that is free from paradoxes. The mathematicians who launched major attempts to form a consistent mathematics, like Zermelo and Russell and Russell and Whitehead, were largely motivated by the desire to meet the crisis in the foundations of mathematics caused by contradictions, primarily those resulting from the Cantorian theory of sets:

(The) Greek situation, as the latter appears to us through the haze of centuries, is striking. In both situations, crises had developed which threatened the security of mathematics; and in both cases resort was taken to explicit axiomatic statement of the foundations upon which one hoped to build without fear of further charges of inconsistency. (Wilder, 1967, p. 117)

Wilder (1967) also notes that like the Greek mathematicians, the modern mathematicians who launched major attempts to form a consistent mathematics (e.g., Zermelo and Russell) had only one model in mind, albeit a different model. In the case of Zermelo, he focused on “a portion of set theory sufficient for ordinary mathematical purposes, yet carefully limited to avoid the known contradictions” (p. 117). Russell too, like Euclid and Zermelo, “had only one model in mind, and he aimed at making the conditions surrounding his system so stringent that there could be only one model” (p. 118).

Not everyone in the 19<sup>th</sup>-20<sup>th</sup> century gave his or her blessing to the new development in mathematics. Both Poincare and Weyl, for example, while acknowledging the legitimacy and accomplishment of the axiomatic method, were disturbed by the widespread preoccupation with it. They called for a return to problems of “mathematical substance” and argued that its role should be limited to giving precision to already created mathematical entities (Wilder, 1967). Nor were all mathematicians of the 17<sup>th</sup> century enthusiastic about the emergence of a new context of mathematics. The

modern notion of number is a case in point. The emergence of negative numbers raised questions as to the utility of symbols without a concrete referent and especially without a geometrical referent. How is it possible, for example, to subtract a greater quantity from a smaller one, where the mental image of “quantity” is nothing else but a physical amount or a spatial capacity? Moreover, how is it possible to understand such a statement as  $(-1)/1=1/(-1)$ , where the quantity 1 is larger than the quantity  $-1$ , and therefore, the division of 1 by  $-1$  must be greater than the division of  $-1$  by 1? (See Mancosu, 1996.)

#### *Why Are These Historical-Epistemological Factors Relevant?*

What does this brief history account tell us about the learning and teaching of the concept of proof? It is still an open question whether the development of a mathematical concept within an individual student or a community of students parallels the development of that concept in the history of mathematics, though cases of parallel developments have been documented (e.g., Sfard, 1995). If this is the case, one would expect that the path of development would vary from culture to culture. Are there common elements or phases to different paths of development across cultures? Did the development of the concept of proof in, for example, China and India follow a similar path to that of the Western world or was there a leap in time from using perceptual proof schemes to modern axiomatic proof schemes? What are some of the salient social and cultural aspects that might impact the trajectory of these paths? Independent of these particular questions, whose treatment goes beyond the scope and goals of this paper, considerations of historical accounts can evoke other research questions whose answers can potentially direct the development and implementation of instructional treatments of mathematical concepts and ideas. In what follows, we will discuss such questions about the concept of mathematical proof evoked by the historical account outlined above.

#### *Pre-Greek mathematics versus Greek mathematics.*

The Greeks had likely constructed their deductive mathematics on the basis of the mathematics of their predecessors, which was mainly empirical. This mere fact evokes a pedagogical question as to the role of the empirical proof schemes in the development of the deductive proof schemes. The empirical proof schemes are inevitable because natural, everyday thinking utilizes examples so much. Moreover, these schemes have value in the doing and the creating of mathematics. They are even indispensable in enriching one’s images by creating examples and non-examples, which, in turn, can help generate ideas and give insights. The question is how to help students utilize their existing proof schemes, largely empirical and external, to help develop deductive proof schemes? As a historian might ask what events—social, cultural, and intellectual—necessitated the transition from pre-Greek mathematics to Greek mathematics, a mathematics educator should ask what instructional interventions can bring students to see a need to refine and alter their existing external and empirical proof schemes into deductive proof schemes.

As was indicated, while pre-Greek mathematics was concerned merely with *actual* physical entities, and its proving was governed by the empirical proof schemes, Greek mathematics dealt with idealizations of spatial and quantitative realities and its proving was deductive. There seems to be a cognitive and epistemological dependency between the nature of the entities considered and the nature of proving applied. To what

extent and in what ways is the nature of the entities intertwined with the nature of proving? For example, students' ability to construct an image of a point as a dimensionless geometric entity might impact their ability to develop the Greek axiomatic proof scheme, and vice versa. As far as we know, this interdependency has not been explicitly addressed and its implications for instruction have not been considered.

The motive for the Greeks' construction of their geometric edifice, according to the historical account presented earlier, was their desire to create a consistent system that was free from paradoxes such as those of Zeno. For example, to avoid Zeno's paradoxes the Greeks based their geometric proofs strictly in the context of static concepts. What does this tell us about how to help students see a need for the construction of a geometric structure, particularly that of Euclid? What is the cognitive or social mechanism by which deductive proving can be necessitated for the students?

#### *Greek mathematics versus modern mathematics.*

As was noted, in constructing their mathematics the Greeks had only one model in mind—that of imageries of idealized physical and quantitative realities. Further, it is these imageries that determined the axioms and postulates on which their geometric and arithmetic edifices stood. In contrast, in modern mathematics entities can be quite arbitrary and one's images of these entities are governed by a set of axioms. This sharp construct is best manifested by Euclid's *Elements*, on the one hand, and Hilbert's *Grundlagen*, on the other: While the former is restricted to a single interpretation, the *Grundlagen* is open to different realizations. Here too the question of intellectual necessity for the transition from Greek mathematics to modern mathematics is of profound pedagogical importance. Historians like Klein and others characterize this transition as revolutionary: It marks a monumental conceptual change in humans' mathematical ways of thinking. More research is needed to systematically document these difficulties to better understand their nature and their implications for instruction.

Another critical difference between Greek mathematics and modern mathematics has to do with *form* versus *content*. In Greek mathematics, the *form* of the proof could not be completely detached from the *content* of the spatial or quantitative context. In contrast, in modern mathematics proof is valid by virtue of its form alone. Here, too, we know of no studies that document systematically students' mathematical behaviors in relation to these fundamental characteristics of Greek mathematics—when students learn Euclidean geometry, for example, as compared to when they learn finite geometry as a case of modern mathematics.

#### *Post-Greek mathematics.*

Symbolic algebra, which began with Vieta's work, seems to have played a critical role in the transition from Greek mathematics to modern mathematics, particularly in relation to the reconceptualization of mathematical proof as a sequence of arguments valid by virtue of their form, not content. In the new concept of proof, one would begin with identities and by virtue of rules of symbolic definitional substitutions proceed through a finite number of steps until the theorem is proved. With symbolic algebra, mathematicians shifted their attention from results of operations (e.g., whether and how an angle can be bisected or trisected) to the operations themselves (e.g., the underlying difference between bisecting and trisecting an angle). A critical outcome of this shift was

the discovery that all Euclidean geometry problems can be solved by a single approach, that of reducing the problems to equations and applying algebraic techniques to solve them.

The role of symbolic algebra in the reconceptualization of mathematics in general and of proof in particular raises a critical question about the role of symbolic manipulation skills in students' conceptual development of mathematics, in general, and of proof schemes, in particular. Can students develop the modern conception of proof without computational fluency? And in view of the increasing use of electronic technologies in schools, particularly computer algebra systems, one should also ask: Might these tools deprive students of—or, alternatively, provide students with—the opportunity to develop algebraic manipulation skills that might be needed for the development of advanced conception of proof? In addressing this question, it is necessary, we believe, to distinguish between two kinds of symbolic proof schemes: non-referential and referential. As we discussed earlier, in the former scheme, neither the symbols nor the operations one performs on them represent a quantitative reality for the students. Rather, students think of symbols and algebraic operations as if they possess a life of their own without reference to their functional or quantitative meaning. By contrast, in the symbolic referential proof scheme, to prove or refute an assertion or to solve a problem, students learn to represent the statement algebraically and perform symbol manipulations on the resulting expressions. The intention in these symbolic representations and manipulations is to derive relevant information that deepens one's understanding of the statement, and that can potentially lead her or him to a proof or refutation of the assertion or to a solution of the problem. In such an activity, one does not necessarily form referential representations for each of the intermediate expressions and relations that occur in the symbolic manipulation process, but has the ability to attempt to do so in any stage in the process. It is only in critical stages—viewed as such by the person who is carrying out the process—that one forms, or attempts to form, such representations. This ability is potential rather than actual because in many cases the attempt to form quantitative representations may not be successful. Nevertheless, a significant feature of the referential symbolic proof scheme, which is absent from the non-referential symbolic proof scheme, is that one possesses the ability to pause at will to probe into the meaning (quantitative or geometric, for example) of the symbols.

Together with the emergence of symbolic algebra, a new conception of mathematical entity, particularly that of number, began to emerge. A mathematical entity, in this conception, is not necessarily dependent on its “natural” pre-scientific experience but on its connection to other entities within a structure and its function within that structure. For example, while the Greeks were highly selective in their choice of numbers—they rejected irrational numbers, for example—the post-Greek mathematicians began to accept them. This conceptual change was not without difficulty. For example, some mathematicians of the 17th century rejected the utility of negative numbers, which they viewed as symbols without real experiential referents. The conceptual attachment to a context—whether it is the context of intuitive Euclidean space or that of  $R^n$ —was dubbed the *contextual proof scheme*. In this scheme, general statements, intended for varying realities, are interpreted and proved in terms of a restricted context. The question of the developmental inevitability of the contextual proof scheme has not been fully addressed in mathematics education research. Some evidence exists to indicate that even

students in an advanced stage in their mathematical education have not developed this scheme. For example, it has been shown that many mathematics majors enrolled in advanced geometry courses have major difficulties dealing with any geometric structure except the one corresponding to their spatial imageries, and that mathematics majors enrolled in linear algebra courses interpret and justify general assertions about entities in a general vector space in terms of  $R^2$  and  $R^3$  entities (Harel & Sowder, 1989, Harel, 1999). Such findings have major curricular implications. For example, they raise major doubts as to the wisdom of the practice of starting off college geometry courses with finite geometries or of introducing general vector spaces in the first course of linear algebra.

The debate among philosophers during the Renaissance about the scientificness of mathematics and the mathematics practice that ensued is of particular pedagogical significance. As we have outlined, the question was whether the mathematical practice in which “implication” is a mere logical consequence, rather than a demonstration of the cause of the conclusion, is scientifically acceptable. This, in turn, raises questions about the acceptability of proof by contradiction and proof by exhaustion. Were these issues of marginal concern to the mathematicians of the 16<sup>th</sup> and 17<sup>th</sup> centuries, or had they been significantly affected by it? To what extent did the practice of mathematics in the 16<sup>th</sup> and 17<sup>th</sup> centuries reflect global epistemological positions that can be traced back to Aristotle's specifications for perfect science? These are important questions, if we are to draw a parallel between the individual's epistemology of mathematics and that of the community. As noted by Mancosu (1996), this debate had a deep and profound impact on the practice of mathematics during the 15<sup>th</sup> to 18<sup>th</sup> centuries. For example, the practices of Cavalieri, Guldin, Descartes, and Wallis reflected a deep concern with these issues by, for example, explicitly avoiding proofs by contradiction in order to conform to the Aristotelian position on what constitutes perfect science. This history shows that the modern conception of proof was born out of an intellectual struggle—a struggle in which Aristotelian causality seems to have played a significant role. Is it possible that the development of students' conception of proof includes some of these epistemological obstacles (in the sense of Brousseau, 1997)—obstacles that may be unavoidable, for they are inherent to the meaning of concepts in relation to humans' current schemes?

We conclude this section on the relevance of history and philosophy of mathematics to the learning and teaching of proof with questions pertaining to the idea of “genetic definitions”—mathematical definitions that utilize motion to generate magnitudes—from the 17<sup>th</sup> century. As we have indicated, the use of such definitions was viewed by some important mathematicians of the time to conform to the Aristotelian epistemological position on the centrality of causality in science. Can this account for the positive impact that dynamic geometry environments might have on advancing students' proof schemes? What is exactly the conceptual basis for the relationship between motion and causal proofs? In a later section we will report on several studies that have examined this effect.

### *Functions of Proof*

Earlier, in the second section, we described a portion of our taxonomy of proof schemes. The discussion that followed brought up several other schemes that emerged in the history of mathematics. In the rest of this section, we depict all the schemes

mentioned in this paper (Table 1) and discuss their functions within mathematics. This list is not complete; we only depict those schemes that are needed for the discussion in this paper (for the complete taxonomy, see Harel & Sowder, 1998).

Table 1

*Proof Schemes*

<u>External Conviction</u>	<u>Empirical</u>	<u>Deductive</u>
Authoritative	Inductive	Transformational
Ritual	Perceptual	Causality
Non-referential		Greek axiomatic
		Modern axiomatic

As we indicated earlier (see the second section), de Villiers (1999) built on the work of others scholars—particularly Hanna (1990), Balacheff (1988), Bell (1976), Hersh (1993)—to address important questions about the role of proof. Specifically, what different functions does proof have within mathematics itself and how can these functions be effectively utilized in the classroom to make proof a more meaningful activity? de Villiers suggests that mathematical proof has six not mutually exclusive roles:

- verification
- explanation
- discovery
- systematization
- intellectual challenge
- communication

In what follows, we will show that all of these functions but one (intellectual challenge) are describable in terms of the proof scheme construct. Some of the proof schemes used to interpret these functions appeared in the taxonomy presented above. A description of each—with our additions and modifications—follows.

*Verification* refers to the role of proof as a means to demonstrate the truth of an assertion according to a predetermined set of rules of logic and premises—the *axiomatic proof scheme*.

*Explanation* is different from verification in that for a mathematician it is usually insufficient to know only that a statement is true. He or she is likely to seek insight into why the assertion is true. We referred to this as the *causality proof scheme*.

