EDUCATIONAL PSYCHOLOGIST, *23*(2), 145–166

When Good Teaching Leads to Bad Results: The Disasters of "Well-Taught" Mathematics Courses

Alan H. Schoenfeld
University of California

This article describes a case study in mathematics instruction, focusing on the development of mathematical understandings that took place in a 10-grade geometry class. Two pictures of the instruction and its results emerged from the study. On the one hand, almost everything that took place in the classroom went as intended—both in terms of the curriculum and in terms of the quality of the instruction. The class was well managed and well taught, and the students did well on standard performance measures. Seen from this perspective, the class was quite successful. Yet from another perspective, the class was an important and illustrative failure. There were significant ways in which, from the mathematician's point of view, having taken the course may have done the students as much harm as good. Despite gaining proficiency at certain kinds of procedures, the students gained at best a fragmented sense of the subject matter and understood few if any of the connections that tie together the procedures that they had studied. More importantly, the students developed perspectives regarding the nature of mathematics that were not only inaccurate, but were likely to impede their acquisition and use of other mathematical knowledge. The implications of these findings for reseach on teaching and learning are discussed.

INTRODUCTION AND OVERVIEW

This article reports the results of a year-long intensive study of the teaching and learning that took place in a 10th-grade geometry class, which will be called the target class. The class took place during the 1983–1984 academic

Requests for reprints should be sent to Alan H. Schoenfeld, Departments of Education and Mathematics, University of California, Berkeley, CA 94720.

year in a highly regarded suburban school district in upstate New York. The study included periodic observations of the target class and of 11 other mathematics classes, interviews with students and teachers, and questionnaire analyses of students' perspectives regarding the nature of mathematics. The target class was observed at least once a week, and was videotaped periodically for subsequent detailed analysis. Two weeks of instruction near the end of the course, dealing with locus and construction problems in geometry, were videotaped in their entirety. Our analyses focused both on the mathematics that was learned, and on what the students learned about the mathematics—including how and when they would use, or fail to use, the mathematics that they had studied.

We begin with a brief discussion of the literature relevant to the work described in this article. Due to space constraints, the discussion of the literature on teaching is extremely telegraphic. A major point, however, is that only a small portion of that literature is directly relevant to the present study. Most of the literature relegates the subject matter being taught to the status of context variable and does not, therefore, discuss it in great depth. Subject-matter understanding, and the influence of classroom practice on the development of that understanding, are the focus of this article. In the literature section we explore the difference between becoming competent at performing the symbolic manipulation procedures in a mathematical domain and grasping the underlying mathematical ideas in that domain. Our discussion suggests that broader definitions of mathematical understanding than "mastering symbol manipulation procedures" should be used, and that there are dangers to the narrow assessments of competency that are currently employed.

The article then turns to a description of the main study. Most of the discussion is qualitative and interpretive, although some data from the questionnaire analyses are given. It is argued that, despite the fact that the class was well taught and that the students did well on the relevant performance measure (the New York State Regents examination), the students in the target class learned some inappropriate and counterproductive conceptualizations of the nature of mathematics as a direct result of their mathematics instruction. Detailed descriptions of the instructional origins of these counterproductive notions are given. Finally, some suggestions regarding directions for subsequent research are considered.

THE CONTEXT: BRIEF COMMENTS ON THE TEACHING LITERATURE AND ON UNDERSTANDING MATHEMATICS

The past decade has seen a radical shift in theories of learning, brought in large part by progress in the cognitive sciences. Through perhaps the mid-1970s, learning theories were for the most part domain-independent. Such theories attempted to characterize general principles of learning, the specifics of which were hypothetically applicable in different domains such as reading, social studies, and mathematics. The details of the subject matter were not important in such theories, for the most part playing the role of "context variables" that (as the theory had it) could be taken into account in experimental design. The corresponding paradigms for investigating research on teaching are described by Corno (this issue); see also Doyle (1978), Dunkin and Biddle (1974), and Shulman (1986). The first major paradigm described by Corno, process-product research, largely used correlational methods to explore relationships between teacher classroom behavior and student learning. The classroom behaviors explored were, for the most part, straightforward and easily quantifiable (e.g., time spent in questioning, "active learning time," amount of praise, amount of feedback). Other classroom variables included type of ability grouping, whether students worked in small or large groups, and so on. Learning was operationally defined as performance on achievement tests—tests which, as wel shall see later, may fail in significant ways to measure subject-matter understanding.

Mediating process research (see, e.g., Corno, this issue) provides a means of overcoming some of the significant limitations of the process-product paradigm. Such work signals the beginning of a rapprochement with cognitive science research on learning, specifically with its focus on the child as active interpreter of its experience. Doyle's study in this issue provides some compelling examples of the importance of this perspective. Doyle suggests that the presentation of subject matter as familiar work—routinized exercises that can be worked out of context and without significant understanding of the subject matter—can trivialize that subject matter and deprive students of the opportunity to understand and use what they have studied. That suggestion is explored in length in this article.

