MATHEMATICS TEACHING AND LEARNING

Alan H. Schoenfeld

Elizabeth and Edward Conner Professor of Education

Graduate School of Education

University of California

Berkeley, CA 94720-1670, USA

Email: alans@socrates.berkeley.edu

Draft T: April 27, 2005

A draft for the *Handbook of Educational Psychology, Second Edition*

It was the best of times, it was the worst of times,

it was the age of wisdom, it was the age of foolishness,

it was the epoch of belief, it was the epoch of doubt,

it was the season of Light, it was the season of Darkness,

it was the spring of hope, it was the winter of despair,

we had everything before us, we had nothing before us.

Charles Dickens, *A Tale of Two Cities*

Overview

This chapter focuses on advances in the study of mathematics teaching and learning since the publication of the first edition of the *Handbook of Educational Psychology* (Berliner & Calfee, editors) in 1996. Because of the scope of the review, comprehensive coverage is not possible. In what follows I have chosen to focus thematically on major areas in which progress has been made or where issues at the boundaries of theory and practice are controversial.[1] These areas include: research focusing on issues of teacher knowledge and aspects of professional development; issues of curriculum development, implementation, and assessment; issues of equity and diversity; and issues of learning in context(s). The chapter concludes with a discussion of the state of the field and its contextual surround.

Teacher Knowledge

Significant progress has been made over the past decade in understanding mathematics teachers' knowledge, how it plays out in practice, and how it can be developed. The field can

[1] This approach, like any approach to mapping out a huge territory, results in some unfortunate omissions. Many fine pieces of work, specifically, many studies that focus on learning and conceptual growth in particular mathematical topic areas, are not discussed here. Nor is the role of technology in mathematics learning. Readers with specific interests in these topics will want to consult the forthcoming *Second Handbook of Research on Mathematics Teaching and Learning* (Lester, in preparation). The previous editions of that *Handbook* (Grouws, 1992) and this one (Berliner & Calfee, 1996) provide relevant background and context for this article.

boast of two major books and two additional programmatic bodies of work, all of which add significantly to our understanding. Over the past decade, two major works have emerged that expand the field's conception of the nature and complexity of the knowledge that teachers bring to the classroom. Liping Ma's 1999 book *Knowing and teaching elementary mathematics* demonstrated the unique character of highly accomplished mathematics teachers' knowledge – a knowledge clearly different from knowledge of the mathematics alone. Magdalene Lampert's 2001 book *Teaching Problems and the Problem of Teaching* offers a remarkably detailed empirical and theoretical examination of the multiple levels of knowledge, planning, and decision-making entailed in a year's teaching. Next, I briefly describe Deborah Ball, Hyman Bass, and their colleagues' studies of the mathematical knowledge that supports effective teaching, and the work of Miriam Sherin in describing teachers' *professional vision.* Like the work described before it, this work sheds light on the character of knowledge that enables teachers to interact effectively with students over substantial mathematics.

This work is followed by a description of the work by the Teacher Model Group at Berkeley, which has worked to characterize both the nature of teacher knowledge and the ways that it works in practice. Like the work of Ball, Bass, and colleagues, this work characterizes teaching as problem solving. It contributes to the problem solving and teaching literatures by describing, at a theoretical level of mechanism, the kinds of decision-making in which teachers engage as they work to solve the problems of teaching.

Pedagogical content knowledge.

The study of teacher knowledge was revitalized in the mid-1980s when Lee Shulman (1986, 1987) introduced the notion of *pedagogical content knowledge.* Although the term was not clearly defined at the beginning, the very notion of *specialized content-related knowledge for* *teaching* caught the field's imagination and opened up significant new arenas for both research and practice. We shall begin by exemplifying the concept and indicating its practical implications, after which we turn to contemporary research.

Here is an example familiar to any algebra teacher. Relatively early in the course, one may use the distributive property to show that $(a + b)^2 = a^2 + 2ab + b^2$. One may also suggest the truth of the formula with an "area model." But, one also knows (after having taught the course once) that, later in the course, when students do their homework or one writes the expression $(x + y)^2$ on the blackboard, a significant proportion of the students will complete the expression by writing, incorrectly, $(x + y)^2 = x^2 + y^2$. The first time this happens, a beginning teacher may be taken aback. But with a little experience, the teacher knows to anticipate this, and to be ready with either examples or explanations. For example, the question "Why don't you try your formula with x = 3 and y = 4?" can lead the student to see the mistake. It exemplifies yet another valuable strategy (testing formulas with examples if one is unsure), and can set the stage for a more meaningful reprise of the reasons that the formula works the way it does.