*Discovery* refers to the situations where through the process of proving, new results may be discovered. For example, one might realize that some of the statement conditions can be relaxed, thereby generalizing the statement to a larger class of cases. Or, conversely, through the proving process, one might discover counterexamples to the assertion, which, in turn, would lead to a refinement of the assertion by adding necessary restrictions that would eliminate counterexamples. Lakatos' (1976) thought experiment on the proof of Euler's theorem for polyhedra best illustrates this process. In some cases one may ask whether a certain axiom is needed to establish a certain result, or what form the result would have if a certain axiom is omitted. We considered this as a case of the *axiomatizing proof scheme*.

*Systematization* refers to the presentation of verifications in organized forms, where each result is derived sequentially from previously established results, definitions,

axioms, and primary terms. This too is a case of the axiomatic proof scheme. The difference between systematization and verification is in the extent of formality.

*Communication* refers to the social interaction about the meaning, validity, and importance of the mathematical knowledge offered by the proof produced.

Communication can be viewed in the context of the two subprocesses that define proving: ascertaining and persuading.

*Intellectual challenge* refers to the mental state of self-realization and fulfillment one can derive from constructing a proof. As we mentioned earlier, this role does not correspond to any of our proof schemes.

With the notion of proof scheme as an organizing concept—appended with these functions—we will now present selected findings reported in the literature that pertain to students' conceptions of proof. Of course, we are unable to describe without speculation most of these findings in terms of the proof scheme construct, because these studies had not been conceptualized or designed with our proof scheme construct in mind. However, as we will see, much can be said about these research findings in relation to proof schemes.

#### Cognitive Factors: A Literature Review of Status Studies with Interpretations in Terms of Students' Proof Schemes

Examining student performance in an area of mathematics is a natural first step in judging whether curricular and teaching efforts in that area seem to be adequate and not to require any special attention. Hence, we first look at several status studies of proof performance. We interpret the results in terms of the (apparent) proof schemes involved in the results, following the proof schemes focus of this chapter. The first two parts of this section will report, respectively, on precollege and college students' conceptions of proof. The findings provide evidence that the pervasive proof schemes among the two populations of students are those belonging to the external conviction proof scheme class and the empirical proof scheme class.

With the deductive proof scheme playing a significant role in mathematics, there is also relevant information from examining performance in logical inferences—a critical characteristic of this scheme. Any description of mathematical proof, the ultimate in justification, will include some mention of logical inferencing ability. Someone not versed in, for example, the mathematical meaning of an “if... then...” statement, or who is not comfortable with reasoning patterns like modus tollens, or who does not understand quantified statements and the role of a counterexample, or who cannot deal correctly at some stage with negating “...and...” and “...or...” statements, or who cannot explain the logic of an indirect proof, surely cannot carry out or even understand many mathematical proofs (including disproofs). Hence logical inferencing ability is a basic tool for the process of proving in mathematics and likely enters also into many justifications of a less sophisticated sort. But logic is central to the deductive proof schemes. For example, the transformational proof scheme, which constitutes the essence of the proving process in mathematics and is expected to develop with at least college-bound students and mathematics major students, should be present in students' mathematical behavior. As we have discussed earlier, “logical inference” is one of the three essential characteristics of this scheme. A third part of this section will include findings reported in the literature that seem germane to the characteristics of these deductive proof schemes.

*Status Studies: Precollege Students**NAEP studies*

A first place to look for data is the periodic National Assessment of Educational Progress in the United States [NAEP], involving, typically, students at ages 9, 13, and 17 (and now reported by grade: 4, 8, and 12). The wide geographic sampling and the large sample sizes, usually many thousands of students, in a NAEP indicates the U.S. picture, which can be further checked with smaller studies and contrasted with studies outside the U.S. For our purposes, however, one limitation of the NAEP has been that only a few items on proof or logic can be included, because of the scope of the tests. Another limitation is that since all the NAEP questions considered here are multiple-choice items, it is difficult to pinpoint the actual reasoning students employed in answering them.

In the planning stage, the first NAEP mathematics assessment (1972-1973) included mathematical proof and logic as objectives to be evaluated, although only a few items tested these areas (Carpenter, Coburn, Reys, & Wilson, 1978, p. 10). Later NAEPs used different designs for setting objectives to be tested. The fourth NAEP (1985-1986), for example, included items testing “mathematical methods,” with a few intended to include “a general understanding of the nature of proof and axiomatic systems, and logic” (Carpenter, 1989, p. 3). Analysts of the results concluded that “most 11<sup>th</sup>-grade students demonstrated little understanding of the nature and methods of mathematical argumentation and proof” (Silver & Carpenter, 1989, p. 11), citing results on items requiring the recognition of counterexamples (with success rates of 31%-39%—see Figure 1 for two items), on items testing understanding of the terms “axiom” and “theorem” (fewer than one-fourth and about half correct, respectively), and on items dealing with an undisclosed but “straightforward” item on indirect proof (about one-third correct) and mathematical induction (similar results) ( pp. 17-18). Even when only those students who had taken geometry were considered, results were just slightly better than those for students who had taken mathematics only through first-year algebra. The overall performance led the analysts to conclude that “the generally poor performance on these items dealing with proof and proof-related methods suggests the extent to which students’ experiences in school mathematics, even for students in college-preparatory courses, may often fail to acquaint them with the fundamental nature and methods of the discipline” (p. 18).

A. Larry says that  $n^2 \geq n$  for all real numbers. Of the following, which value of  $n$  shows the statement to be FALSE?

-1/2 (23%)      0 (29%)      1/10 (39%)      1 (9%) (9% not responding)

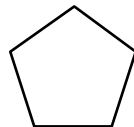
B. Jim says, “If a 4-sided figure has all equal sides, it is a square.” Which figure might be used to prove that Jim is wrong?



(15%)



(31%)



(24%)



(31%)

(46% not responding)



Fig. 1. Sample NAEP items involving recognition of a counterexample (percents of those responding are in parentheses (after Silver & Carpenter, 1989, p. 17). Note: Percents may total more than 100 because of rounding.

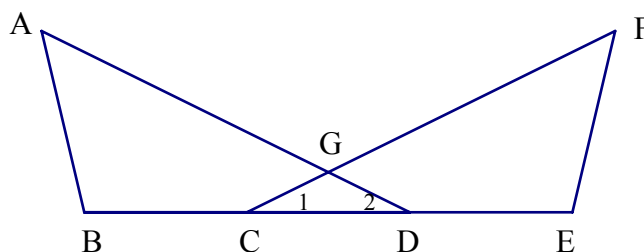
The test framework for the sixth NAEP (in 1992) did not focus specific attention on proof and logic items. But in commenting on performances on geometry items, the analysts (Strutchens & Blume, 1997, ch. 7) noted that many results could be explained by assuming that the students were basing their responses on the appearances in drawings—the perceptual proof scheme—rather than through reasoning. For example, when asked to choose from a set of triangles one that did not have a particular property, only 21% of the students made the correct choice (p. 180), but when asked whether four given statements applied to an illustrated construction of an angle bisector, 51% were correct on all four, leading the analysts to write, “Often a figure can foster correct reasoning....” (pp. 181-182).

Seventy-seven percent of the Grade 8 students, however, were successful at recognizing a counterexample to a false statement about quadrilaterals. It is interesting to note that the first NAEP (1972-1973) included a multiple-choice item on logic that involved recognizing the logical equivalent of “All good drivers are alert.” Only about one-half of the 17-year-olds, 11th graders, chose the correct alternative (“A person who is not alert is not a good driver”). The analysts for those NAEP results noted, however, “The logic exercises were probably answered as much by the semantic context of the problems as by any knowledge of logic” (Carpenter et al., 1978, pp. 126-127). This is in line with the earlier observation that proof schemes can vary from context to context.

Overall, then, the limited picture about proof understanding from the NAEPs is that at best only a small percent of high school students are equipped to deal effectively with the deductive proof schemes, with most apparently relying on the empirical proof schemes.

#### *Other status studies in the U. S.*

Studies in which justification and proof have been foci are perhaps more telling than the limited messages from the NAEPs. Some have been large-scale studies and hence are particularly significant. For example, Senk’s (1985) study involved 1520 U.S. geometry students in 74 classes in 11 high schools in 5 states within a month of the end of the course. After items in which the student was asked to supply missing reasons or statements, each student was asked to give four proofs, with the first two requiring only one deduction beyond those from the given information, as in the example in Figure 2 (72% correct). But only 32% could prove the textbook theorem, The diagonals of a rectangle are congruent.



Given:  $\overline{BD} \cong \overline{EC}$   
 $\angle 1 \cong \angle 2$   
 $\angle B \cong \angle E$   
 Prove:  $\overline{AB} \cong \overline{EF}$

Fig. 2. A sample geometry proof item from Senk's study (1985, p. 451).

The overall results were dismaying for the course in the U. S. in which deductive proof schemes should be expected to develop: "[The] data suggest that approximately 30 percent of the students in full-year geometry courses that teach proof reach a 75-percent mastery level in proof writing...29 percent of the sample could not write a single valid proof" (Senk, 1985, p. 453). At the time, about half of high school graduates took a course in geometry, so it is interesting to speculate about what more recent performances might be, when roughly 80% of U. S. high schoolers take geometry (U. S. Department of Education, 2000, p. 122).

Other, smaller studies in the U.S. have also involved students in geometry or later courses, since virtually all formal work with proof has traditionally been introduced in the geometry course (9<sup>th</sup> or 10<sup>th</sup> year is most common). But, for example, the interviewees in Tinto's (1988) study felt that proof was used only to verify facts that they already knew—an antithesis to the discovery or explanation functions of proof discussed earlier.

Thompson's study (1991) is of special interest since her subjects were all advanced students taking the last course of a curriculum targeting university-bound students and emphasizing reasoning and proof as a major strand. Yet Thompson expressed concern about the number of students who "proved" a statement by providing a specific example—a manifestation of the inductive proof scheme, and only about one-third of her subjects could find a counterexample to a number theory statement (For all integers  $a$  and  $b$ , if  $a^2$  is divisible by  $b$  then  $a$  is divisible by  $b$ ), in a "prove or disprove" context. Thompson also referred to the "enormous difficulty that students had with indirect proof" (1991, p. 23), with only 3% able to complete one indirect proof (that the sum of a rational number and an irrational number is an irrational number). Difficulties with indirect proof could well be related to the earlier discussion of the causal proof scheme.

Knuth, Slaughter, Choppin, and Sutherland (2002) found that 70% of roughly 350 students in grades 6-8 used examples (the empirical proof scheme) in justifying the truth of two statements (show that the sum of two consecutive numbers is always an odd number; show that when you add any two even numbers, your answer is always even, p. 1696). Only a few students attempted general arguments. On a more positive note, across grade levels, the students did show an increasing sensitivity to adhering to a given definition.

Since the focus of these studies was proof performance, they provide even more striking evidence than the NAEP studies did, that most U. S. students, even those in college-preparatory programs, do not seem to utilize deductive proof schemes.

*Status studies outside the U. S.*

Another large-scale study of student performance in justification and proof shows that weak performance and reliance on less mature proof schemes is not solely a U.S. phenomenon. Healy and Hoyles (1998, 2000) conducted a study involving 2459 English and Welsh 14-15 year-olds (finishing ninth graders in 94 classes in 90 schools). That the students were in the top 20%-25% on a national test is noteworthy, as is the fact that the English-Welsh national curriculum gives attention to conjecturing and explaining/justifying conclusions and is also, in U.S. terms, integrated with attention to both algebra and geometry, so presumably the items were reasonable for those grade 9 students. The students were asked to describe what proof means and what it is for, and to judge given proofs, as well as to decide on the truth or falsity and to construct their own proofs for two algebra and two geometry items. Over a quarter of these able students had little or no sense of the purpose of proof or its meaning (2000, pp. 417-418). The average score on the constructed-proof items was less than half the maximum, with 14%-62% of the students, depending on the item, not even able to start a proof (1998, p. 2). Of those able to start a proof, 28%-56% could proceed only minimally (1998, p. 2). It is comforting that the students were better at selecting correct proofs, favoring general arguments and clear and explanatory arguments that they found convincing (1998, p. 3). When the stated purpose was to get the best mark, however, the students often felt that more formal—e.g., algebraic—arguments might be preferable to their first choices (1998, p. 3). We interpret this last finding as an indication of the authoritarian proof scheme and the ritual proof scheme, in that in the eyes of these students proof must have a certain appearance (ritual) as determined by the teacher (authority).

It is interesting that, in the large, the findings of this study are consistent with those of Coe and Ruthven (1994), who looked at the proof performances of a much smaller group of “advanced level” British students in university preparation schools, on a project for an end-of-course assessment. These students, like those in the Healy and Hoyles study, relied predominantly on examples and techniques for analyzing numerical data, but showed little feel for the purposes of proof. Sample items from Healy and Hoyles’ and Coe and Ruthven’s studies are presented in Figure 3.

Proof or disproof construction (Healy & Hoyles, 1998)	Prove whether the following statements are true or false. Write down your answer in the way that would get you the best mark you can.  1. When you add any 2 odd numbers, your answer is always even.  2. If $p$ and $q$ are any two odd numbers, $(p + q) \times (p - q)$ is always a multiple of 4.
“Starting point” item (Coe & Ruthven, 1994, p. 46)	Any four-digit number is rewritten with its digits arranged in ascending and descending order of size. The smaller is then subtracted from the larger and the process is repeated. Example: 7345 gives $7543 - 3457 = 4086$ 4086 gives $8640 - 0468 = 8172$ Then write $7543 \rightarrow 8640 \rightarrow 8721 \rightarrow \dots$ Complete this chain and then try again with other four-digit numbers. Now try with 3 or 5 or 6 or $n$ digits.

Fig. 3: Sample proof items (adapted from Healy & Hoyles (1998, 2000) and Coe & Ruthven (1994)).

An older study by Reynolds (reported in Lovell, 1971) also examined proof ideas held by British students. Most revealing were the responses of 153 6th formers (final year of secondary school) in mathematics, with about 20% endorsing 16 examples as establishing a given statement—a manifestation of the empirical proof scheme. On the other hand, about 70% of these 6th-formers could complete an indirect argument involving a geometric situation (to show the inequality of a pair of alternate interior angles in a drawing with non-parallel lines) that had been started (p. 75), in contrast to the very weak (3%) performance of Thompson’s (1991) U. S. students mentioned above.