Recent cognitive research on learning diverges from the domain-independent work described earlier in that it lays a much greater emphasis on the particulars of the subject matter being studied. In elementary arithmetic, for example, Brown and Burton (1978) developed a diagnostic test that could predict, about 50% of the time, the incorrect answers that a particular student would obtain to a subtraction problem—before the student worked the problem! The literature indicates that misconceptions in arithmetic, in algebra, in physics, and other domains, are quite common and consistent (see, e.g., Helms & Novak, 1984). From this and related work follow two main consequences. The first consequence is that one of the treasured pedagogical principles on which much current instruction is based is, if not plain wrong, certainly inadequate. The predominant model of current instruction is based on what Romberg and Carpenter (1986) called the absorption theory of learning: "The traditional classroom focuses on competition, management, and group aptitudes; the mathematics taught is assumed to be a fixed body of knowledge, and it is taught under the assumption that learners absorb what has been covered" (p. 26). According to this view, the good teacher is the one who has 10 different ways to say the same thing; the student is sure to "get it" sooner or later. However, the misconceptions literature indicates that the students may well have "gotten" something else—and that what the student has gotten may be resistant to change. Dealing with this reality calls for a significantly different perspective on the part of the teacher. It also calls for different perspectives regarding the appropriate domain of study of research on teaching, and different measures of competence. The second consequence is that one must look at the subject matter in detail. Arithmetic mistakes differ from misconceptions in algebra and physics, and from misapprehensions about reading; we will understand each of these only by studying it on its own terms. Thus studies of learning and teaching in particular subject areas must be grounded in analyses of what it means to understand the subject matter being taught. It is to that kind of analysis in mathematics that we now turn. Some relevant research on mathematical cognition and teaching may be found in Leinhardt and Smith (1984), Resnick (1983), Romberg and Carpenter (1986), Schoenfeld (1985), and Silver (1985).

The issue of classroom practice and its relation to students' understanding of mathematical structure was one of the main themes of Wertheimer's (1959) *Productive Thinking,* which provides our first two examples. In the first, Wertheimer asked elementary school students to solve problems like

$$\frac{274+274+274}{3} \quad \text{or} \quad \frac{812+812+812+812+812}{5}$$

Many of the students, who were fluent in all four of the basic arithmetic operations, solved such problems by laboriously adding the terms in the numerator and then performing the indicated division. By virtue of obtaining the correct answer, the students indicated that they had mastered the procedures of the discipline. However, they had clearly not mastered the underlying substance: If you see repeated addition as equivalent to multiplication and you see division as the inverse of multiplication (i.e., the multiplication and division by the same number cancel each other out), there is no need to calculate at all. This example illustrates that being able to perform the appropriate algorithmic procedures, although important, does not necessarily indicate any depth of understanding. (We note here that virtually all standardized testing for arithmetic competency—and, de facto, much standard instruction in arithmetic—focuses primarily if not exclusively on procedural mastery.)

Wertheimer's more famous example comes from his observations of classroom sessions devoted to "the parallelogram problem," the problem of determining the area of a parallelogram of base B and altitude H. The students had been taught the standard procedure, in which cutting off and moving a specific triangle converts the parallelogram to a rectangle whose area is easy to calculate. They did quite well at the lesson, and they were able to reproduce the argument in mathematically correct form. But when Wertheimer asked the students to find the area of a parallelogram in nonstandard position, or to find the area of a parallelogramlike figure to which the same argument applied, the students were stymied. Wertheimer argued that although the students had memorized the proof, they had failed to understand the reason that it worked; although they had memorized the formula, they used it without deep understanding. With that understanding, he argued, the students would have been able to answer his questions without difficulty; without it they could solve certain well-specified exercises but in reality had acquired only the superficial appearance of competence. (We note again that typical achievement tests, which examine students' ability to reproduce the standard arguments, are unlikely to examine the kinds of understandings Wertheimer considered fundamental.)

There are numerous contemporary parallels to these examples. For instance, word problems of the following type are a major focus of the elementary mathematics curriculum: "John has eight apples. He gives five to Mary. How many apples does John have left?" Perhaps the most commonly used instructional procedure to help students solve such problems is the "key word procedure," which is used as follows. The student is told that certain words in problem statements provide the "key" to selecting which arithmetic operation to employ. For example, the key word in the problem just quoted is *left,* which indicates subtraction. One can "solve" the problem by identifying the two numbers in the problem statement, and then—because the key word is *left*—subtracting one from the other. Note that one can do so without even reading the whole problem, and without understanding the situation it describes. Research indicates that many students work the problems in precisely that fashion. In interviews some students revealed that they circled the numbers in the problem statement and then read the problem statement from the last word backwards, because the key word usually appears near the end of the problem! Thus the key word procedure, initially introduced to help students make sense of word problems, had (at least in these cases) precisely the opposite effects. It allowed students to obtain the right answers without understanding—and gave them the option of not seeking understanding at all. Worse, it may have suggested to them that understanding is not necessary when solving mathematics problems; one simply follows the procedure, whether it makes sense or not.

The most extensive documentation of students' performance on word problems, without understanding, comes from the third National Assessment of Educational Progress (NAEP; Carpenter, Lindquist, Matthews, & Silver, 1983). On the NAEP mathematics exam, which used a stratified national sample of 45,000 students, 13-year-olds were given the following problem: "An army bus holds 36 soldiers. If 1,128 soldiers are being bused to their training site, how many buses are needed?" Seventy percent of the students who worked the problem performed the long division algorithm correctly. However, 29% of the students wrote that the number of buses needed is "31 remainder 12" and another 18% wrote that the number of buses needed is 31. Only 23% gave the correct answer. Thus fewer than one third of the students who selected and carried out the appropriate algorithm produced the right answer—a step that required a trivial analysis of the meaning of the problem statement. There are a number of plausible explanations for this behavior, one of which will be suggested in the case study described in the sequel (see also Silver, 1986, for a discussion of related problems). But data of this type document an almost universal phenomenon: Students who are capable of performing symbolic operations in a classroom context, demonstrating "mastery" of certain subject matter, often fail to map the results of the symbolic operations they have performed to the systems that have been described symbolically. That they fail to connect their formal symbol manipulation procedures with the "real-world" objects represented by the symbols constitutes a dramatic failure of instruction.