This kind of knowledge – knowing to anticipate specific student understandings and mis-understandings in specific instructional contexts, and having strategies ready to employ when students demonstrate those (mis)understandings, is an example of pedagogical content knowledge (PCK). PCK differs from general pedagogical knowledge, in that it is tied to content. A general suggestion such as "generate examples and non-examples of important concepts" may seem close, but it hardly arms the teacher with the knowledge for this particular situation (or thousands of others like it). There is a critical aspect of fine-grained domain specificity here: in *this* situation, this kind of example is likely to prove necessary and useful. PCK also differs from "straight" content knowledge: one can understand the correct ways to derive the algebraic relationship under discussion without knowing to anticipate student errors.

The concept is critically important because it points to a form of knowledge that is now understood to be a central aspect of competent teaching – and, one that is at variance with simple notions of teacher "training." Some policy-makers and others have a strongly held belief that what is needed for competent teaching in any domain is a combination of subject matter knowledge and either "common sense" or general pedagogical training. This belief is part of the support structure for a wide range of programs aimed at taking professionals in various mathematical and scientific fields and getting them into the classroom rapidly – the expectation being that a bit of pedagogical training and/or common sense will suffice to prepare those who have solid subject matter backgrounds for the classroom. An understanding of the true bases of pedagogical competency is essential as an antidote to such "quick fixes," and as a precondition for bolstering teacher preparation programs in ways that allow them to prepare prospective teachers more adequately. (For more extended discussions of this issue, see National Academy of Education, 2005.)

Liping Ma's discussion of "Profound Understanding of Fundamental Mathematics"

In simplest terms, Liping Ma's 1999 book *Knowing and teaching elementary mathematics* is a comparison of the knowledge possessed by a relatively small sample of elementary school mathematics teachers in the U.S. and Mainland China. Ma studied 23 "above average" teachers in the U.S. and 72 teachers from a range of schools in China. Her finding was that the most accomplished teachers in China (approximately 10% of those interviewed) had a form of pedagogical content knowledge she calls "profound understanding of fundamental mathematics" or PUFM – a richly connected web of understandings that gave them a deep understanding of the domain and of ways to help students learn it. Broadly speaking, such knowledge was not present in the U.S. teachers Ma interviewed.

In four substantive chapters, Ma studies teachers' understandings of: approaches to teaching subtraction with regrouping, student mistakes in multi-digit multiplication, the generation of meaningful contexts and representations to help students understand division by fractions; and explorations of the relationships between perimeters and areas of rectangular figures. Here I shall describe the third of these, division by fractions, and use it as a vehicle for discussing PUFM in general.

Ma offered her interviewees the following scenario:

> People seem to have different approaches to solving problems involving division with fractions. How do you solve a problem like this one?
>
> $$1\frac{3}{4} \div \frac{1}{2} =$$
>
> Imagine that you are teaching division with fractions. To make this meaningful for kids, something that many teachers try to do is relate mathematics to other things. Sometimes they try to come up with real-world situations or story-problems to show the application of some particular piece of content. What would you say would be a good story or model for $1\frac{3}{4} \div \frac{1}{2}$?
>
> (Ma, 1999, p. 55)

Only 9 of the 21 U.S. teachers who worked the problem produced the correct numerical answer to the division problem. This clearly points to a problem with the teachers' algorithmic competency. In contrast, all 72 Chinese teachers performed the computation correctly. Some teachers made formal arguments, in effect justifying the algorithm by either explaining why one inverts $\frac{1}{2}$ and multiplies by its reciprocal, 2. Some converted to decimals. Of much greater interest, however, are the ways in which teachers offered different story representations for the divisions. In trying to produce stories that could motivate and represent the division, 10 of the U.S. teachers confounded division by $\frac{1}{2}$ with division by 2 – they discussed two people sharing $1\frac{3}{4}$ pies, or other objects, equally between them. Six more confounded division by $\frac{1}{2}$ with multiplication by $\frac{1}{2}$. Only one of the U.S. teachers generated a story that corresponded correctly to the given division. In contrast, 90% of the Chinese teachers generated appropriate stories for the division.

Many of the teachers worked through their stories, giving meaning to the mathematical processes thereby. They demonstrated a wide range of ways to think through, and give meaning to, what it means to divide by one half.

Ma proceeds from her empirical description of the Chinese teachers' knowledge to a theoretical description. She characterizes their understandings of various mathematical topics (for teaching) as "knowledge packages" – tightly bound collections of information that include the meaning of a given concept and related mathematical concepts, representations of the concept and related mathematical concepts, skills (algorithms) and their conceptual underpinnings; and relationships between all of the above. She argues that a *Profound Understanding of Fundamental Mathematics* (PUFM) is built up of a well organized collection of such knowledge packages, and she goes on to suggest ways in which teachers develop such understandings.