Still other studies outside the U.S. also indicate difficulties with certain proof ideas. Galbraith (1981) studied the conceptions of proof held by 12-17-year-old Australian students, most of whom were 13-15 years old. He found that 28% of 130 responses indicated a lack of understanding of what a given counterexample told one about a given statement. On another problem, only 18% of his 73 students felt that one counterexample was sufficient to disprove a conjecture. In commenting on how many counterexamples are sufficient, one student even suggested “Sometimes one is enough and sometimes it isn’t” (p. 19). The confusion about the role of examples in direct proofs versus in indirect proofs versus in disproofs is consistent with our own findings (Harel & Sowder, 1998), a phenomenon we accounted for in terms of students’ empirical proof scheme. Many students in Galbraith’s study were guided by their mental pictures of what a geometric term suggested rather than a given definition, and relied on just one diagram without considering other possibilities (p. 23)—a clear manifestation of the perceptual proof scheme

Porteous (1986, 1990) gave questionnaires to about 400 British students, ages 11-16, and interviewed 50 of them. From the questionnaires, he found that more than 40% of the responses endorsed completely a generalization on the basis of examples only, and only about 10% then offered proofs on their own when asked to explain their decisions (1990, p. 591). Furthermore, 83% of those not offering a proof initially claimed to understand a given proof of an assertion, but only 61% of all the students, including the ones offering a proof on their own, were sure that the author was correct about the assertion after seeing a proof (1996, p. 8). From the students' reactions to given proofs, Porteous concluded that "It is only when a pupil devises a proof for himself that he is convinced of the truth of a statement. A proof provided by a teacher may have some effect, but it is small in comparison with a d(i)scovers proof...we need to encourage pupils to investigate relationships for themselves, in order to produce their own reasons, or proof, for general statements. This is not to say, of course, that empirical work has no real part to play in the learning process" (1996, pp. 21-22).

Williams (1980) interviewed 11th grade Canadian students in a college preparatory program and concluded that fewer than 30% showed a grasp of the meaning of mathematical proof, that about half of the students saw no need to prove a statement that they regarded as obvious, roughly 70% did not distinguish between inductive and deductive arguments, and fewer than 20% understood how indirect proof works.

Spanish researchers Recio and Godino (2001) conducted a study involving two groups ( $n = 429$  and  $n = 193$ ) of beginning university students. They found that about a third of the 429 students and less than a quarter of the 193 students could prove both of two elementary statements (the difference between the squares of consecutive natural numbers is odd and equal to the sum of the numbers, and the bisectors of adjacent supplementary angles are perpendicular). Also, fewer than half of the group of 429 were successful on each individual proof. Roughly 40% used empirical reasoning: "Empirical inductive [proof] schemes were the spontaneous type of argumentation in a high percentage of students when they were confronted with new problems, in which it was necessary to develop new proof strategies, different from the learned procedures" (p. 91).

In Israel, Fischbein and Kedem (1982) found that their high school students did not appear to understand that no further examples need be checked, once a proof was given, a finding confirmed by Vinner (1983), who also noted that many high school students (35%), even high attaining ones (39%), seem to regard a given proof as the method to examine and verify a later particular case.

Similar results were found in a recent Japanese study (cited in Fujita & Jones, 2003). The official curriculum of Japan calls for students to "understand the significance and methodology of proof." However, even though most 14 to 15 year old students are successful at proof writing, "around 70% cannot understand why proofs are needed" (p. 9).

Internationally, overall the most positive conclusion seems to be that the proof glass is not completely empty but that it is by no means even close to full. Even "good" performances may be tainted by little understanding or appreciation of the functions of proof. The prevalence of empirical proof schemes for most students seems to be international.

*Status Studies: University Students*

One might expect that university students would perform better than secondary students on proof activities. As the Recio and Godino (2001) study indicated, however, performance is disappointingly low in many entering university students. There are only a few formal studies that have examined university students' later conceptions of proof. In one example, up to 80% of prospective elementary school teachers derived the truth of a general statement from examples (Goetting, 1995), confirming a result by Martin and Harel (1989) about the pervasiveness of the empirical proof scheme among students. Martin and Harel had emphasized proof in their course for prospective elementary school teachers (who had had geometry in high school) and then asked them to judge whether particular arguments were mathematical proofs for a statement covered and proved in the course (If the sum of the digits of a whole number is divisible by 3, then the number is divisible by 3) or an unfamiliar one (If  $a$  divides  $b$ , and  $b$  divides  $c$ , then  $a$  divides  $c$ ). The arguments were either inductive (based on specific instances) or deductive (assertions via general statements). More than half of the 101 students considered inductive arguments to be mathematical proofs. As many as half of these college students also accepted a false deductive argument as being a mathematical proof, apparently reacting to the appearance of the argument—indicating the dominance of the ritual proof scheme—and over a third found both inductive and deductive arguments to be acceptable mathematical proofs. Morris (2002) echoed this last result, with 40% of her pre-service teachers affirming that both inductive and deductive arguments assured certainty and failing to distinguish between the two types of arguments.

But prospective elementary teachers are often not very sophisticated in mathematics, so one might expect majors in technical fields, especially mathematics majors, to perform much better in proof settings. Such is the case, but only for some students, judging from anecdotes from university mathematics faculty. Furthermore, there are studies that point out areas of weakness, usually with small numbers of students. For example, case studies show that although some students may possess the transformational proof scheme, others seem to enter—and leave—university with little or limited abilities with, and views of, proof, mostly demonstrating external and empirical proof schemes (Sowder & Harel, 2003).

For some university students, even recognizing whether a given argument constitutes a proof cannot be taken as a given. In interviews, Selden and Selden (2003) asked their eight mathematics majors to judge whether four given mathematical arguments were valid, with only one of the arguments valid. The students performed only at a 46% correct level with their first judgments but did improve to 81% correct by the end of the interview with their fourth judgments, with the Seldens crediting the improvement to the added experience and reflection during the interview. Although the study did not examine long-term effects, it suggests that repeated validating activities may be a valuable growth opportunity.

Knowledge about what factors are important in devising proofs is limited. In an expert-novice study, with four graduate students in mathematics as the experts and four undergraduates as the novices, Weber (2001) noticed that “an understanding of mathematical proof and a syntactic knowledge of the facts of a domain are not sufficient for one to be a competent theorem prover” (p. 107); experts have what Weber calls “strategic knowledge,” knowledge that undergraduates do not exhibit—knowledge of the

domain's proof techniques (p. 111), knowledge of which theorems are important and when they are useful (p. 112), and knowledge of when, and when not, to use strategies based on symbol manipulation rather than deeper knowledge (p. 113). In a similar vein, Raman's (2003) interviews of mathematics students and faculty suggested the importance of a "key idea"—"an heuristic idea which one can map to a formal proof with appropriate sense of rigor" (p. 323). She concluded, "For mathematicians, proof is essentially about key ideas; for many students, it is not" (p. 324). In expert-novice studies, one often does not know how the experts acquired their expertise (or whether there is a selection factor involved), but knowing what differences exist may give ideas for instructional emphases. But in questioning university mathematics faculty about university mathematics majors' proof understanding, a "twice is nice" theme emerged: Exposure to the same material twice allows the student, on the second exposure, to focus on proof methods (Sowder, 2004). Perhaps these second exposures are helpful in attaining other aspects of Weber's strategic knowledge and Raman's key ideas, and in growing in deductive proof schemes. Marty (1991) felt that his explicit attention to proof methods, rather than the common focus on new mathematical content, helped his college students succeed in later mathematics courses.

University students' distinctions among axioms, definitions, and theorems are not sharp. Vinner (1977), for example, found that only about half of a group of Berkeley sophomores and juniors in mathematics could correctly identify all of three statements about exponents as definitions, as opposed to theorems or laws or axioms. Since much of algebra focuses on algorithms and these students had first studied the material in high school, perhaps the lack of such distinctions is expected. Yet, when Brumfiel (1973) questioned a class of University of Michigan juniors and seniors, nearly all of whom had complete a university course in formal geometry, he found that their mastery of similar distinctions were shockingly deficient. For example, collectively the students could recall only one axiom (two points determine a line), about half called a definition of isosceles triangle a theorem, and all were certain that, for two (given) independent postulates about points and lines, one could be deduced from the other. In Israel, Linchevsky, Vinner, and Karsenty (1992) found that only about one-fourth of their university mathematics majors understood that it is possible to have alternate definitions for concepts. These studies speak to about-proof topics. But about-proof topics cannot be mastered without understanding the proof topics themselves. In this case the topics in question involve the meaning and role of axioms and definitions. These studies, then, suggest a weak, or even absent, axiomatic proof scheme among mathematics majors—an observation that is consistent with findings of our study with mathematics majors (Harel & Sowder, 1998).

### *Findings Germane to the Transformational Proof Scheme*

Recall that the deductive proof scheme class consists of two schemes: the transformational proof scheme and the axiomatic proof scheme. The former is a special case of, and a conceptual prerequisite for, the former. The research findings we have just reported paint a gloomy picture as to the quality of students' proof schemes. The failure of mathematics instruction to help many students—even college-bound students and mathematics majors—construct the transformational proof scheme is evidenced in these

reports. In this section we provide a closer look at the characteristics of the transformational proof scheme, particularly the deduction process characteristic.

As was defined earlier, the transformational proof scheme has three characteristics: *generality*, *operational thought*, and *logical inference*. As the language may suggest, our definition of the transformational proof scheme naturally draws one to Piaget's theory of intellectual developmental (for example, Inhelder & Piaget, 1958). Battista and Clements (1992) focused on Piagetian thought in their discussion of proof in an earlier handbook (pp. 439-440), identifying three levels of justification/proof, with the third level being marked by the student being "capable of formal, deductive reasoning based on *any* assumptions" (p. 439)—the generality characteristic—whereas at the second level, the student's reasoning has an empirical reality, with the student not realizing that an argument may cover all cases. The third level of thinking is often identified with Piaget's "formal operational" level—the operational thought characteristic. Although the nominal approximate age for the development of formal operational thought is early adolescence and there are different aspects of formal operational thought (e.g., proportional thinking, ability to control variables, propositional logic), studies suggest that many high school and college students have not reached the formal operational level in some areas important in mathematics. For example, Lawson, Karplus, and Adi (1978) gave tasks in several areas, including propositional logic (one such item is Task 1 in Figure 4), to students at grades 6, 8, 10, 12, and 13-14. They found "the lack of clear and substantial improvement with age for the propositional logic items," with percents correct by grade on the Task 1 item being 1.0%, 2.1%, 10.1%, 12.0%, and 16.2% (p. 469). A study of 16-year-old English students found that only 30% tested at a level consistent with formal operational thought (cited in Adey, 1999). It appears that it is risky to assume that logical reasoning at a high level develops automatically in the course of schooling, whether within or outside mathematical contexts.

Psychologists other than Piaget have examined logical thinking for many years, with especially fascinating work done during the last thirty years (for example, Evans, 1982; Rips, 1994; Wason & Johnson-Laird, 1972). More recently, Johnson-Laird (2001) has observed that "deductive reasoning is under intense investigation. The field is fast moving and controversial" (p. 441), with at least some of the controversy coming from different hypothesized mechanisms of reasoning for selected tasks. There does not, however, seem to be controversy about the following findings. Performance on logical inferences involving modus ponens is usually reasonably good, but performance on those tasks involving modus tollens is weak, as is a full understanding of inferences involving if-then statements. Students are too willing to use invalid inference patterns like affirming the consequent (i.e., "if  $p$  then  $q$ " and " $q$ " yielding " $p$ "). In general, humans do not seem to process negative statements as readily as affirmative ones, and context is an extremely important variable in performance on deductive reasoning tasks (Wason & Johnson-Laird, 1972). Figure 4, for example, gives logically isomorphic tasks on which the success rates of adults are strikingly different. Only 2 out of 24 (about 8%) of the respondents correctly identified the two cases that must be considered for Task 1, but 88% did so on Task 2 (reported in Wason & Johnson-Laird, 1972). Thus, our discussion here focuses primarily on findings and studies dealing with logic in mathematical settings, calling on known work from psychology when it seems particularly pertinent.



<p>Task 1. Given four envelopes with a letter on the front side and a number on the back, select just the envelopes definitely needed to be turned over to find out whether they violate the rule. Separate envelopes show on front: D and C, and on back: 5 and 4.</p> <p>Rule to test: If a letter has a D on one side, then it has a 5 on the other side. Percent correct (D, 4): 8%</p>	<p>Task 2. Given four envelopes with a space for a stamp on one side and sealed or not, select just the envelopes definitely need to be turned over to find out whether they violate the rule.</p> <p>Envelopes show (a) back of sealed envelope; (b) unsealed envelope with flap up; (c) front of an envelope with stamp; (d) front of an unstamped envelope.</p> <p>Rule to test: If a letter is sealed, then it has a 5 pence stamp on it. Percent correct (a, d): 88%</p>
---	---

Figure 4. Two logically isomorphic tasks with “abstract” [Task 1] and “concrete” [Task 2] contexts (after Wason & Johnson-Laird, 1972, pp. 191-192).

Although our focus is mainly on studies dealing directly with mathematics, one sobering study involving everyday contexts deserves special note, because some of the data come from prospective secondary school mathematics teachers. Using items from Eisenberg and McGinty (1974), Easterday and Galloway (1995) compared the performance of last-year university students planning to teach middle school or high school mathematics with those of 7-8<sup>th</sup> graders and 12<sup>th</sup> graders on a variety of reasoning tasks. The tasks were based on everyday contexts that should be non-suggestive (e.g., “If John is big, then Jane is big. John is big. Is Jane big? Yes/no/maybe”). In particular, on modus tollens tasks, the college students scored 47%, about the same as the 12<sup>th</sup> grade calculus students but 20% *less* than the 7-8<sup>th</sup> graders’ 67%. These particular 7-8<sup>th</sup> graders were studying geometry and may therefore have been exposed to logic, but an earlier comparison in 1986 with 7-8<sup>th</sup> graders in advanced sections had given a similar result, with the 7-8<sup>th</sup> graders scoring 62% and the college students only 40% on the modus tollens tasks (Easterday & Galloway, 1995, p. 433). The authors concluded, “College students are barely performing better than children whom they may one day teach” (p. 435), and neither group was by any means topping out in performance.