A set of similar phenomena motivated the present study. Since 1979 I have conducted a series of studies exploring students' understandings of geometry. Those studies have focused, in particular, on the relationship between geometric proofs and geometric constructions. To summarize, I had found that high school and college students who had taken a full year of high school geometry, which focuses on proving theorems about geometric objects, uniformly approached geometric construction problems as empiricists. They engaged in empirical guess-and-test loops, completely ignoring their proof-related knowledge. In one series of interviews, for example, college students were asked to work two related problems. The first was a proof problem. Solving this problem directly provided the answer to the second, a construction problem (the second problem asked how to construct a circle whose properties had been completely determined in the first). Yet, after solving the first problem, nearly a third of the students began the second problem by making conjectures that flatly violated the results they had just proved!

Such behavior indicated that these students saw little or no connection between their "proof knowledge," abstract mathematical knowledge about geometric figures obtained by formal deductive means, and their "construction knowledge," procedures and information they had mastered in the very same class for working straightedge and compass construction problems. I make this statement more provocatively as Belief 1 given below; some other typical beliefs are also given. I conjecture students may develop these beliefs as a result of their experiences with mathematics. (Extended discussions of the students' beliefs may be found in Chapters 5 and 10 of Schoenfeld, 1985. A discussion of the "ideal" relationship between geometric empiricism and deduction may be found in Schoenfeld, 1986).

Belief 1: The processes of formal mathematics (e.g., "proof") have little or nothing to do with discovery or invention. Corollary: Students fail to use information from formal mathematics when they are in "problem-solving mode."

Belief 2: Students who understand the subject matter can solve assigned mathematics problems in five minutes or less. Corollary: Students stop working on a problem after just a few minutes because, if they haven't solved it, they didn't understand the material (and therefore will not solve it).

Belief 3: Only geniuses are capable of discovering, creating, or really understanding mathematics. Corollary: Mathematics is studied passively, with students accepting what is passed down "from above" without the expectation that they can make sense of it for themselves.

In listing these beliefs we note the parallel to research Doyle describes in this issue. Doyle describes a student who took a teacher's instructions for an assignment as a recipe for completing the task, rather than a way of learning the material. In terms more provocative than Doyle might like, one can characterize that student's perspective as follows:

Belief 4: One succeeds in school by performing the tasks, to the letter, as described by the teacher. Corollary: Learning is an incidental by-product to "getting the work done."

The purpose of the following study was to explore the presence and robustness of such beliefs, and to seek their possible origins in mathematics instruction.

THE INSTRUCTIONAL ROOTS OF STUDENTS' VIEWS OF MATHEMATICS

At the beginning of the school year the teacher of the target class issued an open invitation. My assistants and I were welcome to visit any of his classes, any time, without prior notice. We were free to videotape any of the classes for later analysis, and interview any of his students if mutually convenient times could be arranged. One class was chosen, and it was observed once or more a week for the entire year. Two weeks of instruction, focusing on locus and construction problems, were videotaped and analyzed in detail. An 80-item questionnaire was filled out by the 20 students in the target class. It was also completed by 210 other students in 11 other classes. Those classes were also observed periodically, to determine whether the students and instruction in the target class could be considered typical.

Both the teacher and the students would be ranked well according to any of the measures typically employed in classroom research. To begin with a performance measure, the class scored in the top 15% on the New York State Regents geometry examination, a statewide uniform examination given for the course. The class was well run. Early in the term the teacher established the rules of protocol for classroom interactions, and they were adhered to throughout the term. The relationship between teacher and students was cordial and respectful, and discipline was never a problem in class. Control was maintained in a low-key manner, with humor. Straight lecture was kept to a minimum. The vast majority of classes focused on working problems, with students presenting their solutions at the board. Such discussions were usually Socratic, with the teacher leading the student to the correct answer if he or she had not obtained it. Questions were invited, and reliable feedback was given. During such sessions the students were attentive; the class would do well on standard measures such as "time on task."

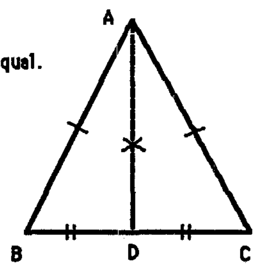
A classroom observer unfamiliar with mathematics would necessarily give the class high marks. The teacher followed the curriculum and ran the class well, and the students scored well on standardized examinations. Even so, some very unhealthy things took place in the course. In particular, there is strong reason to believe that, as a direct result of their experience in the course, the students developed (or, at least, were reinforced in) the kinds of beliefs described in the previous section. In the sequel, four aspects of the classroom interactions, and their results, are described. The first is described at some length, the others more briefly.

First Result: Students Learn to Separate the Worlds of Deductive and Constructive Geometry

The subject matter covered in the course was prescribed by the Math 10 Regents curriculum, and the curriculum was followed quite closely. Despite the amount of freedom hypothetically allowed to them, most teachers stick pretty close to the textbook (Romberg & Carpenter, 1986). In New York, strict adherence to the curriculum was even more likely because of the statewide exam. Performance on the Regents examination is the primary measure of both teacher and student success in almost all New York school districts. The primary goal of instruction, therefore, was to have students do well on the exam. The curriculum and the examinations were well established and quite consistent from year to year. Thus the amount of attention to give to each topic, and the way to teach it (for "mastery" as measured by the exams), were essentially prescribed.

The curriculum contained a dozen "required" proofs, one of which appeared on the Regents exam and was worth 10 points (of 100). An example of a required proof, written in the standard two-column format, is given in Figure 1. Many other problems on the exam called for proof skills. Proof was thus central to the curriculum, and it received great emphasis. In

Theorem: If two sides of a triangle are equal, the angles opposite those sides are equal.