PUFM is fundamentally mathematical – the core ideas are about mathematical structure. But it is also fundamentally pedagogical, with an organization aimed at meaning-making and deep understanding. In this sense, the knowledge possessed by an accomplished teacher overlaps with, but is different from, that of an accomplished mathematician. There are likely to be aspects of elementary mathematics such as rational number (fractions) that any mathematician knows, and that a highly accomplished teacher does not know – for example, the formal definition of the rational numbers as equivalence classes of ordered pairs of integers. But, there are also aspects of elementary mathematics that teachers with PUFM possess, and professional mathematicians do not. These include having a substantial number of ways of giving meaning to mathematical operations and concepts, and seeing and fostering connections among them. PUFM represents a deeper, more connected understanding of elementary mathematical sense-making than mathematicians are likely to know. It is a different (though related) form of knowledge.

Magdalene Lampert's "Teaching Problems and the Problem of Teaching"

In *Teaching Problems and the Problem of Teaching,* Lampert (2001) takes on the extraordinarily difficult challenge of unraveling the complexities of teaching – of portraying the complex knowledge, planning, and decision-making in which she engaged, over the course of a year, as she taught a class in fifth grade mathematics. This book is an eloquent and elegant antidote to simplistic views of the teaching process. Lampert writes:

> One reason teaching is a complex practice is that many of the problems a teacher must address to get students to learn occur simultaneously, not one after another. Because of this simultaneity, several different problems must be addressed by a single action. And a teacher's actions are not taken independently; there are inter-actions with students, individually and as a group.…
>
> When I am teaching fifth-grade mathematics, for example, I teach a mathematical idea or procedure to a student while also teaching that student to be civil to classmates and to me, to

> complete the tasks assigned, and to think of herself or himself and everyone else in the class as capable of learning, no matter what their gender, race, or parents' income. As I work to get students to learn something like "improper fractions," I know I will also need to be teaching them the meaning of division, how division relates to other operations, and the nature of our number system. (Lampert, 2001, p. 2).

Lampert views and portrays her teaching through multiple lenses. She begins close up, with a view of a specific lesson. Lampert describes individual students in the class, and how they began to work on a problem she assigned. She zooms in on one particular interaction, which occurred when a student wrote something on the board that she did not understand. She asked the class if others could explain where that answer might have come from. The student she called on asked instead if she could explain her own solution. This raised a series of dilemmas for Lampert. Which train of reasoning should be followed? Whom does she run the risk of enfranchising or disenfranchising with her choice, and what implications will this have for the power relationships developing in the classroom? Which aspects of the mathematics will be publicly aired, helping other students to connect not only to the "correct" answer but to think through the various ways of understanding the problem? As she wrestles with these issues, the first student asks to change what he has written. He does, and the number he places on the board is close to the right answer. Now Lampert faces yet another choice. How can she "unpack" this student's thinking, so the class can see how and why he arrived at it, and orchestrate a classroom conversation that will result in the student and the class figuring out the right answer? How can she do so in a way that teaches meta-lessons about reviewing and verifying one's work, that connects to as many of the students' understandings as possible, and that reinforces the classroom's norms of respectful and substantive mathematical interactions?

All this and more happens in one segment of one lesson. And, a lesson is a very small part of a year (which, it should be noted, is 10% of a fifth-grader's life-to-date, so personal as well as intellectual development is a very big issue!). The art of Lampert's book is that she presents the incidents in enough detail to allow one to experience them, at least vicariously; then she steps back, providing an analytic commentary on what took place. Over the course of the book, Lampert displays and reflects upon multiple aspects of her teaching, at various levels of grain size. In an early chapter, she presents her reflections and notes on how to get the year started. She identifies her major goals. She compiles a list of productive activities. She views the year through a content lens – students will need to learn the concept of fraction, long division and multiplication, and more. She considers issues related to "learning the practice of mathematics, things like: revision; hypothesizing; giving evidence, explanation; representation." There are issues of physical environment. These are planned in some detail, and then revised in response to ongoing reality – who the students are, and how things seem to be working. Here too, Lampert presents a substantial amount of detail. If, for example, you want students to learn how to make conjectures public, and then to work through those conjectures respectfully (including challenging others' ideas and/or retracting one's own when it turns out not to be right), one must pick problems that will support rich interactions, and work on establishing the right classroom norms.

As noted above, classroom considerations for a fifth grade teacher go far beyond issues of content. A chapter of Lampert's book is devoted to "teaching students to be people who study in school." How does one realize goals such as "teaching intellectual courage, intellectual honesty, and wise restraint" – having students learn to be willing to take considered risks, be ready to change their position with regard to an issue on the basis of new evidence, but weighing evidence carefully before taking or revising a position? How does one define accomplishment, and establish classroom norms consistent with that definition? Here too, Lampert stakes out a particular kind of territory and then explains how she works toward the goals he has defined.