There seem to be only a few data on students’ abilities with specific ideas from logic and within mathematics. Mentioned earlier were the limited results from NAEP, and there the picture was clouded by the use of a non-mathematical context. The large-scale Longitudinal Proof Project in England has, however, looked at students’ performance on if-then statements involving number theory ideas (Hoyles & Kuchemann, 2002). Among their findings was that 62% of 14-year-olds thought that a given if-then statement (for example, if the product of two numbers is odd, then the sum of the numbers is even) “said the same thing” as its converse. However, the longitudinal nature of the study also allowed Hoyles and Kuchemann to find a 7% improvement on this item over that of the same students at age 13. (In passing, it may be worth noting that so many theorems, especially in geometry, do have true converses, so it is easy to see how

students may be insensitive to the logical difference between an if-then statement and its converse.)

Even advanced students have difficulty with quantifiers. Dubinsky and Yiparaki (2000) found that university students at various levels, including some in an abstract algebra course, had much greater trouble giving the mathematical meaning of a doubly quantified statement when the existential quantifier appeared before the universal quantifier (“There is a positive number  $b$  such that for every positive number  $a$   $b \leq a$ ”—19% correct) than when the quantifiers were reversed, universal before existential (“For every positive number  $a$  there is a positive number  $b$  such that  $b < a$ ”—59%). Selden and Selden (1995) noted that university mathematics students may have difficulties in even restating mathematical statements precisely, with their largely third- and fourth-year students often giving incorrect responses when quantifiers were involved. Thus, although there may be areas of apparent strength in the use of logic (for example, the use of modus ponens), there appear to be many areas of weakness as well, and at a wide gamut of levels of schooling.

An important point is that everyday usage of logical expressions may differ considerably from the precise usage in mathematics. Epp (2003) has summarized the differences in a compelling way, and O’Brien, Shapiro, and Reali (1971) have referred to “child’s logic” in describing some of the differences. For example, “or” in everyday usage is most often in the exclusive sense (“I’ll wear my sandals or my tennis shoes”), in contrast with the inclusive convention common in mathematics. An everyday if-then statement (for example, “If you finish your work, then you can watch the game”) often connotes what would be an if-and-only-if statement in mathematics and to many children seems to be an “and” statement. The disparities between everyday usages and mathematical usages are so marked that explicit instruction in logic as used in mathematics would seem to be necessary, with contrasts to the less precise everyday usages pointed out, yet, as Epp contends (2003), perhaps exploiting non-mathematical usages that do reflect the mathematically precise ones, as exemplars for the latter.

Another interesting disparity between everyday usage and mathematical usage is that of indirect proof. According to Freudenthal (1973) indirect proof is a very common activity. Seven to eight year old children used contradiction in game playing and checking conjectures (Reid & Dobbin, 1998). Antonini (2003) even found that indirect argumentations occurred spontaneously by students in his interviews with them about mathematical assertions. Yet research has shown that students experience difficulties with proof by contradiction in mathematics. Leron (1985), for example, observed that despite the simple and elegant form of certain proofs by contradiction, students experience what seem insurmountable difficulties. Lin, Lee, and Wu Yu (2003) see the ability to negate a statement as a prerequisite ability for succeeding at a proof by contradiction. They found that the difficulty levels of students’ negating a statement can be ordered decreasingly as negating statements without quantifiers, negating “some,” negating “all,” and negating “only one.”

In general, then, there are many weak spots in students’ likely grasp of the logical reasoning used in advanced proof schemes. Is it a chicken-egg question, or can logical thinking and proof performance grow together? Later in this section we summarize several studies in which explicit instruction in logical principles was incorporated into high school geometry courses.

### Instructional-Social-Cultural Factors: Evidence Pointing to Curriculum and Instruction

Students' instructional history is without question an important variable in studying their performance in justification/proof settings. Hoyles (1997), in particular, makes a compelling case that studies of advanced learners' performance are relatively meaningless without knowing what curricula (and what teaching emphases) they have received. Two important dimensions of instructional history are the curriculum and the teaching. We begin with studies of teaching and teachers, giving some evidence of the *delivered* curricula. Following this, we review a few studies that highlight the vast differences in curricula, primarily the *intended* curricula as evidenced by national or regional guides. The earlier report of several status studies of students' performance have given a glimpse of the *learned* curricula. We then attend to some studies that give evidence of what might be achieved through *revised* curricula and teaching, as delivered by teachers with a different perspective on teaching, with the section including a review of a few studies that show the possibilities offered by technology. We conclude with several studies that have focused explicitly on logic as a vital component for what we call the deductive proof schemes.

#### *Current Status*

##### *Teachers and teaching.*

The emphasis that teachers place on justification and proof no doubt plays an important role in shaping students' proof schemes. "You get what you teach" is a common aphorism, and the intended curriculum may differ markedly from the delivered curriculum, especially across classrooms. The essentiality of opportunity to learn must be recognized not only at the intended curriculum level but also in the teachers' enacted curriculum. Yet, a study involving 62 mathematics and science teachers in 18 high schools in six states allowed Porter (1993) to note that for the mathematics teachers, "On average, no instructional time is allocated to students learning to develop proofs, not even in geometry" (p. 4).

Results on items allied to the areas of justification and proof in the 1996 NAEP mathematics assessment indicate not only that students perform much worse on items requiring explanation and justification, but also that at grade 8, the students of teachers who devote more time to "developing reasoning and analytical ability to solve unique problems" have noticeably higher overall scores than other students (Silver & Kenney, 2000). The importance of the role of the teacher (and the curriculum) in fostering justification and proof is further highlighted by the videotape study in the Third International Mathematics and Science Study (TIMSS), in which 30 eighth-grade sessions in each of Germany, Japan, and the U.S. were analyzed. The report reaches the following conclusion.

The most striking finding in this review of 90 classes was the rarity of explicit mathematical reasoning in the classes. The almost total absence of explicit mathematical reasoning in Algebra and Before Algebra courses raises serious questions about the ways in which those subjects are taught. In order for these courses to help introduce students to mathematical ways of knowing, some of the logical foundations of mathematical knowledge should be explicit. Of course, the total absence of any instances of inductive or deductive reasoning in the analyzed

United States classes cries out for curriculum developers to address this aspect of learning mathematics. (Manaster, 1998, p. 803)

It should be clear that only an authoritarian proof scheme is likely to be fostered in these classrooms.

Despite Porter's (1993) finding of little attention to proof in high school, one would certainly expect more explicit attention to proof in the mathematics at those grade levels where proof most often is a conscious part of the curriculum. Senk (1985) noted, however, that there were consistent differences across schools in the geometry students' performances on her proof tasks. Tinto (1988) too noted that one of her four teachers seemed markedly different from the others in his approach to geometry. In their large-scale study of proof in British classrooms, Healy and Hoyles (1998) found that students who had been expected to write proofs and who had classes in which proof was taught as a separate topic performed somewhat better on proof items than other students. Thompson (1991), on the other hand, did not notice differences across her nine teachers and schools; her sample, however, included three private schools and two magnet schools and so was perhaps not representative. Overall, it appears that at least some of the deficiencies in students' acquisition of more sophisticated proof schemes may stem from the lack of opportunity to engage in proof-fostering activities, even in courses where one would expect much attention to proof.

The evidence from the status studies of university students' proof knowledge suggests that some, if not many, precollege teachers are unlikely to teach proof well, perhaps because their own grasp of proof was probably limited in college and may not have grown since then. Knuth (2002a) examined the conceptions of proof of 16 practicing secondary school mathematics teachers, most with backgrounds that would pass a face-validity test for knowledge of mathematics. In interview settings, Knuth asked the teachers to respond to general questions about proof (e.g., What purpose does proof serve in mathematics?), to evaluate given arguments (both proofs and non-proofs), and to identify the arguments that were most convincing. Although all of the teachers endorsed the verification role of proof, none mentioned the explanatory role of proof (see the section on the concept of proof scheme). Six of the 16 thought it might be possible to find contradictory evidence of a (non-specified) statement that had been proved. Four of the 16 tested a statement with a given, endorsed proof with further examples (cf. Fischbein & Kedem, 1982), even though all of the teachers eventually acknowledged that it would not be possible to find a counterexample. Even though the teachers collectively correctly identified 93% of the correct arguments as being proofs, over a third of the non-proof arguments were rated as being proofs! Ten of the 16 accepted the proof of the converse of one statement as a proof of the statement. Thirteen of the 16 teachers found arguments based on examples or visual presentations to be most convincing. Although Knuth felt that their responses may have been directed toward *personally*-convincing rather than *mathematically*-convincing, that mathematics teachers would be convinced by such arguments more than by a mathematical proof is significant, because it reveals an apparent dominance of the empirical proof schemes among the teachers. Knuth (2002b) further examined these teachers' ideas about proof in the context of school mathematics (versus the earlier just-in-mathematics). In view of the NCTM (2000) recommendation that reasoning and proof be considered fundamental aspects of the study of mathematics at all levels of study, it is disappointing that the teachers in Knuth's study "...tended to

view proof as an appropriate goal for the mathematics education of a minority of students” (2002b, p. 83), with 14 not considering proof to be a central concern in school mathematics. And, to repeat, none even explicitly mentioned the explanatory role of a proof, although seven did mention its verification role—that a proof shows why a statement is true. Thirteen regarded the development of logical thinking or reasoning as the primary role of proof in school mathematics. Yet, the teachers did not completely reject justifications, but were willing to rely on informal “proofs” (e.g., examples and drawings) to support results, a practice that may mislead students into thinking that such are acceptable mathematical “proofs,” and reinforcing the acceptability of their empirical proof schemes.

### *Curricula.*

That mathematics curricula differ in their treatments of proof is by no means a recent phenomenon. For example, in his study of students’ proof explanations, Bell (1976) found that proof is the topic that shows the greatest variation in approaches internationally. He noted that this variation can be attributed to the tension between the recognition among teachers that deduction is essential to mathematics but that only the most capable students develop a good understanding of it. There is evidence that this condition remains true today as well. For example, Fujita and Jones (2003) compared the textbook treatments of geometry in lower secondary schools in Japan and Scotland and concluded the following.

Our analysis indicates that...Japanese textbooks set out to develop students’ deductive reasoning skills though the explicit teaching of proof in geometry, whereas comparative UK [United Kingdom] textbooks tend, at this level, to concentrate on finding angles, measurement, drawing, and so on, coupled with a modicum of opportunities for conjecturing and inductive reasoning. (p. 1)

It is, perhaps, natural to expect great variation in the treatment of curricular topics within countries that do not have national curricular and educational guidelines. In the U. S., for example, how geometry, the primary locus of proof efforts until recently, should be handled has led to vastly different opinions and occasionally to different approaches or emphases in school geometry (cf., e.g., Hoffer, 1981; NCTM, 1973, 1987; Usiskin, 1987).

### *Examples of Curricula Aimed toward Enhancing Students’ Proof Schemes*

We do not claim that the studies touched on below use our language of proof schemes, but it will be clear that students who have had such treatments should possess different proof schemes from those that apparently result from traditional teaching. Of particular interest and importance, we think, are the feasibility studies that have involved children in elementary schools.

#### *Earlier feasibility studies.*

Perhaps because of the spirit of the “New Math” times, and the call of the Cambridge Conference Report to move toward more sophisticated mathematics earlier in the curriculum, some studies in the 1970s focused on proof with younger children. To gauge whether and when children seem able to cope with proof or proof-like tasks, Lester (1975) devised a computer-delivered deductive system in a game-like form, with well-

defined rules (i.e., postulates) and target configurations (i.e., theorems) to be attained by applying the rules (i.e., with proofs). Hence, the students were dealing implicitly with an axiomatic system.

Lester (1975) sampled 19 students from each of four groups, one group from grades 1-3, a second from grades 4-6, a third from grades 7-9, and the fourth from grades 10-12, gave them practice with the rules, and studied their performances on the target tasks. His grades 7-9 students performed as well as the students from grades 10-12, and the students from grades 4-6 solved about as many tasks as the older students, but took somewhat longer. Lester suggested that “even students in the upper elementary grades can be successful at mathematical activities that are closely related to proof” (p. 23). Indeed, King (1970, 1973) thoroughly developed a 17-day unit dealing with some elementary number theory results (e.g., a number which is a factor of two numbers is also a factor of their sum). He found that a group of 10 above-average sixth graders could reproduce the proofs initially developed with considerable teacher support (in contrast to a non-equivalent control group), but the evidence also suggested that the proofs were given from rote memory.

Fawcett’s study of the late 1930s deserves special mention. The title and sub-title give a good summary: *The Nature of Proof, A Description and Evaluation of Certain Procedures Used in a Senior High School to Develop an Understanding of the Nature of Proof*. Whatever the reason—World War II, the usual inertia in curriculum—this study seemed to have had little impact, even though it was reported in a yearbook of the NCTM (Fawcett, 1938/1995). His approach was surprisingly modern in tone. Fawcett’s summary includes the following.

The theorems [of geometry] are not important in themselves. It is the *method* by which they are established that is important, and in this study geometric theorems are used only for the purpose of illustrating this method. The procedures used are derived from four basic assumptions:

1. That a senior high school student has reasoned and reasoned accurately before he begins the study of demonstrative geometry.
2. That he should have the opportunity to reason about the subject matter of geometry in his own way.
3. That the logical processes which should guide the development of the work should be those of the student and not those of the teacher.
4. That opportunity be provided for the application of the postulational method to non-mathematical material.

Non-mathematical situations of interest to the pupils were used to introduce them to the importance of definition and to the fact that conclusions depend on assumptions, many of which are often unrecognized. To make definitions and assumptions and to investigate their implications is to have firsthand experience with the method of mathematics... (p. 117)

Fawcett’s teaching experiment, with a non-equivalent control group, continued through two school years, with the report covering just the first year. As the excerpt above suggests, the students eventually composed, collectively, their lists of undefined terms, definitions, and assumptions. The need for such elements arose in discussing everyday situations, such as the importance of definition in discussing how the governor of Ohio handled a particular bill (pp. 51-52). Of course the teacher played a major role in

initiating such discussions and in providing fruitful leads for particular results, but in the large the students were responsible for conjecturing results and then proving them. The evaluation of the experiment was based on a state geometry test and a test of the ability to analyze non-mathematical material. Even though the experimental students, after one year of a two-year treatment, had not covered the usual material in the standard course, their performance (although not reported thoroughly) seemed satisfactory on the 80-point state geometry test: Median 52.0, state median 36.5 (p. 102). More telling was the experimental students' performance on the analysis of non-mathematical material, where they out-performed by far the control group (change score of 7.5 vs a change score of 1.0; maximum possible not given) (p. 103). Fawcett also quoted the laudatory reactions of visitors and of the students themselves, contrasting the students' final remarks with the largely indifferent attitudes expressed at the beginning of the experiment.