Given: AB = AC

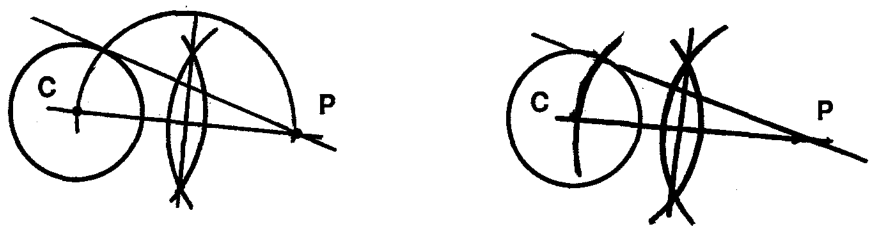
Prove: ∠ABD = ∠ACD

Statements	Reasons
1. AB = AC	1. Given
2. Draw the median AD to side BC of the triangle	2. Construction Principle
3. BD = CD	3. Definition of median
4. AD = AD	4. Reflexive property
5. △ABD ≅ △ABC	5. SSS*
6. ∠ABD = ∠ACD	6. CPCT●

* SSS is the customary abbreviation for the postulate "If the three sides of one triangle are congruent to the three sides of a second triangle, then the two triangles are congruent."

● CPCT is the customary abbreviation for "corresponding parts of congruent triangles are congruent."

Figure 1 A typical two-column proof.



A. A correct answer

(Note that no explanation is required.)

B. An incorrect answer

Problem: To construct, using straightedge and compass, a tangent to the given circle C that passes through the given point P

Figure 2 A typical construction problem.

contrast, constructions were not emphasized. One "required" construction, worth only 2 points, appeared on the end-of-year exam. A typical construction problem, with a correct solution, is given in Figure 2. What the exam required, in fact, was that the student perform the construction accurately. No explanations were asked for; the student received full credit for producing the right sequence of arcs on the page with a reasonable degree of precision. Developing this ability was the focus of classroom instruction.

In the unit of constructions—taught, incidentally, near the end of the school year so the students would not have time to forget their recently acquired skills before the Regents exam—the training the students received was geared toward the mastery of a physical rather than an intellectual skill. (The "bottom line" was quite clear: If the construction was inaccurate although correct, the student did not receive credit.) The reliance on empirical standards was made clear at the beginning of the unit. The students had been told to bring their construction tools (straightedge and compass) to class. Those who arrived in class without them were sent to their lockers to pick them up, and a discussion of constructions did not begin until all of the students had the tools in hand. From the very beginning, the constructions were taught as a step-by-step procedure to be memorized. For example, here is the beginning of the teacher's presentation of a new construction, to copy an angle.

> Construction Number 2 is [the following]: Given an angle, to construct an angle equal, or congruent, to that angle.
>
> *Step 1,* choose any point on your work line. That's going to be the vertex of your new angle. *Step 2,* mark an arc on the given angle with the point of the compass on the vertex formed by the two intersecting lines. *Step 3,* move over

> to the work line, and without changing the compass setting, make an arc with the point of the compass at the point that's going to be the vertex of your new angle . . .

This introduction set the tone for the classroom discussions of constructions. Most of the time the teacher introduced a construction and demonstrated it at the blackboard. As he demonstrated the steps involved, the students copied the constructions into their notebooks. When students demonstrated their solutions to homework problems at the board, the other students carefully copied the constructions into their notebooks. Approximately 90% of classroom time during the unit on constructions was spent with straightedge and compass in hand, practicing the constructions.

The first day of the unit the students were quite slow at the constructions. The teacher told them that they would have to get faster, but that they would not find it difficult: "Mainly with constructions it is all going home and practicing." The idea that practice is essential, and that the students should have the constructions committed to memory, was repeated throughout the unit. The message to the students was quite clear, as indicated by a comment the teacher made just before a test: "You'll have to know all your constructions cold. . . . This is where practice at home comes in."

Another indication of the empirical focus during the unit on constructions came at the beginning of a class period. The class started in typical fashion. The teacher sent six students to the board, to present their solutions to homework problems. He then realized that there was a problem: "I forgot we're doing constructions. For constructions we use a straightedge and compass, and I only have one compass." He sent five of the students back to their seats, and had the first student demonstrate her construction at the board with the appropriate tools. The rest then followed, one at a time, each student working his or her construction at the blackboard with exquisite care. The students in their seats copied the construction into their notebooks with comparable care. This session took 20 min, with the exclusive focus being on the accuracy of the constructions.

This may seem to be a minor matter, but it is not. All six of the students could have sketched their answers to the problems at the beginning of class; when called upon, they could have discussed why their constructions were right. Such a discussion could have focused on understanding and could have helped to link the formal mathematics with the empirical and constructive aspects of it. Instead, the discussion focused on carefully performing the sequence of steps that constituted the constructions. The "message" the students got is that accuracy is what counts. They did not get to see the overt linkage of the construction unit with the subject matter they had studied for most of the term.

The goals of instruction were accuracy and speed, and understanding was sacrificed (not intentionally!) in the process. Class time was spent almost exclusively in practicing the sequence of steps required to complete a construction. Did a construction have to be accurate? "Just line up your points. That's the important thing." What standards of accuracy would be used to grade constructions on a test? "As long as I can see all the marks so that I can follow your construction, and the construction is correct, then I will not take off if it's just off by a minimal distance. . . . But more than that and I will take off, yes."