In a final theoretical chapter, Lampert presents an elaborated model of teaching practice. There she reframes the problems of teaching multiple students at the same time, and the social complexities of practice; the problems of teaching over time; the complexities of teaching content with a curriculum that is largely problem-based; and the complexities of teaching in an environment where all the actors – students as well as the teacher – are taken seriously as contributors to a goal-oriented, emergent agenda. This model, and the book, raise far more questions than they resolve. But that is as it should be. Lampert has taken an ill-understood domain and portrayed its complexity. She has done so in a structured and theoretical way, which makes that complexity accessible and identifies key dimensions of teaching performance and goals. Now that the framework exists, further work by others should move toward the elaboration of the model and toward practical research questions of teacher development toward the kinds of competencies described in it.

Ball, Bass, and Colleagues' Study of Mathematical Knowledge for Teaching.

Deborah Ball, Hyman Bass, and colleagues have embarked on a number of projects aimed at understanding the mathematical competencies that underlie teaching. Like the work described above, this growing body of work is predicated on the assumption that mathematics teaching is a deeply mathematical act that is built on a base of mathematical understanding and that also calls for different types of knowledge.

The group's research agenda, writ large, is to understand the mathematical underpinnings for a broad range of pedagogical undertakings, to understand how the teachers' knowledge shapes their classroom practices, and how those practices ultimately affect student learning in mathematics. Papers that describe this agenda and document some progress toward its achievement, include Ball & Bass, 2000, 2003b; Cohen, Raudenbush, & Ball, 2003; Hill & Ball, 2004; Hill, Rowan, & Ball, in press; the RAND mathematics study panel report, 2002; and the Study of Instructional Improvement, 2002.

A central component of this enterprise is the creation of a series of measures that serve to document teacher knowledge and its impact (see http://www.sii.soe.umich.edu/instruments.html, and Study of Instructional Improvement, 2002). For example, one of the project's released assessment items shows three hypothetical students' work on multiplying multi-digit numbers:

Student A	Student B	Student C
35	35	35
x 25	x25	X25
125	175	25
+75	+700	150
875	875	100
		+600
		875

The item asks teachers to identify which of the students might be using a method that could be used in general. Answering the item correctly involves inferring the procedure used by the student in each case, and judging whether it will always produce a correct answer. This involves substantial mathematical problem solving, and extends far beyond knowing and being able to demonstrate the standard procedure. Other items under development examine key aspects of competency in central areas of the elementary mathematics curriculum. Ball and colleagues are beginning to use such measures to document the impact of professional development interventions in mathematics: see Hill & Ball, 2004.

Sherin's studies of teachers' professional vision.

A body of studies that sits squarely at the intersection of teacher knowledge and teacher learning has been conducted for some years by Miriam Sherin (Sherin, 2001, 2002, 2004; Sherin & Han, 2004). A key construct employed by Sherin, reflecting an important part of accomplished teachers' knowledge, is an adaptation of Charles Goodwin's (1994) notion of *professional vision.* Sherin argues that teachers, like other professionals, develop a particular type of perception common to their profession. Archaeologists recognize the remnants of structures where others see piles of rocks. Doctors recognize clusters of symptoms where laypeople may note individual symptoms or none at all. Similarly, mechanics see and hear functions and malfunctions in mechanical devices; architects note structural stability and other characteristics of buildings; and so on. In the case of teaching, Sherin argues that one form of professional vision is a shift from a focus on pedagogy (examining the moves teachers make in particular circumstances) to a perspective that includes a more intense and critical focus on students' thinking. Sherin & Han (2004) document the use of *video clubs* (meetings in which one or more teachers, in collaboration with university researchers, examine videotapes of the participating teachers' classrooms) as a powerful form of professional development, which can serve as a catalyst for this kind of change. In these video clubs, initial conversations about stimulus videotapes typically involved teachers commenting on pedagogy and researchers focusing on student thinking. Over time, the balance changed. Toward the end of the yearlong series of conversations, the bulk of teacher-initiated comments focused on student thinking – and, teachers' comments both explored the meanings of students' statements and synthesized student ideas. This feeds into pedagogy, of course – but into a diagnostic, student-based pedagogy, which is more typical of accomplished teachers.

The Berkeley Teacher Model Group's Modeling of the Teaching Process.

The next research discussed lies at the intersection of research on problem solving and on teaching. Although it overlaps substantially with other work described in this section, it also differs substantially in style.