One can only conclude from these studies that upper elementary school children can deal with proof ideas or actions, and that high school students can develop meaningful understandings of proof if they are taught appropriately.

*More recent feasibility studies.*

Several researchers have emphasized a sociocultural perspective in investigating and enhancing the development of students' proof schemes. According to this perspective, the development of any higher voluntary form of human knowledge cannot be understood apart from the social context in which it occurs. As such, learning is necessarily a product of social interaction. Key to this perspective is Vygotsky's (1978) notion of "zone of proximal development" (ZPD): the difference between what the students can do under adult guidance or in collaboration with more capable peers and what they can do without guidance. A direct and critical implication of this perspective is *instructional scaffolding* (Mercer, 1995), which refers to the provision of guidance and support that is increased or withdrawn in response to the developing competence of the learner. Blanton and Stylianou (2003) further suggest:

Since learning is viewed as a product of interaction, it follows that one's development within the ZPD is affected by the intellectual quality and developmental appropriateness of these interactions (Diaz, Neal, & Amaya-Williams, 1999). In other words, the extent of one's development within the ZPD is predicated in part upon how the more knowing other organizes, or scaffolds, the task at hand. Thus, if we intend to understand development within the ZPD, we must be thinking about if and how tasks can be scaffolded to extend one's learning. (p. 114)

Blanton and Stylianou (2003) have investigated the nature and role of this scaffolding in the learning and teaching of proof. Specifically, they addressed two questions: (a) What is the nature and meaning of instructional scaffolding in the classroom in the development of students' proof ability? and (b) How do different types of scaffolding prompts from the teacher affect students' self-regulatory thinking about proof? Their results suggest

... students who engage in whole-class discussions that include metacognitive acts as well as transactive discussions about metacognitive acts make gains in their ability to construct mathematical proofs. Moreover, students' capacity to engage in these types of discussion is a habit of mind that can be scaffolded

through the teacher's transactive prompts and facilitative utterances. ... [This] suggests that students can internalize public argumentation in ways that facilitate private proof construction if instructional scaffolding is appropriately designed to support this. (p. 119)

Some of the most promising work directly or indirectly related to students' growth in elements of proof has been with elementary school children, as early as the primary grades and most often in problem-based settings where questions of proof have come up as intended by-products of the investigations. Evidence exists to indicate that a classroom environment that is conducive to social interactions among students and between the teacher and the students can be productive. Hence, these studies have often incorporated small-group work and whole-class discussions sharing how different children were thinking (that is, justifying their work) rather than just focusing on the numerical result. For example, Zack (2002) had two groups of fifth graders work on the following problem: Find all the squares on an  $n$  by  $n$  chessboard, first with  $n = 4$ , then  $n = 5$ , then  $n = 10$ . What if it was a 60 by 60 square? Can you prove that you have them all? After working on this problem, her students offered eight different counter-arguments to an answer of 2310 (based on 6 times the established answer of 385 for a 10 by 10). Surely students who experience such instruction will develop different proof schemes than will children whose teachers always judge the correctness of an answer. In some cases (Cobb et al., 1991) researchers worked closely with classroom teachers in designing the tasks and studying the effects of the children's work and their discussions. In other cases (e.g., Carpenter, Franke, Jacobs, Fennema, & Empson, 1998), the teachers were provided with a great deal of background on children's thinking strategies but left to their own in designing and carrying out lessons; however, an earmark of the classroom work was to call for students to give justifications.

Other examples of early appearances of proof ideas come from work aiming toward algebra in the elementary grades (Schifter, Monk, Russell, Bastable, & Earnest, 2003). Some of their instruction deliberately but indirectly deals with such generalizations as commutativity of addition, although that language may not be used in the early exposures. Key are the tasks and teacher questions. Questions such as "Will it always work that way?" or "Why does it work out that way?" are a natural part of the instruction and indeed become an expected part of the lesson. Their work shows that children's conviction about at least some generalizations may follow an interesting path. For example, first graders, when asked "Will it always work that way?" after noting a specific example for commutativity of addition, may be uncertain that the idea will always work. Later, they may be confident that the answer is "Yes" because they have tried lots of examples (an empirical proof scheme), but still later, they may return to uncertainty because they are aware that there are many untested cases and that they have not tried them all. And, not unexpectedly, some students will over-generalize, accepting commutativity of subtraction as well.

Several Italian studies have also examined the nature of students' justifications as encountered in problem-based lessons (e.g., Boero, Chiappini, Garuti, & Sibilla, 1995; Boero, Garuti, & Mariotti, 1996). One example is their grade 8 students' study of shadows. Using data that the students collected over a period of time about the shadows of two vertical sticks, they examined the question of whether the shadows of a vertical stick and an oblique stick can be parallel, and if so, when. Small-group work, with



teacher help, led to clearly stated conjectures (e.g., If the sun's rays belong to the vertical plane of the oblique stick, then the shadows are parallel) that were then examined further, with an eye toward establishing them "in general." In the analysis of these subsequent attempts at general arguments (i.e., proofs), the researchers noticed that in the successful proofs, there were connections with key observations made during the conjecture-forming stage. The researchers' collective work has led to their hypothesis of "cognitive unity," emphasizing the close connection between the reasoning during the formation of a conjecture and the reasoning in an eventual proof:

(D)uring the production of a conjecture, the student progressively works out his/her statement through an intensive argumentative activity functionally intermingling with the justification of the plausibility of his/her choices; during the subsequent statement proving stage, the student links up with this process in a coherent way, organizing some of the justifications ('arguments') produced during the construction of the statements according to a logical chain (Boero, Garuti, Lemut, & Mariotti, 1996, p. 119-120).

They also argue that such a process is followed by many mathematicians--during the conjecturing stage, the mathematician uses arguments that can later be adapted to support his or her mathematical proof—and make the case that much more instruction in mathematics should involve conjecturing.

One series of studies (Maher, 2002; Maher & Martino, 1996; Martino & Maher, 1999) carried out with instruction in a similar problem-based vein, is notable because of its long-term nature (occasional sessions over 14 years, usually separate from the regular mathematics classes) and because of proof behaviors—proof by contradiction, proof by cases, proof by mathematical induction—that arose naturally, if informally, even in elementary school, at least on the part of some students. The nature and flavor of the sessions is communicated by these retrospections of a participant:

Well, we break up into groups...like five groups of three, say, and everyone in their own groups would have their own ideas, and you'd argue within your own group, about what you knew, what I thought the answer was, what you thought the answer was and then from there, we'd all get together and present our ideas, and then this group would argue with this group about who was right with this... (Maher, 2002, p. 37)... You didn't come in and say, "this is what we were learning today and this is how you're going to figure out the problem." We were figuring out how we were going to figure out the problem. We weren't attaching names to that but we could see the commonness between what we were working on there and maybe what we had done in school at some point in time and been able to put those things together and come up with stuff and to do these problems to come up with, what would be our own formulas because we didn't know that other people had done them before. We were just kind of doing our own thing trying to come up with an answer that was legitimate and that no matter how you tried to attack it, we could still answer it... (Maher, 2002, p. 32).

As the excerpts illustrate, all of these student-centered, problem-based studies have involved a way of teaching that is in stark contrast to the stereotype of a mathematics class: Students check homework, teacher illustrates something new,

students then do seat-work or homework to practice the new material. The didactical contract (Brousseau, 1997) in the experimental classes was obviously quite different from that in the stereotypical one. In particular, the “social norms” were quite different: Students were expected to work together rather than singly, students were to explain their solution methods, and students were to listen carefully and evaluate the explanations of other students--and hence perhaps learn different proof schemes.

Yackel and Cobb (1996) have sharpened the analysis of social norms to identify “sociomathematical norms,” those social norms that refer specifically to *mathematical* activity. For example, coming to accept that an explanation is expected might be a general social norm, whereas what constitutes an acceptable mathematical explanation would be a sociomathematical norm (Yackel & Cobb, 1996). Other sociomathematical norms might include norms dealing with when different explanations are mathematically different, or when a justification is acceptable, or when justifications or explanations convey efficiency or elegance. Indeed, Yackel, Rasmussen, and King (2000) focused on the norms that developed during a problem-based undergraduate differential equations course. They found that students “frequently explained their reasoning without prompting, offered alternative explanations and attempted to make sense of other students’ reasoning and explanations, despite the fact that their prior experiences were with traditional approaches to mathematics instruction” (p. 276) and, in particular, there was evidence of the development of such sociomathematical norms as, what is an acceptable mathematical justification and what makes up a mathematically different explanation. That the students, who had likely experienced mathematics classrooms with quite different social norms, developed such norms is particularly encouraging. Yackel and Cobb point out that sociomathematical norms should be examined in teaching modes that are different from the inquiry mode in which they have been studied, because the classroom conduct, whether stereotypic or not, will automatically convey what is an acceptable norm. For example, one can wonder whether some students, young or experienced, might be more comfortable with a more directed approach. Dweck (1999) has identified different goal orientations on the part of students. Some students may be guided primarily by performance goals like grades, parent/teacher approval, high marks, or status with others, whereas others have what Dweck calls “learning goals”—their primary interest is in understanding or mastery. A student’s extreme preference for performance goals may be an unfortunate aspect of schooling or society as it exists, of course, but it might also result in resistance to sociomathematical norms that do not obviously support perceived performance. It is an interesting question as to whether different sorts of teaching might shape a student’s learning goals.

In most of the studies above, outside support for the teacher was particularly important. The question of what is feasible in classrooms without further teacher preparation or researcher involvement is crucial, as, for example, Yackel and Hanna point out (2003). It is nonetheless exciting to envision learners, starting in the primary grades and continuing through high school and college, developing the social and sociomathematical norms about proof, and the proof schemes, that one might wish, but it is daunting to think of the changes needed in curricula and teachers (and testing programs) to support the development of such norms. As an indication of areas of teacher preparation that are important, Martino and Maher (1999) suggest, based on their videotape analysis of their multi-year study, that there is a “strong relationship between

(1) [the teacher's] careful monitoring of students' constructions leading to a problem solution, and (2) the posing of a timely question which can challenge learners to advance their understanding" (p. 53). And there may be useful norm-supporting ideas even for classes in which problem-based or inquiry-based instruction is not the mode. For example, Blanton, Stylianou, and David (2003), based on their work with a college discrete mathematics course and extending the work mentioned earlier (Blanton & Stylianou, 2002), believe that an instructor's attention to likely student understanding (and hence a zone of proximal development) can enable an instructional scaffolding that, along with a careful discussion of (preferably) students' proofs, can support students' understanding of proof.

Again, all of these studies illustrate how appropriate mathematical curricula, together with appropriate teacher intervention, can help students to develop critical elements of deductive reasoning and an openness to the ideas of mathematical proof. As feasibility studies, these efforts have shown that even elementary school children can be remarkably able to make sense of mathematics, if given opportunities. Their explanations not only help them in their communication skills, but the explanations are likely also to help their classmates, to give to their teachers insights into the children's level of understanding, and to engender or foster students' growth in their proof schemes and their grasp of sociomathematical norms.

#### *Use of technology.*

The increasing use of dynamic geometry software during the last couple of decades has provoked researchers to look closely into the learning benefits as well as the potential risks of this tool. In this section we focus on studies conducted to investigate the impact of dynamic geometry environments (DGEs) on the learning proof, primarily because those environments have been involved in several studies.

Generating and measuring many examples, as is now possible and easy with DGEs, would seem only to support the idea that examples prove a result (the empirical proof scheme) and hence interfere with any need for deductive proof. Chazan (1993a,b) noted and corroborated that the ideas that evidence is proof and that proof is merely evidence were widespread. He interviewed 17 students who were taught geometry in a DGE environment (1993b) to see how students fared when they experienced a geometry curriculum that involved both measurement of cases and deductive proof, with the curriculum including attention to the different types of argumentation involved with the two methods. He found both evidence-is-proof (the empirical proof scheme) and proof-is-merely-evidence postures on the part of the students, and hence he noted that the "comparison and contrast of verification and deductive proof certainly deserves explicit attention in mathematics classrooms" (1993b, p. 382), with the teachers involved in the study feeling that more time should have been spent in dealing with students' doubts about deductive proofs (1993a, p. 109). However, he also noticed that "fewer students considered evidence to be proof... whereas more students were skeptical about the limits of applicability of deductive proofs... Some students seemed to become more skeptical about deductive proofs as a result of becoming more skeptical about measurement of examples" (1993a, p. 109). He also suggested, based on interviewee comments, "that the explanatory aspect of proofs is a useful starting point for a discussion of the value of deductive proofs" (1993b, p. 383), that some students have no idea of what deductive

proofs are intended to do, and that some students resist the idea that a deductive argument can assure that there cannot be counterexamples.

A special issue of the journal, *Educational Studies in Mathematics* (2000, 44), included four reports on teaching experiments involving DGE. The first four are by Mariotti, by Jones, by Marrades and Gutierrez, and by Hadas, Hershkowitz, and Schwartz. The last paper is by Laborde, and synthesizes the previous four, providing a connecting theme among them by using Brousseau's construct of "milieu." We continue with these studies.

One of the significant results of the four studies is that their findings address a major concern regarding the use of DGEs in the teaching of geometry in school: "the opportunity offered by [DGE] to 'see' mathematical properties so easily might reduce or even kill any need for proof and thus any learning of how to develop a proof" (Laborde, 2000, p. 151).