The whole tenor of the course led students to rely on empirical standards for the correctness of a construction. The strongest example of such emphasis, to the exclusion of the formal mathematics that *guarantees* the correct answer, came when the teacher presented a quick trick for distinguishing between the constructions that yield the inscribed and circumscribed circles of a given triangle. In the class there was no discussion whatsoever of why each of the constructions worked. Rather, the teacher told the students that they would have trouble remembering which constructions use angle bisectors and which use perpendicular bisectors. The following method, he suggested, would help them decide. They should draw a scalene triangle and sketch in its angle bisectors. They could then check the lengths from the point of intersection to the three vertices. Those segments would have to be equal for the circumscribed circle. But in the sketch it will be obvious that those lines are not equal. That resolves the issue: The bisectors must yield the inscribed circle, because they don't yield the circumscribed circle. Of course, if the students don't remember and choose the wrong construction "It will be obvious to you right away if you made a mistake because that circle is going to come out either way inside or way outside." Here, too, the message was that accuracy is the final arbiter of correctness. (See pp. 361–367 of Schoenfeld, 1985, for details.)

With the stress on accuracy, the students learned to rely exclusively on empirical standards to assess the correctness of their construction attempts. They accepted the "bottom line" of the evaluation scheme and came to believe that a construction is right only if it is cleanly carried out and produces results that are empirically correct (i.e., within the bounds of error considered acceptable for the tools one has for performing the construction).

Returning to the statement of Belief 1, what if anything—from the student's point of view—does mathematical argumentation, or proof, have to do with constructions? In brief, virtually nothing. In these students' experience, proofs had always served as confirmation of information that someone (usually the teacher or mathematicians at large) already knew to be true; they provided the "justifications" for constructions. But ask these students to discover a construction, and they do not see that any proof arguments are relevant at all. For these students, a construction is right when it "works." They are in "discovery mode," and proofs have never helped them to discover. Confronted with a construction problem they make their best guess, and then test it by trying it out and seeing if their attempt meets their empirical standards. Such behavior was learned, alas, as an unintended by-product of their instruction.

Second Result: Students Perceive That the Form of a Mathematical Answer Is What Counts

For mathematicians, a "proof" is a coherent chain of argumentation in which one or more conclusions are deduced, in accord with certain well-specified rules of deduction, from two sets of "givens": (a) a set of hypotheses and (b) a set of "accepted facts" consisting of either axioms or results that are known to have been proven true. Save for certain domain-specific constraints (e.g., formal proofs in certain branches of logic that follow rigidly prescribed formats), there is a great deal of flexibility in the way a proof argument can be written. Put simply, what matters to the mathematical community is the argument's coherence and correctness. Consider, for example, the following proof that the base angles of an isosceles triangle are equal.

> Consider a triangle ABC, where AB = AC, as in the diagram at the top of Figure 1. Draw the median AD to the midpoint D of side BC. Because AB = AC, AD equals itself, and BD = DC, the triangles ABD and ACD are congruent. Hence the corresponding angles ABD and ACD are equal.

Most mathematicians would consider this argument to be fully adequate and appropriate proof of the desired result. In the vast majority of high school geometry classrooms across the country, however, that argument would not be accepted as either adequate or appropriate. In most 10th-grade geometry classes there is a strict protocol for writing proofs, and students are expected to follow the protocol closely. One begins by stating the problem at the top of the page, listing what is given and what is to be proved. One then draws a large *T*, which divides the space below the problem statement into two columns. The column on the left is labeled "Statements," and the column on the right is labeled "Reasons." In the left-hand column one writes a series of statements, beginning with the "givens" from the problem statement. These statements are numbered in order, with only one statement per line. Each statement must be justified separately. The right-hand column contains the justifications, which are numbered to correspond to the statements. The last entry on the list of statements, of course, is the result to be proved. Figure 1, discussed earlier, gives a "correct" version of the proof.

The two-column format is a matter of convention. In its defense, there are arguments in favor of that convention; for example, an organized format may help students to be orderly and to keep track of their arguments. (There are also arguments that the rigidity of format obscures the thought processes that provide good mathematical arguments, and that alternate formats are superior; see, e.g., Anderson, Boyle, Franklin, & Reiser, 1985.) Either way, however, there is nothing sacred about the form; it is simply an agreed-upon means of communication. In many high school classrooms, however, it has taken on nearly sacred status. In most classes—in particular, in all of the classes we observed, including the target class—and on the New York State Regents exam, arguments must be expressed in that form. In instruction, and on the examinations, substantial partial credit is deducted for deviations from the proper form.

The use of such a form for proof strikes most students as being arbitrary and capricious. The teacher usually overcomes this initial resistance, but at the cost of a significant amount of time and effort. At the beginning of the term the form is presented as something that must be used, something that the students will simply have to get accustomed to. When students put their homework on the board during the first few weeks of the unit on proofs, a tremendous amount of time is spent correcting the form of the students' answers. There are discussions of completeness ("Do we really have to list the givens? After all, they appear in the problem statement." Answer: Yes). There are discussions of format ("if more than one statement has the same reason, can't we write them on the same line?" Answer: No). And there are discussions of acceptable abbreviations ("Do we have to write out the whole statement about corresponding parts of congruent triangles?" Answer: No, you can use CPCT). A great deal of class time is spent in the consideration of what is legitimate and acceptable. In one of our videotapes, for example, 22 of the 37 min spent discussing the students' blackboard work was spent in discussions of form rather than on the correctness of the students' work.