A major goal of the Teacher Model Group at Berkeley has been to move toward the modeling of increasingly complex behavior – first problem solving in the laboratory (see, e.g., Schoenfeld, 1985), then in tutoring, and finally in teaching. There are at least two dimensions of complexity here. The first is the complexity of the task. Problem solvers in the laboratory had essentially one goal: solve the given problem. Tutors' goals are more complex, as they hope to facilitate learning and must take many other factors related to their students' knowledge into account. And, as Lampert's book makes abundantly clear, teachers are working toward many goals at once: among them having students learn the content under discussion, connect it to other content, learn to become good students, learn to interact productively with other, and develop productively as people. The task of teaching is far more multi-dimensional than the task of solving mathematics problems. Second, mathematics problems (at least in the laboratory) are static. In contrast, the problems one encounters while teaching are highly interactive and contingent: new issues arise constantly and must be dealt with.

A second major goal has been to address the one major theoretical problem remaining in research on problem solving. Research through the 1980s produced a *framework* for the analysis of mathematical problem solving – one that included aspects of the knowledge base (knowledge and strategies), of decision-making (including monitoring and self-regulation), and of beliefs (which shape the problem solver's choice of actions). What was lacking was a *theory* – a specification of how all this fit together, and explained how and why individuals made the problem solving choices they did, on a moment-to-moment basis.

The Teacher Model Group (TMG) has worked to address these issues by building a theory of teaching that produces analytic models of teachers' classroom behavior. These models have the specificity typical of cognitive modeling. They seek to capture how and why, on a moment-by-moment basis, teachers make the decisions they do in the midst of their teaching.

In a series of papers, Schoenfeld and the teacher model group (Arcavi & Schoenfeld, 1992; Schoenfeld, 1998a, 1999, 2000, 2002a, 2005; Schoenfeld, Gamoran, Kessel, Leonard, Orbach, & Arcavi, 1992; Schoenfeld, Minstrell & van Zee, 2000) used a theory-based approach to model an increasingly complex and widely varying set of tutoring and teaching episodes.

The basic idea is that a teacher's decision-making can be represented by a goal-driven architecture, in which ongoing decision-making (problem solving) is a function of that teachers' knowledge, goals, and beliefs. The teacher enters the classroom with a particular set of goals in mind, and some plans for achieving them. At any given time, activated goals may include short-term goals (having students learn the particular content intended for this lesson), medium-term goals (creating and maintaining a supportive climate in which students feel that they can take risks and interact in substantial ways over subject matter), and long-term goals (having students come to see the discipline as a form of sense-making; aiding in their intellectual and personal development). Plans are chosen by the teacher on the basis of his or her beliefs and values. That is, if a teacher believes that skills are crucially important, the plan may include a fair amount of drill. If the teacher wants to foster a certain kind of conceptual understanding, then the activities chosen for the class will reflect that. The teacher then sets things in motion and monitors lesson progress. If there are no untoward or unusual events, various goals are satisfied and other goals and activities take their place as planned. If something unusual does take place, then a decision is called for – the teacher will decide whether to set a new goal on the basis of what he or she believes is important at the moment. If a new high-priority goal is established, the teacher will search through his or her knowledge base for actions to meet that goal (and perhaps other high priority goals as well). This results in a change of direction, with a new top-level goal. When that goal is satisfied, there may be a return to the previously suspended goal, or a re-prioritization.

The analyses of teaching conducted by TMG work at a line-by-line level. Space allows just for one example. In an introductory lesson, Jim Minstrell's class has been discussing ways of computing the "best number" to represent a collection of data. Students have discussed whether outliers should be included in the data set; they have begun to discuss representing the data with a single number – by the average, by the mode. At that point a student raises her hand and says

> This is a little complicated but I mean it might work. If you see that 107 shows up 4 times, you give it a coefficient of 4, and then 107.5 only shows up one time, you give it a coefficient of one, you add all those up and then you divide by the number of coefficients you have.

Note that a teacher in this kind of situation has many options, ranging from sticking to his lesson plan, telling the student they'll discuss the issue after class, to clearing up the issue with a "mini-lecture," to putting his lesson plan on hold to pursue the issue raised by the student. The question: is enough known about Minstrell, in a principled way, to explain what he does (or even to predict what he is likely to do)?

TMG's model of Minstrell includes his goals and beliefs, which include having students experience physics as a sense-making activity; creating a classroom climate in which students feel free to (and rewarded for) pursuing content-related ideas in sensible ways; and having students learn to sort such things out. Minstrell also believes in minimizing teacher "telling," and has developed a technique he calls "reflective tosses" in which he often answers questions with questions, clarifying things but leaving responsibility for generating (at least partial) answers to them with the students. Thus, confronted with the question from the student, the model acts as follows. The question is germane and substantive. It reflects serious engagement on the part of the student, and its clarification will be a clear act of sense-making. Addressing it will provide an opportunity to instantiate the sense-making values Minstrell espouses. So he will address it – now, and in full. How? Via reflective tosses. He will ask the student to clarify what she means, and ask her and the class how her proposed formula relates to the simple version of average ("add up all the numbers and divide by the number of numbers you have") that the class had already discussed.