Jones' study was a teaching experiment with 12-year-old students. The intervention involved students working in pairs or small groups on the classification of quadrilaterals. The instructional activities involved tasks where the students were to reproduce a figure that could not be "messed up" by dragging any of its components (a vertex or segment). More advanced activities included tasks of producing a figure that can be transformed into another specified figure by dragging (e.g, from a rectangle to a square). In each case the students had to explain why their constructed figure was the expected one. Laborde (2000) points out that this type of explanation consists of giving the conditions that imply that the constructed figure is the expected type of quadrilateral, a necessary activity to understand how proof works. From our perspective, these explanations in the context of a DGE involve movement, and, therefore—as we have discussed in earlier in the section—they are causal, and hence deductive. Accordingly, Jones' (2000) study suggests that the DGE does not necessarily eliminate the need for proof in the students' eyes but can enhance students' deductive proof scheme. The developmental path of students' conception in this study started where students were. Initially, according to Jones, students lacked the capability to describe or explain in precise mathematical language. The instructional emphasis in this stage was on description rather than explanation, where students utilized perception rather than mathematical language to describe their observations. As the emphasis extended to explanations, students' language became more precise but was mediated by the DGE terminology (e.g., the term "dragging"). By end of the teaching experiment, students' explanations related entirely to the mathematical context.

Mariotti's study was a long-term teaching experiment with 15-16-year-old students. Students were engaged in DGE activities through which the students themselves constructed a geometric system of axioms and theorems as a system of Cabri Geometry commands. Similar to Jones, Mariotti emphasized activities where the task is for pairs of students to construct geometric figures, describe the construction procedure, and justify why the procedure produces the expected figure. The basic conceptual change that Mariotti's (2000) study achieved was in students' status of justification, which transitioned from an "intuitive" geometry—a collection of self-evident properties—to a theoretical geometry—a system of statements validated by proof. The theoretical geometry that Mariotti's students had constructed seems to be more than a deductive system, in that students were not only constructing and proving theorems but also

establishing the axioms on which these theorems rest, and thereby laying the foundations for their axiomatic proof scheme.

Hadas, Hershkowitz and Schwarz' (2000) study was done in the context of a geometry course that emphasized the concept of proof. They developed instructional activities involving students making assertions about certain geometric relations and later checking them with a DGE. The choice and sequence of activities were such that upon checking their assertions with the DGE the students would find them to be false—a realization that would make the students curious as to the reason for the falsity of the conjecture. For example, in one of the activities the students began with two tasks. The first task was to measure (with the software) the sum of the interior angles in polygons as the number of sides increases, generalizing their observation, and then explaining their conclusion. The second task was to measure (with the software) the sum of the exterior angles of a quadrilateral. Following this, the students were asked to hypothesize the sum of the exterior angles for polygons as the number of sides increases, and to check their hypothesis by measuring (with the software) and explain what they found. Hadas, Hershkowitz and Schwartz succeeded in creating in the students a need to find out the cause for their assertion to be untrue. Laborde (2000) points out that such an achievement would have been impossible without the use of a dynamic geometry system, for “the false conjectures came after students were convinced of other properties thanks to the DG system. ... [The] interplay of conjectures and checks, of certainty and uncertainty, was made possible by the exploration power and checking facilities offered by the DG environment” (p. 154).

Marrades and Gutierrez investigated how DGEs can help secondary-school students (aged 15-16 years) enhance their proof schemes. As in Hadas, Hershkowitz and Schwarz' study, Marrades and Gutierrez showed that a DGE can help students realize the need for formal proofs in mathematics. By interpreting their results in terms of our taxonomy of proof schemes, an important observation reported in their study is that students' transition to deductive proof schemes is very slow; the total teaching experiment lasted 30 weeks, with two 55 minute class per week. Of particular importance is their finding that for this transition to take place instruction must not ignore students' current empirical proof schemes and must institute a didactical contract that attempts to suppress the authoritative proof scheme. Their method was to repeatedly emphasize “the need to organize justifications by using definitions and results (theorems) previously known and accepted by the class” (p. 120). Finally, another significant finding of this study is that the ability to produce deductive proof evolves hand in hand with students' understanding of subject matter: the concepts and properties related to the topic being studied. This is consistent with other findings. Simon and Blume (1996), for example, illustrated that a learner may not fully understand another's proof because of a limited grasp of the concepts addressed in the proof (p. 29). One can argue that such exposure might lead to a disequilibrium and eventually a greater understanding of both the concepts and the proof.

Hence, DGEs are a promising tool, but they do not automatically or easily lead to improved proof schemes. Accomplishing that sort of growth apparently requires a carefully laid-out curriculum (cf. de Villiers, 1999) and considerable adjustment by a teacher accustomed only to telling as the mode of instruction. Lampert (1993), for example, described some of the difficulties encountered by teachers who allowed

conjecturing with a DGE. (Making conjectures, itself, may be difficult for students new to the expectation. Koedinger [1998], with an eye toward possible software activities, noted that with his task of writing a conjecture about kites, given the definition of kite, about a quarter of the roughly 60 geometry students could not come up with a non-trivial conjecture within 20 minutes.) Lampert noted that the change from “sage on the stage” to “guide on the side” required adjustments both for the teacher and the students, since a different sort of didactical contract was involved. Teachers were also concerned about the coverage of a standard body of content (for external testing purposes, or for later courses), as well as the departure from the usual, familiar axiomatic development that often eventuated.

### *Does Explicit Teaching of Logic Work?*

The use of deductive proof schemes, at least implicitly, involves logic. Calls for the explicit teaching of logic are easy to find. More than 40 years ago, consensus at the future-oriented Cambridge Conference on School Mathematics (1963) was that “it is hardly possible to do anything in the direction of mathematical proofs without the vocabulary of logic and explicit recognition of the inference schemes” (p. 39). There have been occasional caveats, as in Suppes' remarks:

I would not advocate an excessive emphasis on logic as a self-contained discipline... What I do feel is important is that students be taught in an explicit fashion classical rules of logical inference, learn how to use these rules in deriving theorems from given axioms, and come to feel as much at home with simple principles of inference like *modus ponendo ponens* as they do with elementary algorithms of arithmetic” (1966, p. 72).

Studies of the effects of an explicit attention to logic have not, however, indicated that there is then a pay-off in proof-writing. School geometry has long had a goal of improving students' logical thinking (cf. NCTM, 1970), and several studies have looked at the influence of including an explicit, concentrated treatment of logic in that course (e.g., Deer, 1969; Mueller, 1975; Platt, 1967). Yet, the usual outcome, even when the logic treatment has involved up to four weeks (Platt) and/or a reasonable number of students (Mueller, 146 students; Platt, 12 classes), is a finding of no-significant-differences in proof performance.

It is reasonable to expect that the teacher's emphasis on logical reasoning, even in the absence of explicit treatment in a curriculum, might influence the students' use of logic themselves. In a study of the influence of teachers' language on performance, Gregory and Osborne (1975) explored whether the teachers' use of logic influenced their students' performance on a logic test, by examining audiotapes of five lessons of 20 junior high school mathematics teachers. Although acknowledging some design problems, they found a significant positive correlation between logic test performance for some types of inferences (e.g., when negatives were involved) and the usages of if-then statements by the five teachers with the greatest average use vs. usages of the five teachers with the lowest. Eye-catching, however, was the range in the averages of the number of usages per lesson of if-then sorts of statements: 8.3 to 40.6, with individual lessons giving from 3 to 48 usages (p. 28). Unfortunately, there was no control for content covered (some teachers were teaching number theory, some word problems, and some properties of geometric figures), so the range is only suggestive of what might be

an important element in students' growth in deductive reasoning: the frequency of use of logical statements and logical reasoning in the classroom.

Thus there is much evidence that some important elements of deductive reasoning are not natural parts of students' repertoires, at a variety of school levels. How to develop these elements so that they are recognized and utilized in mathematical proofs is an open question, and may not be realized through a single, short unit. As the 2000 NCTM Principles and Standards assert, "Reasoning and proof cannot simply be taught in a single unit on logic, for example, or by 'doing proofs' in geometry" (NCTM, 2000, p. 56).

### Reflection

In this chapter we have presented a comprehensive perspective on proof learning—a perspective that addresses mathematical, historical-epistemological, cognitive, sociological, and instructional factors. Comprehensive perspectives on proof are needed, we argued, in order to better understand the nature and roots of students' difficulties with proof so that effective instructional treatments can be designed and implemented to advance students' conceptions of and attitudes toward proof. Our perspective grew out of a decade of investigations—empirical as well as theoretical—into students' conceptions of proof. In various periods and stages of these investigations we have repeatedly confronted questions that collectively address a combination of the five factors mentioned above.

The notion of "proof scheme" serves as the main lens of our comprehensive perspective on proof. Through it, for example, we analyze and interpret students' proving behavior—in their individual work as well as in their interaction with others—and understand the development of proof in the history of mathematics. *A proof scheme consists of what constitutes ascertaining and persuading for a person (or community).* This definition was born out of cognitive, epistemological, and instructional considerations. Specifically, a critical observation in our and other scholars' work is that proof schemes vary from person to person and from community to community—in the classroom, with individual students as well as the class as a whole, and throughout history. "Proof," when viewed in this subjective sense, highlights the student as learner. As a result, teachers must take into account what constitutes ascertainment and persuasion for their students and offer, accordingly, instructional activities that can help them gradually refine and modify their proof schemes into desirable ones. This subjective view of proof emerged from our studies and impacted many of the conclusions we drew from them. For example, it influenced our conclusions as to the implications of the epistemology of proof in the history of mathematics to the conceptual development of proof with students, the implications of the way mathematicians construct proofs to instructional treatments of proof, and the implications of the everyday justification and argumentation on students' proving behaviors in mathematical contexts. The subjective notion of proof scheme is not in conflict with our insistence on unambiguous goals in the teaching of proof—namely, to gradually help students develop an understanding of proof that is consistent with that shared and practiced by the mathematicians of today. The question of critical importance is: What instructional interventions can bring students to see an intellectual need to refine and alter their current proof schemes into deductive proof schemes (Harel, 2001)?

The status studies we have reviewed and presented in this paper show the absence of the deductive proof scheme and the pervasiveness of the empirical proof scheme among students at all levels. Students base their responses on the appearances in drawings, and mental pictures alone constitute the meaning of geometric terms. They prove mathematical statements by providing specific examples, not able to distinguish between inductive and deductive arguments. Even more able students may not understand that no further examples are needed, once a proof has been given. Students' preference for proof is ritualistically and authoritatively based. For example, when the stated purpose was to get the best mark, they often felt that more formal—e.g., algebraic—arguments might be preferable to their first choices. These studies also show a lack of understanding of the functions of proof in mathematics, often even among students who had taken geometry and among students for whom the curriculum pays special attention to conjecturing and explaining or justifying conclusions in both algebra and geometry. Students believe proofs are used only to verify facts that they already know, and have no sense of a purpose of proof or of its meaning. Students have difficulty understanding the role of counterexamples; many do not understand that one counterexample is sufficient to disprove a conjecture. Students do not see any need to prove a mathematical proposition, especially those they consider to be intuitively obvious. This is the case even in a country like Japan where the official curriculum emphasizes proof. They view proof as the method to examine and verify a later particular case. Finally, the studies show that students have difficulty writing valid simple proofs and constructing, or even starting, simple proofs. They have difficulty with indirect proofs, and only a few can complete an indirect proof that has been started.

We believe there is a need for more longitudinal studies regarding students' proof schemes. The most difficult studies to carry out, for both financial and design reasons are longitudinal studies. Yet we cannot gain a solid understanding of the effect on students of continued attention to justification and proof throughout their studies in mathematics, except through longitudinal studies. A one- or two-year exposure (let alone a one-semester treatment) to instruction and curricula attentive to reason-giving can be dwarfed by a multiple-year focus on instruction and curricula that, to use an extreme example, emphasize rote skills likely to be useful in some external testing program. In the latter cases, being asked to give reasons and arguments might well be viewed as aberrations and irrelevant to the perceived "really important" side of mathematics.

The findings from studies of teachers' conceptions of proof do not look much better than those with students. Overall, teachers seem to acknowledge the verification role of proof, yet for many the empirical proof schemes seem to be the most dominant, even in dealing with mathematical statements, and they do not seem to understand other important roles of proof, most noticeably its explanatory role. Some teachers tend to view proof as an appropriate goal only for the mathematics education of a minority of students, not considering proof and justification to be a central concern in school mathematics, as has been repeatedly called for by the mathematics education leadership (e.g., NCTM, 1989, 2000). Studies show that little or no instructional time is allocated to the development of the deductive proof schemes, not even in geometry. In the U.S. explicit mathematical reasoning in mathematics classes is rare, and in algebra and pre-algebra courses it is virtually absent. Many teachers are unlikely to teach proof well, since their own grasp of proof is limited. It is important to determine better the extent to



which teachers are equipped to deliver a curriculum in which proof is central. Results from studies like those of Knuth (2002a) and Manaster (1998), if indeed typical of a widespread performance of mathematics teachers, demand attention both on the part of university mathematics departments, which have a primary responsibility for the preparation of mathematics teachers, and on school districts, which support the continued development of their existing faculty.

The bright side of the findings is that students who receive more instructional time on developing analytical reasoning by solving unique problems fare noticeably better on overall test scores. Likewise, students who have been expected to write proofs and who have had classes that emphasized proof were somewhat better than other students. It also seems possible to establish desirable sociomathematical norms relevant to proof, through careful instruction, often featuring the student role in proof-giving. There has been a concern that the ease with which technology can generate a large number of examples naturally could undercut any student-felt need for deductive proof schemes. Fortunately, several studies have shown that with careful, non-trivial planning and instruction over a period of time, progress toward deductive proof schemes is possible in technology environments, where such desiderata as making conjectures and definitions occur.