The focus on form decreased as the students mastered it, but by that time the emphasis had had its effect. In a number of our videotapes of high school and college students, the students solved a proof problem in 2 or 3 min. They then made comments like "All right, now let's do it properly" and spent as long as 10 to 15 min making sure that the argument was written in correct form, with all of the proper abbreviations. As a result of their instruction these students came to believe that it is the form of expression, as much as the substance of the mathematics, that is important. That was not a good lesson for them to have learned.

Third Result: Students Come to Believe That All Problems Can Be Solved in Just a Few Minutes

The structure of homework assignments and of test problems in the target class was essentially identical to the structure of homework assignments and test problems in the other classes we observed. This structure, in turn, was essentially the same (making allowances for students' age) throughout the whole school system. Over the period of a full school year, none of the students in any of the dozen classes we observed worked mathematical tasks that could seriously be called problems. What the students worked were exercises: tasks designed to indicate mastery of relatively small chunks of subject matter, and to be completed in a short amount of time. In a typical 5-day sequence, for example, students were given homework assignments that consisted of 28, 45, 18, 27, and 30 "problems," respectively. (A typical problem was as follows. The students had learned the following construction: Given a point and a line, construct a line passing through the given point that is parallel to the given line. The problem: Given a triangle and a point marked on one side of the triangle, construct a line passing through the given point that is parallel to the base of the triangle. Solving this problem calls only for recognizing that it is identical to the known construction, and applying the same procedure.) The teacher's practice was to have students present solutions to as many of the homework problems as possible at the board. Given the length of his assignments, that means that he expected the students to be able to work 20 or more "problems" in a 54-min class period. Indeed, the unit test on locus and construction problems (a uniform exam in Math 10 classes at the school) contained 25 problems—giving students an average 2 min and 10 sec to work each problem. The teacher's advice to the students summed things up in a nutshell: "You'll have to know all your constructions cold *so you don't spend a lot of time thinking about them* [emphasis added]."

More time was spent on proof problems, of course, but even those were expected to be worked with dispatch; 10 or 15 min was as long as a student was expected to spend on any problem. Once again, we note that this behavior was not a local aberration. The New York State Regents exam, a 3 hr-long exam, contains an average some 30 to 40 questions (including proof questions). New York is neither alone nor atypical in this regard. As a standard of reference one may consider standardized tests of any sort, such as Student Aptitude Tests (SATs) or the College Board examinations.

In sum, students who have finished a full 12 years of mathematics have worked thousands upon thousands of "problems," virtually none of which were expected to take the students more than a few minutes to complete. The presumption underlying the assignments was as follows: If you understand the material, you can work the exercises. If you can't work the exercises within a reasonable amount of time, then you don't understand the material. That's a sign that you should seek help.

Whether or not the message is intended, students get it. One of the open-ended items on our questionnaire, administered to students in 12 high school mathematics classes in Grades 9 through 12, read as follows: "If you understand the material, how long should it take to answer a typical homework problem? What is a reasonable amount of time to work on a problem before you know it's impossible?" Means for the two parts of the question were 2.2 min (n = 221) and 11.7 min (n = 227), respectively. (The two different values of n arose because of a small number of nonquantitative responses like "a few.")

I do not wish to suggest that all mathematics problems should be long, time-consuming blockbusters. There is a role for both exercises and exploratory problems. Students must learn basic facts and procedures, of course, but it is also essential for them to engage in real mathematical thinking—in trying to make progress on difficult problems, in engaging in the give-and-take of making sense of complex situations, in learning that some problems take time, hard work, and a bit of luck to solve. We have done a serious disservice to any student who emerges from the classroom thinking that mathematics only applies to situations that can be solved in just a few minutes—and that if you can't solve a problem in a short amount of time, you should simply give up.

Fourth Result: Students View Themselves as Passive Consumers of Others' Mathematics

One of the most vivid memories of my education comes from an upper-division probability class, when my instructor was about to introduce the binomial theorem. She stopped writing the statement of theorem at the point where she needed to write the formula. "I never remember this formula," she said, "but it's so easy to derive that you don't need it anyway." Then she showed us how to derive the formula. What she showed us made sense. To this day I can't remember the formula, but I can derive it, either when I need it (which is rare) or because the thought of it brings back pleasant memories. The idea that was brought home in that class—that mathematics really makes sense, and that you can figure something out if you need to—was exhilarating. It is (or should be) part of the pleasure of learning mathematics.

Such moments were rare in my experience as a student, and they were almost completely absent from the classes we observed. The mathematics instruction that we observed consisted almost exclusively of training in skill acquisition. For each of the years K through 12 (and beyond; calculus instruction in college is pretty much the same), there was an agreed-upon body of knowledge, consisting of facts and procedures, that comprised the curriculum. In each course, the task of the teacher was to get students to master the curriculum. That meant that subject matter was presented, explained, and rehearsed; students practiced it until they got it (if they were lucky). There was little sense of exploration, or of the possibility that the students could make sense of the mathematics for themselves. Instead, the students were presented the material in bite-sized pieces so that it would be easy for them to master. As an example, recall the step-by-step procedure for constructions, described earlier, that was used by the teacher of the target class. Constructions were introduced that way, and students were given practice that way. When, for example, a student had difficulty with a particular problem, the teacher reminded him that the problem called for a construction with which the student was familiar. He then asked: "In your construction, what is Step 1?" The student replied correctly. The teacher continued. "Good. In your construction, what is Step 2?" And so on. In this way, students got the clear impression that someone else's mathematics was theirs to memorize and spit back. Nor was step-by-step memorization limited to constructions. Recall that the Regents' exam had required proofs as well; students were told to commit them to memory. This was standard practice, and was promoted as being both efficient and desirable. For example, an advertisement for a bestselling series of review books for the Regents' exams proudly announced: "Students like these books because they offer step-by-step solutions."