Evidence from the range of cases that have been modeled suggests that the underlying architecture of TMG's model, and the theory it instantiates – that teachers' decision-making and problem-solving are a function of the teachers' knowledge, goals, and beliefs – are robust. This in turn suggests a series of practical applications. First, the better one understands how something is done, the better one can diagnose it and assist others in their professional growth. Second, it may be possible to delineate typical "developmental trajectories" of teaching skill, aiding in professional development. Third, close analysis has revealed some surprising similarities and common teaching routines in what, on the surface, seem to be the very different classroom action by teachers such as Deborah Ball, Jim Minstrell, and myself. These routines may be things that novice teachers can learn. (See, e.g., Schoenfeld, 2002a).

Teacher Learning – Issues of Professional Development

The content of the preceding section on teacher knowledge leads naturally to the issue of the growth and change of teacher knowledge – and hence to issues of teacher learning and professional development. In this regard, it is important to recall injunction found in *How People Learn* (National Research Council, 2002a, expanded edition) that learning is learning, whether the learner is child or adult (or, specifically, a teacher). That is, the mechanisms by which adults and children learn are the same – as are issues of identity, engagement, conceptual understanding, and the development of productive practices. Some of the most interesting approaches to professional development are those that take the notion of teacher learning seriously.

In various ways, a focus on student thinking is a hallmark of the most noted approaches to professional development. We discuss three such approaches. Discussions of the professional development workshops are largely pragmatic, but the efforts are grounded in research.

Developing Mathematical Ideas (DMI).

The DMI program (Cohen, 2004; Schifter, 1993, 1998; Schifter, Bastable, & Russell, 1999; Schifter & Fosnot, 1993; Schifter, Russell, & Bastable, 1990) is a professional development seminar for elementary school teachers of mathematics. In a section of Schifter (2001) entitled "What mathematical skills do teachers need?" Schifter identifies and exemplifies four critical skills that, she says, are often absent:

Skill 1: Attending to the mathematics in what one's students are saying and doing. This may sound obvious, but, as Sherin's work indicates, focusing on student thinking is actually a learned skill – and not necessarily one that teachers have when they emerge from their teacher preparation programs.

Skill 2: Assessing the mathematical validity of students' ideas. Recall the example from the Study of Instructional Improvement (2002), given above, which showed three different ways that students might find the product (35 × 25) The issue is: even if the work looks non-standard, is the mathematics correct?

Skill 3: Listening for the sense in students' mathematical thinking – even when something is amiss. Once one is alert to the mathematical possibilities in student thinking, one can often find the core of a correct mathematical approach in something that produces an incorrect answer. This gives something to build on.

Skill 4: Identifying the conceptual issues the students are working on. Schifter provides an example of a student responding to a problem with a strange combination of arithmetic operations. Upon closer examination, the student's work is seen to represent an incorrect generalization of a strategy that was useful in a different context. This provided the basis for an interesting mathematical conversation with the student.

The DMI program attempts to provide a series of experiences that help teachers develop these skills. Cohen (2004) provides a detailed description of one of the DMI seminars and its impact on the participants. . There is much to learn from the close and sympathetic examination of adult learners

Cognitively Guided Instruction.

Focusing on student thinking lies at the core of one of the most widely known programs of professional development, Cognitively Guided Instruction, or CGI (Carpenter, Fennema, and Franke, 1996). CGI is based on an extensive body of developmental research on students' understanding of elementary mathematical situations – for example, the mathematically isomorphic *change* situation in "join" form:

> Connie had 5 marbles. Jim gave her 8 more marbles. How many marbles does Connie have altogether?

and in "separate" form:

> Connie had 13 marbles. She gave 5 marbles to Jim. How many marbles does she have left?