An important element in deductive proof schemes is of course the use of logical reasoning. Yet there is evidence that many students, and possibly even many teachers, do not have a good grasp or appreciation of some important principles of logic. Nor is it clear as to how best to devise instruction to improve performance with logic in mathematical (and non-mathematical) contexts. It is unclear to us how best to prepare students to deal with the logical reasoning essential in mathematical proofs and valuable in even informal justifications—osmosis, or explicit attention? And, in particular, the knowledge of teachers of mathematics about logical reasoning may be a matter of concern (e.g., Easterday & Galloway, 1995). We see a need for the incorporation of items on proof or logic (even multiple-choice ones) into the periodic National Assessments of Education Progress. From the practical viewpoint, the NAEPs exist, and they offer a view of performance across the U.S. Even more pleasing would be to see large-scale efforts devoted explicitly to the study of performance in proof and logic, like those in Great Britain (Healy & Hoyles, 1998, 2000; Hoyles & Kuchemann, 2002). A deep look at the students and teachers' knowledge of proof, on the one hand, and at the development of the deductive proof scheme in the history of mathematics, on the other, has provided us with important insights as to what might account for students' difficulties in constructing this scheme and what instructional approaches can facilitate its construction. In particular, considerations of historical-epistemological developments have led us to new research questions with direct bearing on the learning and teaching of proofs. For example:

1. To what extent and in what ways is the nature of the content intertwined with the nature of proving? In geometry, for example, does students' ability to construct an image of a point as a dimensionless geometric entity impact their ability to develop the Greek axiomatic proof scheme?
2. What are the cognitive and social mechanisms by which deductive proving can be necessitated for the students? The Greek's construction of their geometric edifice seems to have been a result of their desire to create a

consistent system that was free from paradoxes. Would paradoxes of the same nature create a similar intellectual need with students?

3. Students encounter difficulties in moving between proof schemes, particularly from the Greek's axiomatic proof scheme (the one they construct in honors high-school geometry, for example) to the modern axiomatic proof scheme (the one they need to succeed in a real analysis course, for example). Exactly what are these difficulties? What role does the emphasis on *form* rather than *content* in modern mathematics (as opposed to Greek mathematics where content is more prominent) play in this transition? Can students develop the modern axiomatic proof without computational fluency? What role does the causality proof scheme play in this transition?

These questions are examples of what we have delineated as important contributions from the history of mathematics to our thinking about students' proof schemes. But there are likely to be other valuable ideas from further study of the growth or development of proof ideas in the history of mathematics. How mathematical proof arose in other cultures—e.g., the basis for the Japanese temple drawings (Rothman & Fukagawa, 1998)—would in itself be fascinating and potentially instructive about how proof ideas might be introduced or developed today. In this respect, our effort to form a comprehensive perspective on proof is an attempt to understand what might be called the “proving conceptual field,” a term analogous to Vergnaud's (1983, 1988) “multiplicative conceptual field.” Like the multiplicative conceptual field, the proving conceptual field may be thought of as a set of problems and situations for which closely connected concepts, procedures, and representations are necessary.

## References

- Adey, P. (1999). *The science of thinking and science for thinking: A description of logical acceleration through science education*. (INNODATA Monographs-2). Geneva, Switzerland: International Bureau of Education.
- Antonini, S. (2003). Non-examples and proof by contradiction. In N. Pateman, B. J. Dougherty, & J. Zilliox (Eds.), *Proceedings of the 27<sup>th</sup> Annual Meeting of the International Group for Psychology in Mathematics Education (Vol. 2)*, pp. 49-55). Honolulu: University of Hawaii.
- Aristotle (1941). *The Basic Works of Aristotle*, R. McKeon (Ed. & Translator). New York: Random House.
- Babai, L. (1992). Transparent proofs. *Focus*, 12(3), 1-2.
- Balacheff, N. (1988). Aspects of proof in pupils' practice of school mathematics. In D. Pimm (Ed.), *Mathematics, teachers and children* (pp. 216-235). London: Hodder & Stoughton.
- Balacheff, N. (1991). The benefits and limits of social interaction: The case of mathematical proof. In A. Bishop, E. Mellin-Olsen, & J. van Dormolen (Eds.), *Mathematical knowledge: Its growth through teaching* (pp. 175-192). Dordrecht, The Netherlands: Kluwer.
- Balacheff, N. (2002). A researcher epistemology: a deadlock for educational research on proof. *Proceedings of 2002 International Conference on Mathematics: Understanding Proving and Proving to Understand* (pp. 23-44). Department of Mathematics, National Taiwan Normal University.
- Ball, D., & Bass, H. (2003). Making mathematics reasonable in school. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 27-44). Reston, VA: NCTM.
- Battista, M., & Clements, D. (1992). Geometry and spatial reasoning. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 420-464). New York: Macmillan.
- Bell, A. W. (1976). A study of pupils' proof-explanations in mathematical situations. *Educational Studies in Mathematics*, 7, 23-40.
- Blanton, M. L., & Stylianou, D. A. (2002). Exploring sociocultural aspects of undergraduate students' transition to mathematical proof. In D. S. Mewborn, P. Sztajn, D. Y. White, H. G. Wiegel, R. L. Bryant, & K. Nooney (Eds.), *Proceedings of the 24<sup>th</sup> Annual Meeting for Psychology in Mathematics Education –North America, Vol. 4*, (pp. 1673-1680). Athens, GA: University of Georgia.
- Blanton, M. L., Stylianou, D. A., & David, M. M. (2003). The nature of scaffolding in undergraduate students' transition to mathematical proof. In N. Pateman, B. J. Dougherty, & J. Zilliox (Eds.), *Proceedings of the 27<sup>th</sup> Annual Meeting of the International Group for Psychology in Mathematics Education, Vol. 2* (pp. 113-120). Honolulu: University of Hawaii.
- Boero, P., & Garuti, R. (1994). Approaching rational geometry: from physical relationships to conditional statements. In J. Ponte & J. Matos (Eds.), *Proceedings of the Eighteenth Annual Meeting of the International Group for Psychology in Mathematics Education, Vol. II* (pp. 96-103). Lisbon: University of Lisbon.

- Boero, P., Chiappini, G., Garuti, R., & Sibilla, A. (1995). Towards statements and proofs in elementary arithmetic: An exploratory study about the role of teachers and the behaviour of students. In L. Meira & D. Carraher (Eds.), *Proceedings of the 19<sup>th</sup> Psychology in Mathematics Education Conference*, vol. 3 (pp. 129-136). Recife, Brazil: Universidade Federal de Pernambuco.
- Boero, P., Garuti, R., & Mariotti, M. (1996). Some dynamic mental processes underlying producing and proving conjectures. In L. Puig & A. Gutierrez (Eds.), *Proceedings of the 20<sup>th</sup> Conference of the International Group for the Psychology of Mathematics Education*, vol. 2 (pp. 121-128). Valencia, Spain: Universitat de Valencia.
- Boero, P., Garuti, R., Lemut, E., & Mariotti, M. (1996). Challenging the traditional school approach to the theorems: A hypothesis about the cognitive unity of theorems. In L. Puig & A. Gutierrez (Eds.), *Proceedings of the 20<sup>th</sup> Conference of the International Group for the Psychology of Mathematics Education*, vol. 2 (pp. 113-120). Valencia, Spain: Universitat de Valencia.
- Brousseau, G. (1997). *Theory of didactical situations in mathematics*. N. Balacheff, M. Cooper, R. Sutherland, & V. Warfield (Eds. & Trans.). Dordrecht, The Netherlands: Kluwer.
- Brumfiel, C. (1973). Conventional approaches using synthetic Euclidean geometry. In K. Henderson (Ed.), *Geometry in the mathematics curriculum* (pp. 95-115). Thirty-sixth Yearbook of the National Council of Teachers of Mathematics. Reston, VA: NCTM.
- Cambridge Conference on School Mathematics. (1963). *Goals for school mathematics*. Boston: Houghton Mifflin.
- Carpenter, T. P. (1989). Introduction. In M. M. Lindquist (Ed.), *Results from the fourth mathematics assessment of the National Assessment of Educational Progress* (pp. 1-9). Reston, VA: NCTM.
- Carpenter, T. P., Coburn, T. G., Reys, R. E., & Wilson, J. W. (1978). *Results from the first mathematics assessment of the National Assessment of Educational Progress*. Reston, VA: NCTM.
- Carpenter, T. P., Franke, M. L., Jacobs, V. R., Fennema, E., & Empson, S. (1998). A longitudinal study of invention and understanding in children's multidigit addition and subtraction. *Journal for Research in Mathematics Education*, 29, 3-20.
- Chazan, D. (1993a). Instructional implications of students' understandings of the differences between empirical verification and mathematical proof. In J. Schwartz, M. Yerushalmy, & B. Wilson (Eds.), *The Geometric Supposer: What is it a case of?* (pp. 107-116). Hillsdale, NJ: Erlbaum.
- Chazan, D. (1993b). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics*, 24, 359-387.
- Cobb, P. (1994). Theories of mathematical learning and constructivism: A personal view. Paper presented at the Symposium on Trends and Perspectives in Mathematics Education, Institute for Mathematics, University of Klagenfurt, Austria.

- 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.
- Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist*, 31, 175-190.
- Coe, R. & Ruthven, K.. (1994). Proof practices and constructs of advanced mathematics students. *British Educational Research Journal*, 20(1), 41-53.
- Deer, G. W. (1969). The effects of teaching an explicit unit in logic on students' ability to prove theorems in geometry (Doctoral dissertation, Florida State University, 1969). *Dissertation Abstracts International*, 30, 2284-2285B.
- de Villiers, M. D. (1999). *Rethinking proof with the Geometer's Sketchpad*. Emeryville, CA: Key Curriculum Press.
- Dubinsky, E., & Yiparaki, O. (2000). On student understanding of AE and EA quantification. In E. Dubinsky, A. H. Schoenfeld, & J. Kaput (Eds.), *Research in collegiate mathematics education IV*. (pp. 239-286). Providence, RI: American Mathematical Society.
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology Press.
- Easterday, K. E., & Galloway, L. L. (1995). A comparison of sentential logic skills: Are teachers sufficiently prepared to teach logic? *School Science and Mathematics*, 95, 431-436.
- Eisenberg, T. A., & McGinty, R. L. (1974). On comparing error patterns and the effect of maturation in a unit on sentential logic. *Journal for Research in Mathematics Education*, 5, 225-237.
- Epp, S. S. (2003). The role of logic in teaching proof. *American Mathematical Monthly*, 110(10), 886-899.
- Evans, J. St. B. T. (1982). *The psychology of deductive reasoning*. London: Routledge & Kegan Paul.
- Fawcett, H. P. (1995). *The nature of proof*. Thirteenth Yearbook of the NCTM. New York: Teachers College, Columbia University. (Original work published 1938).
- 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* (128-131). Antwerp, Belgium: Universitaire Instelling Antwerpen.
- Freudenthal, H. (1962). The main trends in the foundations of geometry in the 19<sup>th</sup> century. In E. Nagel, P. Suppes, & A. Tarski, *Logic, methodology, and philosophy of science*. Stanford, CA: Stanford University Press.
- Freudenthal, H. (1973). *Mathematics as an educational task*. Dordrecht, Holland: D. Reidel Publishing Company.
- Fujita, T., & Jones, K. (2003, July). *Critical review of geometry in current textbooks in lower secondary schools in Japan and the UK*. Paper presented at the 27<sup>th</sup> annual meeting of the International Group for Psychology of Mathematics Education, Honolulu.
- Galbraith, P. (1981). Aspects of proving: a clinical investigation of process. *Educational Studies in Mathematics*, 12, 1-28.
- Ginsburg, H. J., & Opper, S. (1969). *Piaget's theory of intellectual development*. Englewood Cliffs, NJ: Prentice-Hall.

- Goetting, M. M. (1995). The college students' understanding of mathematical proof (Doctoral dissertation, University of Maryland, 1995). *Dissertation Abstracts International*, 56-A, 3016.
- Gregory, J. W., & Osborne, A. R. (1975). Logical reasoning ability and teacher verbal behavior within the mathematics classroom. *Journal for Research in Mathematics Education*, 6, 26-36.
- Hacking, I. (1980). Proof and eternal truths: Descartes and Leibniz. In *Descartes: Philosophy, Mathematics and Physics*. Gaukroger, S. (Ed.). pp. 169-179. PUBLISHER?
- Hadas, N., Hershkowitz, R., & Schwarz, B. B. (2000). The role of contradiction and uncertainty in promoting the need to prove in dynamic geometry environments. *Educational Studies in Mathematics*, 44, 127-150.
- Haimo, D. T. (1995). Experimentation and conjecture are not enough. *American Mathematical Monthly*, 102(2), 102-112.
- Hanna, G. (1983). Rigorous proof in mathematics education. Toronto: OISE Press.
- Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange*, 21, 6-13.
- Hanna, G., & Jahnke, H. N. (1996). Proof and proving. In A. Bishop, K. Clements, C. Keiterl, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (Part 2, pp. 877-908). Dordrecht, The Netherlands: Kluwer.
- Harel, G. (1999). Students' understanding of proofs: A historical analysis and implications for the teaching of geometry and linear algebra. *Linear Algebra and Its Applications*, 302-303, 601-613.
- Harel, G. (2001). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. Campbell & R. Zazkis (Eds.), *The learning and teaching of number theory* (pp. 185-212). Dordrecht, The Netherlands: Kluwer.
- Harel, G. (in press). Greek versus modern mathematical thought and the role of Aristotelian causality in the mathematics of the Renaissance: Sources for understanding students' conception of proof. In P. Boero (Ed.), *The concept of mathematical proof*. Dordrecht, The Netherlands: Kluwer.
- Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in collegiate mathematics education III* (pp. 234-283). Providence, RI: American Mathematical Society.
- Hart, E. W. (1994). A conceptual analysis of the proof-writing performance of expert and novice students in elementary group theory. In J. Kaput & E. Dubinsky (Eds.), *Research issues in undergraduate mathematics learning*, MAA Notes No. 33 (pp. 49-157). Washington: Mathematical Association of America.
- Hartshorne, R. (2000). *Geometry: Euclid and beyond*. Springer, New York.
- Healy, L., & Hoyles, C. (1998). Justifying and proving in school mathematics, executive summary. London: Institute of Education, University of London.
- Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education*, 31(4), 396-428.
- Heath, T. L. (1956). *The thirteen books of Euclid's Elements* (2<sup>nd</sup> ed.), Vol. I. New York: Dover.

- Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics*, 24, 389-399.
- Hoffer, A. (1981). Geometry is more than proof. *Mathematics Teacher*, 74, 11-18.
- Hoyles, C. (1997). The curricular shaping of students' approaches to proof. *For the Learning of Mathematics*, 17(1), 7-16.
- Hoyles, C., & Kuchemann, D. (2002). Students' understanding of logical implication. *Educational Studies in Mathematics*, 51, 193-223.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence* (A. Parsons & S. Milgram, Trans.). New York: Basic Books.
- Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Sciences*, 5, 434-442.
- Jones, K. (2000). Providing a foundation for deductive reasoning: Students' interpretations when using dynamic geometry software and their evolving mathematical explanations. *Educational Studies in Mathematics*, 44, 55-85.
- King, I. L. (1970). A formative development of a unit on proof for use in the elementary school. (Doctoral dissertation, University of Wisconsin--Madison, 1969). *Dissertation Abstracts*, 31, 680A.
- King, I. L. (1973). A formative development of an elementary school unit on proof. *Journal for Research in Mathematics Education*, 4(1), 57-63.
- Klein, J. (1968). *Greek mathematical thought and the origin of algebra* (E. Brann, Trans.). Cambridge, MA: MIT Press. (Original work published 1934)
- Kleiner, I. (1991). Rigor and proof in mathematics: A historical perspective. *Mathematics Magazine*, 64(5), 291-314.
- Knuth, E. J. (2002a). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education*, 33(5), 379-405.
- Knuth, E. J. (2002b). Teachers' conceptions of proof in the context of secondary school mathematics. *Journal of Mathematics Teacher Education*, 5, 61-88.
- Knuth, E. J., Slaughter, M., Choppin, J., & Sutherland, J. (2002). Mapping the conceptual terrain of middle school students' competencies in justifying and proving. In S. Mewborn, P. Sztajn, D. Y. White, H. G. Wiegel, R. L. Bryant, & K. Nooney (Eds.), *Proceedings of the 24<sup>th</sup> Meeting for PME-NA, Vol. 4*, (pp. 1693-1700). Athens, GA.
- Koedinger, K. R. (1998). Conjecturing and argumentation in high-school geometry students. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments for developing understanding of geometry and space* (pp. 319-347). Mahwah, NJ: Erlbaum.
- Laborde, C. (2000). Dynamic geometry environments as a source of rich learning contexts for the complex activity of proving. *Educational Studies in Mathematics*, 44, 151-161.
- Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery*. Cambridge: Cambridge University Press.
- Lampert, M. (1993). Teachers' thinking about students' thinking about geometry: The effects of new teaching tools. In J. Schwartz, M. Yerushalmy, & B. Wilson (Eds.), *The Geometric Supposer: What is it a case of?* (pp. 107-116). Hillsdale, NJ: Erlbaum.

- Lawson, A. E., Karplus, R., & Adi, H. (1978). The acquisition of propositional logic and formal operational schemata during the secondary school years. *Journal of Research in Science Teaching*, 15, 465-478.
- Leron, U. (1985). A direct approach to indirect proofs. *Educational Studies in Mathematics*, 16, 321-325.
- Lester, F. K. (1975). Developmental aspects of children's ability to understand mathematical proof. *Journal for Research in Mathematics Education*, 6(1), 14-25.
- Lin, F-L., Lee, Y-S., & Wu Yu, J-Y. (2003). Students' understanding of proof by contradiction. In N. Pateman, B. J. Dougherty, & J. Zilliox (Eds.), *Proceedings of the 27<sup>th</sup> Annual Meeting of the International Group for Psychology in Mathematics Education*, Vol. 4 (pp. 443-449). Honolulu: University of Hawaii.
- Linchevsky, L., Vinner, S., & Karsenty, R. (1992). To be or not to be minimal? Students teachers' views about definitions in geometry. In W. Geeslin & K. Graham (Eds.), *Proceedings of the Sixteenth Psychology in Mathematics Education Conference*, vol. II, (pp. 48-55). Durham, NH: University of New Hampshire.
- Lovell, K. (1971). The development of the concept of mathematical proof in abler pupils. In M. Roszkopf, L. Steffe, & S. Taback (Eds.), *Piagetian cognitive-development research and mathematical education* (pp. 66-80). Washington: NCTM.
- Maher, C. A. (2002). How students structure their own investigations and educate us: What we've learned from a fourteen year study. In A. D. Cockburn & E. Nardi (Eds.), *Proceedings of the 26<sup>th</sup> Annual Meeting of the International Group for the Psychology of Mathematics Education*, Vol. 1 (pp. 31-46). Norwich, England: University of East Anglia.
- Maher, C. A., & Martino, A. M. (1996). The development of the idea of mathematical proof: A 5-year case study. *Journal for Research in Mathematics Education*, 27, 194-214.
- Manaster, A. B. (1998). Some characteristics of eighth grade mathematics classes in the TIMSS videotape study. *American Mathematical Monthly*, 108(9), 793-805.
- Mancosu, P. (1996). *Philosophy of mathematical practice in the 17<sup>th</sup> century*. New York: Oxford University Press.
- Marrades, R., & Gutierrez, A. (2000). Proofs produced by secondary school students learning geometry in a dynamic computer environment. *Educational Studies in Mathematics*, 44, 87-125.
- Marriotti, M. A. (2000). Introduction to proof: The mediation of a dynamic software environment. *Educational Studies in Mathematics*, 44, 25-53.
- Martin, W. G., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education*, 10(1), 41-51.
- Martino, A. M., & Maher, C. A. (1999). Teacher questioning to promote justification and generalization in mathematics: What research practice has taught us. *Journal of Mathematical Behavior*, 18, 53-78.
- Marty, R. (1991). Getting to Eureka! Higher order reasoning in math. *College Teaching*, 39(1), 3-6.



- Mercer, N. (1995). *The guided construction of knowledge: Talk amongst teachers and learners*. Philadelphia: Multilingual Matters.
- Morris, A. K. (2002). Mathematical reasoning: Adults' ability to make the inductive-deductive distinction. *Cognition and Instruction*, 20(1), 79-118.
- Mueller, D. J. (1975). Logic and the ability to prove theorems in geometry (Doctoral dissertation, Florida State University, 1969). *Dissertation Abstracts International*, 36, 851A.
- NCTM. (1970). *A history of mathematics education in the United States and Canada*. Thirty-second Yearbook. Washington, DC: Author.
- NCTM. (1973). *Geometry in the mathematics curriculum*. Thirty-sixth Yearbook. Reston, VA: Author.
- NCTM. (1987). *Learning and teaching geometry, K-12*. 1987 Yearbook. Reston, VA: Author.
- NCTM. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.
- NCTM. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- O'Brien, T. C., Shapiro, B. J., & Reali, N. C. (1971). Logical thinking—language and context. *Educational Studies in Mathematics*, 4, 201-209.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.
- Piaget, J. (1973a). *The child and reality: problems of genetic psychology*. New York: Grossman.
- Piaget, J. (1973b). *The language and thought of the child*. London: Routledge and Kegan Paul.
- Piaget, J. (1978). *Success and understanding*. Cambridge, MA: Harvard University Press.
- Platt, J. L. (1967). The effect of the use of mathematical logic in high school geometry: An experimental study [Abstract]. *Dissertation Abstracts International*, 28, 4544-4545A.
- Porteous, K. (1986, July). *Children's appreciation of the significance of proof*. Paper presented at the (tenth) annual Psychology in Mathematics Education conference, London.
- Porteous, K. (1990). What do children really believe? *Educational Studies in Mathematics*, 21, 589-598.
- Porter, A. (1993). Opportunity to learn. Brief No. 7. Madison, WI: Center on Organization and Restructuring of Schools.
- Raman, M. (2003). Key ideas: What are they and how can they help us understand how people view proof? *Educational Studies in Mathematics*, 52, 319-325.
- Recio, A. M., & Godino, J. D. (2001). Institutional and personal meanings of mathematics proof. *Educational Studies in Mathematics*, 48, 83-99.
- Reid, D., & Dobbin, J. (1998). Why is proof by contradiction difficult? *Proceedings of the 22<sup>nd</sup> Psychology in Mathematics Education Conference*, Vol. 4 (pp. 41-48), Stellenbosch, South Africa.
- Rips, L. J. (1994). *The psychology of proof*. Cambridge, MA: MIT Press.
- Rothman, T., & Fukagawa, H. (1998). Japanese temple geometry. *Scientific American*, 278 (5), 84-91.

- Schifter, D., Monk, S., Russell, S. J., Bastable, V., & Earnest, D. (2003, April). Early algebra: What does understanding the laws of arithmetic mean in the elementary grades? Paper presented at the meeting of the National Council of Teachers of Mathematics, San Antonio.
- Schoenfeld, A. H. (1994). What do we know about mathematics curricula? *Journal of Mathematical Behavior*, 13(1), 55-80.
- Selden, A., & Selden, J. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, 29, 123-151.
- Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, 34, 4-36.
- Senk, S. L. (1985). How well do students write geometry proofs? *Mathematics Teacher*, 78(6), 448-456.
- Sfard, A. (1995). The development of algebra—Confronting historical and psychological perspectives. In C. Kieran (Ed.), *New perspectives on school algebra: Papers and discussions of the ICME-7 algebra working group* (special issue). *Journal of Mathematical Behavior*, 14, 15-39.
- Silver, E. A., & Carpenter, T. P. (1989). Mathematical methods. In M. M. Lindquist (Ed.), *Results from the fourth mathematics assessment of the National Assessment of Educational Progress* (pp. 10-18). Reston, VA: NCTM.
- Silver, E. A., & Kenney, P. A. (2000). *Results from the seventh mathematics assessment*. Reston, VA: NCTM.
- Simon, M. A., & Blume, G. W. (1996). Justification in the mathematics classroom: A study of prospective elementary teachers. *Journal of Mathematical Behavior*, 15, 3-31.
- Sowder, L. (2004). University faculty views about mathematics majors' understanding of proof. Unpublished manuscript, San Diego State University.
- Sowder, L., & Harel, G. (2003). Case studies of mathematics majors' proof understanding, production, and appreciation. *Canadian Journal of Science, Mathematics and Technology Education*, 3(2), 251-267.
- Strutchens, M. E., & Blume, G. W. (1997). What do students know about geometry? In P. A. Kenney & E. A. Silver (Eds.), *Results from the sixth mathematics assessment* (pp. 165-193). Reston, VA: NCTM.
- Suppes, P. (1966). The axiomatic method in high school mathematics. In The Conference Board of the Mathematical Sciences, *The role of axiomatics and problem solving in mathematics* (pp. 69-76). Boston: Ginn.
- Tall, D. (Ed.) (1991). *Advanced mathematical thinking*. Dordrecht, The Netherlands: Kluwer.
- Thompson, D. (1991, April). *Reasoning and proof in precalculus and discrete mathematics*. Paper presented at the meeting of the American Educational Research Association, Chicago.
- Tinto, P. (1988, April). *Students' views on learning proof in high school geometry*. Paper presented at the meeting of the American Educational Research Association, New Orleans.
- U.S. Department of Education. (2000). *NAEP 1999 trends in academic progress: three decades of student performance*, by J. R. Campbell, C.M. Hombo, & J. Mazzeo.

- NCES 2000-469. Office of Educational Research and Improvement: Washington, D.C.
- Usiskin, Z. (1987). Resolving the continuing dilemmas in school geometry. In M. M. Lindquist, & A. P. Shulte, (Eds.), *Learning and teaching geometry, K-12* (pp. 17-31). 1987 Yearbook, NCTM. Reston, VA: NCTM.
- Vergnaud, G. (1983). Multiplicative structures. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 127-174). New York: Academic Press.
- Vergnaud, G. (1988). Multiplicative structures. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in the middle grades* (pp. 141-161). Volume 2 of the Research Agenda for Mathematics Education project. Reston, VA: NCTM.
- Vinner, S. (1977). The concept of exponentiation at the undergraduate level and the definitional approach. *Educational Studies in Mathematics*, 8, 17-26.
- Vinner, S. (1983). The notion of proof—some aspects of students' views at the senior high level. In R. Hershkowitz (Ed.), *Proceedings of the Seventh International Conference of the Psychology of Mathematics Education* (pp. 289-294). Rehovot, Israel: Weizmann Institute of Science.
- Vygotsky, L. (1962). Thought and language. Cambridge, MA: MIT Press.
- Vygotsky, L. (1978). Mind in society: The development of the higher psychological processes. Cambridge, MA: The Harvard University Press.
- Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. Cambridge, MA: Harvard University Press.
- Weber, K. (2001). Student difficulty in constructing proof: The need for strategic knowledge. *Educational Studies in Mathematics*, 48, 101-119.
- Wilder, R. (1967). The role of the axiomatic method. *American Mathematical Monthly*, vol. 75, pp. 115-127.
- 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.
- Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, 27, 458-477.
- Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 227-236). Reston, VA: NCTM.
- Yackel, E., Rasmussen, C., & King, K. (2000). Social and sociomathematical norms in an advanced undergraduate mathematics course. *Journal of Mathematical Behavior*, 19, 275-287.
- Zack, V. (2002). Learning from learners: Robust counterarguments in fifth graders' talk about reasoning and proving. In A. Cockburn & E. Nardi (Eds.), *Proceedings of the 26<sup>th</sup> International Group for the Psychology of Mathematics Education*, vol. 4, pp. 434-441. Norwich, UK: University of East Anglia.

---

<sup>1</sup> We wish to acknowledge the helpful comments from Gila Hanna, Carolyn Maher, and Erna Yackel. Preparation of this chapter was supported in part by National Science

---

Foundation Grant No. REC-0310128. Any opinions or conclusions expressed are those of the authors and do not represent an official position of NSF.

<sup>2</sup> To assure a degree of quality control and for practical reasons, we have restricted our survey to writings that have undergone an external review process, except in the case of a few doctoral dissertations that we examined. Examination of at least one case—the teaching and learning of mathematical induction—was omitted here since it has been treated thoroughly elsewhere (Harel, 2001). We regret that no doubt we have unintentionally overlooked other pieces of research or commentary that could have improved this chapter.

<sup>3</sup> All aspects of proof addressed in this paper must always be understood in the context of the learning and teaching of proof. Even when we address mathematical, historical, or philosophical aspects of proof, the goal is to utilize knowledge of these aspects for the purpose of better understanding the processes of learning and teaching of proof. Thus, phrases such as “research on proof,” “perspective on proof,” and the like should always be understood in the context of mathematics education, not of mathematics, history, or philosophy per se.

<sup>4</sup> We are restricting our discussion to classical Greek mathematics and the mathematical developments that grew out of it. The development of deductive reasoning in China, India, and other non-Western cultures is not considered in this paper.