The point I wish to stress here is that students develop their understanding of the mathematics from their classroom experience with it. If the "bottom line" is error-free and mechanical performance, students come to believe that that is what mathematics is all about. In the target class, for example, the teacher talked about how important it was for students to think about the mathematics and to understand it. He pointed out the fact that they should not memorize blindly, because if they did "and forgot a step" they would be in trouble. In truth, however, this rhetoric—in which the teacher honestly believed—was contradicted by what took place in his classroom. The classroom structure provided reinforcement for memorization, and the reward structure promoted it. One of the items on our questionnaire, for example, asked students to agree or disagree with the statement "The math that I learn in school is mostly facts and procedures that have to be memorized." With a scale ranging from *very true* (1) to *not at all true* (4), this item received an average score of 1.75—the third strongest "agree rating" of 70 questions. Yet the statement "When I do a geometry proof I get a better understanding of mathematical thinking" received an average score of 1.99—again, very strong agreement. These data parallel the NAEP secondary school data, where students claimed that mathematics is mostly memorization but that mathematics helps a person to think logically. Our classroom observations supported Carpenter et al.'s (1983, p. 657) conjecture that the "latter attitudes may reflect the beliefs of their teachers or a more general view rather than emerge from their own experience with mathematics." More importantly, the latter attitudes did not influence behavior: When working mathematics problems, the students behaved in accord with the three mathematics beliefs discussed earlier.

DISCUSSION

The preceding section is not intended in any way as an indictment of the teacher of the target class. The teacher did his job—as the system defines it—effectively. Although one could find aspects of his teaching to criticize (e.g., he might have provided better motivation for some of the units), these complaints would be minor. The class was efficiently managed, there was mutual respect between him and his students, and the students learned what they were supposed to (as measured by the Regents exam). The teacher was not considered a "star," but he was well enough regarded by his colleagues.

In short, what we observed was typical, if not better than average, instruction—with typical results. Our observations of other mathematics classes, Grades 9 through 12, showed little variation in substance from class to class; the questionnaire responses from the target class regarding attitudes toward mathematics were similar to those for the 11 other classes, and similar to the pattern of responses on the NAEP exam. Our discussion has been of a geometry course, and has focused on some results particular to geometry. However, it should be stressed that our results apply across the board to mathematics instruction; geometry is just a case in point. The illustrations in the literature review—Wertheimer's (1959) complaints about students' lack of understanding of the four basic operations despite their ability to perform them and his discussion of parallelogram problem, the abuses of the key word procedure, and students' nonsense answers to word problems on the NAEP exam—all serve as indicators that the problems discussed here are widespread, and that they permeate the curriculum at all levels. Indeed, the NAEP report (Carpenter et al., 1983) suggested that we should not take much solace in students' slightly improved abilities at role computational problems.

> [S]tudents may not understand some of the problems they do solve. Most of the routine problems can be mechanically solved by applying a routine computational algorithm. In such problems the students may have no need to understand the problem situation, why the particular computation is appropriate, or whether the answer is reasonable. . . . The errors made on several of the problems indicate that students generally try to use all of the numbers

> given in a problem statement in their calculation, without regard for the relationship of either the given numbers or the resulting answers to the problem situation. (p. 656)

I believe that the issues raised in this article are general, and that the causes of the behavior discussed here are systemic. Mathematics curricula have been chopped into small pieces, which focus on the mastery of algorithmic procedures as isolated skills. Most textbooks present "problems" that can be solved without thinking about the underlying mathematics, but by blindly applying the procedures that have just been studied. Indeed, typical classroom instruction (recall the key word procedure) subverts understanding even further by providing methods for solving problems that allow students to answer problems correctly, without making an attempt to understand them. Good teaching practice can compensate for the inadequacies of the texts, of course. There is evidence to suggest, however, that it does not. In reviewing a series of case studies on mathematics instruction, Romberg and Carpenter (1986) noted that the textbook was consistently seen as

> the authority on knowledge and the guide to learning [and that] . . . many teachers see their job as "covering the text." Further, it was also noted that mathematics and science were seldom "taught as scientific inquiry—all subjects were presented as what experts had found to be true." Ownership of mathematics rests with the textbook authors and not with the classroom teacher. (p. 25)

Other than texts, the major force that drives the curriculum is testing. In New York, the Regents' exams dominate instruction at all levels; students and their teachers are judged by classroom performance on the examinations. Around the country, school district administrators are coming to rely more on achievement tests as measures of their teachers' classroom performance. Unfortunately, virtually all standardized examinations are insensitive to the kinds of issues discussed in this article. (The busing question cited from the NAEP exam is atypical. Indeed, one highly respected mathematics educator complained to me that such questions are unfair. After all, he said, such problems were not in the curriculum, so one would not expect students to do well on them. With all due respect, I reject this as an outrageous assertion. When 29% of our students indicate that the number of buses required for a particular task is "31 remainder 12," something has gone very, very wrong.)

It is quite reasonable to expect teachers to rely on the text materials they have at hand, and to emphasize the skills on which the students (and they!) will ultimately be judged—most often the isolated, mechanical, algorithmic procedures on which the examinations focus. One notes that until quite recently the literature on teaching research has not been terribly helpful in this regard: In general, the process-product paradigm and mediating process research have used achievement tests (or similarly constructed measures) as their measures of instructional "success." A much broader view of mathematics, of curricular goals, and of what students really learn in their instruction, is needed in order to conceptualize and effect change. In this final discussion I shall make some basic assumptions regarding the nature of mathematics and the nature of humans as learners. These, in turn, suggest some directions for research and practice.