Research summarized in Carpenter (1985) documented the kinds of models children construct to represent such situations and a developmental trajectory of the growth in the children's models and their linkages to arithmetic operations. This knowledge base provides a solid grounding with which teachers can interact with students. When a student faces a particular situation, the teacher (guided by a knowledge of developmental trajectories in general, and possessing a repertoire of situation models and ways to formalize them) can determine which understanding the student has, and help the student (a) solve problems based on those understandings, and (b) conceptualize and formalize what he or she knows, thus expanding the student's knowledge base. CGI does not offer a prescriptive pedagogy; rather, it provides the knowledge by which teachers can respond flexibly to and build on their students' current understandings. A large body of research (see, e.g., Carpenter, Fennema, & Franke, 1996; Carpenter, Fennema, Franke, Empson, & Levi, 1999; Carpenter, Franke, Jacobs, & Fennema, 1998; Carpenter & Lehrer, 1999; Franke, Levi, Carpenter, & Fennema, 2001) indicates that as teachers become more familiar with student understanding, they become more flexible in their teaching – and that the effects of the professional development support "generative growth" in teachers' understanding over time (Franke, Levi, Carpenter, & Fennema, 2001).

A difference between CGI and DMI identified by Carpenter, Fennema and Franke (1996) is that CGI explicitly uses an understanding of student work to help teachers develop deeper understandings of the mathematics itself, while DMI uses the study of mathematics to sensitize teachers to a wide range of students' mathematical thought processes. It would seem that these two emphases are part of a productive dialectic in teachers' professional growth. Awareness of student cognition provides an opportunity to think more deeply about mathematics and student conceptions of mathematics. These, in turn, can shape instructional practices, and reflection on those can provide deeper awareness of student cognition.

Lesson study.

Lesson study, an aspect of teacher professionalism in Japan, has the potential to either become the next large-scale educational fad in the U.S. or a powerful form of professional development. The practice has received widespread attention in the West largely as a result of a book entitled *The teaching gap* (Stigler & Hiebert, 1999). Stigler & Hiebert explored possible explanations of mathematics performance data revealed by the Third International Mathematics and Science Study, or TIMSS. A wide range of performance reports (see, e.g., Beaton, Mullis, Martin, Gonzalez, Kelly, & Smith, 1997, Kelley, Mullis, & Martin, 2000; Mullis, Martin, Beaton, Gonzalez, Kelly, and Smith, 1998; Mullis, Martin, Gonzalez, Gregory, Garden, O'Connor, Chrostowski, & Smith, 2000) indicated that the mathematics performance of U.S. students was roughly at the median internationally (and toward the bottom of scores for highly industrialized nations), while Singapore, Korea, and Japan scored consistently at the top. TIMSS video studies (see, e.g., Hiebert, Gallimore, Garnier, Givvin, Hollingsworth, Jacobs, Chui, Wearne, Smith, Kersting, Manaster, Tseng, Etterbeek, Manaster, Gonzales, & Stigler, 2003; Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999) provided compelling evidence that Japanese mathematics lessons were far more coherent than parallel mathematics lessons in the Unites States.

Stigler and Hiebert (1999) argue that teaching is a *cultural activity* – that there is relatively little within-country variation in teaching practices compared to between-country variation. That is, the teaching styles in the U.S. and in Japan are relatively consistent, and different. So are forms of professional development. A key component of instructional improvement is lesson study – the design, implementation, testing, and improvement, *by teachers,* of "research lessons."

The most detailed examination to date of the principles and practices of lesson study can be found in Fernandez & Yoshida (2004). Fernandez and Yoshida provide extensive detail regarding the ways in which a collective of Japanese schoolteachers select a topic of curricular importance, identify the focus of the intended research lesson or lessons, and begin to design the lesson(s). The level of specificity in these deliberations is extraordinary. A typical lesson plan in the U.S. consists of the description of a sequence of intended teacher actions. In contrast, lesson plans developed as part of lesson study include descriptions of: the sequence of intended learning activities; the ways in which students are expected to react to each of these activities; a planned teacher response to each of the likely student reactions; and intellectual foci for the evaluation of the progress of the lesson. Extensively detailed lesson plans are developed and refined by a teacher planning group over a sequence of meetings, as a truly collaborative effort. A group member then volunteers to teach the research lesson, for the collective. The lesson is not seen as the individual teacher's "property." Instead, the teacher is seen as the means of implementation of the group's design. The other teachers observe the trial lesson closely, looking to see what seems to be effective and what is not. The collective then "debriefs" in an elaborate process that leads to the refinement and re-teaching of the lesson by someone else. Ultimately, the tangible product is a shared lesson that is extremely well designed and documented, so that all the participants (and others) can use it. A somewhat intangible but equally important product is the professional growth of those who contribute to the design process. The sequence of lesson design activities consists of thinking hard about the desired content and learning outcomes; about activities intended to promote those outcomes; about student thinking (including anticipating student reactions to activities, and what those reactions mean in terms of student understanding); and about principled revisions to the materials on the basis of careful assessments of student learning. Beginning teachers undertake these activities in the company of more experienced and accomplished colleagues, so there is a natural apprenticeship into the community of skilled practitioners.