My first assumption is simple, though perhaps not uncontroversial: Assumption 1 – A major purpose of mathematics instruction is to help students learn to think mathematically. To elaborate, my assumptions about mathematics are as follows: Assumption 2 – Even at the most elementary levels, mathematics is a complex and highly structured subject; Assumption 3 – Thinking mathematically consists not only of mastering various facts and procedures, but also in understanding connections among them; and Assumption 4 – Thinking mathematically also consists of being able to apply one's formal mathematical knowledge flexibly and meaningfully in situations for which the mathematics is appropriate. Finally, I make some (basically constructivist) assumptions about humans as learners: Assumption 5 – Students are active interpreters of the world around them, constantly building interpretive frameworks to make sense of their experiences; and Assumption 6 – those interpretive frameworks shape the ways that students see the world and act in it – in particular, how they see and use their mathematical knowledge.

If one takes them seriously, the first four assumptions regarding the nature of mathematical thinking call for a reexamination of curricular goals, materials, and measurement tools. On the one hand, it is pretty clear what mathematical thinking is not: the rote memorization of facts and procedures as often practiced in our classrooms, and as reified by current texts and examinations. Defining the appropriate replacements will be no easy task, however. In essence, Assumptions 2 through 4 serve as major items in a research agenda. It is incumbent upon the research community to provide detailed elaborations of the nature of mathematical thinking – to characterize the knowledge and cognitive processes that comprise thinking mathematically, and to describe the cognitive structures that support such thinking. Recent work in cognitive science has begun to take some steps in that direction, but we are still very much in the early stages of such work – research which will require grounding in both mathematics and psychology, and thus the collaboration of mathematicians and cognitive scientists.

Similarly, Assumptions 5 and 6 serve to frame another large task in that research agenda – that of understanding the world from the student's point of view, and developing means of characterizing the effects of instruction on the ways that students' mathematical world views develop. My choice of words may seem extremely broad, but the choice is deliberate. As Doyle's work (this issue) indicates, students learn various lessons from their instruction—for example, that if the lesson learned is that one does tasks to satisfy the teacher, and that the "design specs" of the task (how to get it done) are what count, then the student may not learn much about the subject matter. All of the examples in this article are similar in kind. What the students in the target class learned about geometry extended far beyond their mastery of proof and construction procedures. They developed perspectives on the role of each, which in turn determined which knowledge they used—or failed to use. Similarly, their views about mathematical form, "problems," and their role as passive consumers of others' mathematics, all shaped their mathematical behavior. The other literature examples indicate that the same holds true in all mathematics instruction.

In sum, research on the psychology of teaching and learning needs to be expanded both in scope and in breadth. "Learning outcomes" must be broadly defined if we are to provide adequate characterizations of behaviors such as those described in the previous paragraph. But explorations of learning also need to become more focused and detailed as we begin to elaborate on what it means to think mathematically. It is also essential—both for research purposes and because measurement is the "bottom line" for much real-world instruction—for our efforts to include the development of measures that will adequately characterize this expanded notion of mathematics learning. And if we really intend to have an impact on practice, we will need to become deeply involved in the development and testing of instructional materials.

This list of tasks may seem daunting, but it is not beyond our reach. There is, as noted earlier, an increasing rapprochement between researchers on teaching and cognitive scientists. Similarly, there are closer ties between psychologists of learning and subject-matter experts as the result of perceived need for collaborative efforts. As our sense of the task grows, so does our capacity to deal with it.

REFERENCES

Anderson, J., Boyle, C., Franklin, F., & Reiser, B. (1985). Intelligent tutoring systems. *Science, 228,* 456–462.

Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science, 4,* 379–426.

Carpenter, T. P., Lindquist, M. M., Matthews, W., & Silver, E. A. (1983). Results of the third NAEP mathematics assessment: Secondary school. *Mathematics Teacher, 76*(9), 652–659.

Corno, L. (1988). The study of teaching for mathematics learning: Views through two lenses.

Educational Psychologist, 23, 181–202.

Doyle, W. (1978). Paradigms for research on teacher effectiveness. In L. S. Shulman (Ed.), *Review of research in education* (Vol. 5, pp. 163–198). Itasca, IL: Peacock.

Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist, 23,* 167–180.

Dunkin, M. J., & Biddle, B. J. (1974). *The study of teaching.* New York: Holt, Reinhart & Winston.

Helms, H., & Novak, J. (Eds.). (1984). *Proceedings of the international seminar on misconceptions in science and mathematics.* Ithaca, NY: Cornell University.

Leinhardt, G., & Smith, D. (1984). *Expertise in mathematics instruction: Subject matter knowledge.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Resnick, L. B. (1983). Toward a cognitive theory of instruction. In S. Paris, G. M. Olson, & H. W. Stevenson (Eds.), *Learning and motivation in the classroom* (pp. 5–38). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Romberg, T. A., & Carpenter, T. P. (1986). Research on teaching and learning mathematics: Two disciplines of scientific inquiry. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 850–873). New York: Macmillan.

Schoenfeld, A. H. (1985). *Mathmatical problem solving.* New York: Academic.

Schoenfeld, A. H. (1986). On having and using geometric knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 225–264). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Shulman, L. S. (1986). Paradigms and research programs in the study of teaching: A contemporary perspective. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 3–36). New York: Macmillan.

Silver, E. A. (Ed.). (1985). *Teaching and learning mathematical problem solving: Multiple research perspectives.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Silver, E. A. (1986). Using conceptual and procedural knowledge: A focus on relationships. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 181–198). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Wertheimer, M. (1959). *Productive thinking.* New York: Harper & Row.