There is significant potential for appropriate adaptations of aspects of lesson study to powerful mechanisms for professional development in school systems outside of Japan. At the same time, there is significant potential for the practices of lesson study to be trivialized in ways that render the process superficial and of little or no value. As summarized briefly above, a major component of lesson study involves focusing on student understanding, and then evaluating the pedagogy on the basis of the impact of the designed activities. The lessons design practices described by Yoshida and Fernandez (2004) take place at a very fine level of detail, much more fine-grained than those typically conceptualized by American teachers. As Sherin (2001, 2002) and others have shown, focusing on student work does not necessarily come naturally; one has to learn how to do it. In this author's experience, and that of others who have observed attempts to implement lesson study in the U.S., teachers need to learn how to judge lesson effectiveness. Teachers' first judgments about lesson effectiveness are often global and not grounded in data. One hears statements like "the timing felt pretty good" or " they were engaged most of the time" much more frequently than one hears commentary on the actual content of what students said and did. The question, then, is whether teachers will be provided the support structures (including time, and perhaps external resources until teachers have developed the relevant skills and understandings) that will enable them to bootstrap the skills needed to implement lesson study effectively. Absent such support, the prognosis is not good.

Curriculum development, implementation, controversy, and assessment

The 1990s were (in the U.S. and in places around the globe that are influenced by U.S. curricula) the time of the greatest curricular change and controversy since the "new math" of the 1960s. Research over the latter part of the 20th century produced new understandings about the nature of mathematical thinking and learning – ideas that would result in the reconsideration of the foci and contents of mathematics curricula. New goals for curricula were codified in the National Council of Teachers of Mathematics' 1989 *Curriculum and evaluation standards for school mathematics.* Concurrently, perceptions of national economic crises provided an impetus for the revision of mathematics curricula. With a significant infusion of funding from the National Science Foundation, a number of "standards-based" curricula were developed. Many of these curricula, produced in the 1990s, differed substantially in look, feel, and classroom implementation from "traditional" curricula. Many turned out to be controversial – so much so that the term "math wars" was coined to describe the public controversies that followed. Resolving such controversies depends, or course, on the evidence available. This raises issues of assessment: what should be assessed regarding mathematical thinking and learning, and how does one assess such things? What should be examined when one examines the impact of various curricula? How does one examine curricular impact in rigorous and informative ways? Those are the issues explored in this section.

Context

By the mid-1980s, scholars in mathematics education had reconceived the epistemological foundations of mathematics learning. Broadly speaking, the view of mathematics learning as the "acquisition of knowledge" had been superseded by the perspective that being competent at mathematics meant understanding and being able to use mathematical concepts and procedures – and that in addition, strategic competence, metacognitive ability (including monitoring and self-regulation), and productive beliefs and affect (or disposition) were important aspects of mathematical competence.

In the 1980s, the U.S. was also feeling threatened economically. *A Nation at Risk* (1983) issued a clarion call for the reform of U.S. mathematics and science education. In 1989 The National Council of Teachers of Mathematics (NCTM), a professional organization of mathematics teachers, produced the *Curriculum and Evaluation Standards for School Mathematics* (subsequently known as the *Standards*). The *Standards* were a philosophical as well as a curricular document. The goal of the writers was to "create a coherent vision of what it means to be mathematically literate" in a rapidly changing world, and to "create a set of standards to guide the revision of the school mathematics curriculum." (p. 1). The authors defined "standard" as follows:

> ***Standard.*** A standard is a statement that can be used to judge the quality of a mathematics curriculum or methods of evaluation. Thus, standards are statements about what is valued. (p. 2)

That is, the *Standards* were intended to be a statement of "what matters" in mathematics instruction or testing. They were not intended to be a blueprint for curriculum development. Philosophically, they focused on new goals for students and society: "New social goals for education include (1) mathematically literate workers, (2) lifelong learning, (3) opportunity for all, and (4) an informed electorate." (p. 3) The *Standards* were oriented toward "five general goals for all students: (1) that they learn to value mathematics, (2) that they become confident in their ability to do mathematics, (3) that they become mathematical problem solvers, (4) that they learn to communicate mathematically, and (5) that they learn to reason mathematically." (p. 5)

Here I shall focus on the curricular descriptions found in the *Standards* – the first three sections, which discuss curricular desiderata in kindergarten through grade 4, grades 5-8, and grades 9-12 respectively. Each of the three grade band sections contained 13 or 14 Standards. For the first time in curricular history, a major curriculum document gave as much attention to the *process* aspects of mathematical performance as it did to the mathematical content to be covered in the curriculum. At each of the grade bands, the first four standards were the same: Mathematics as Problem Solving, Mathematics as Communication, Mathematics as Reasoning, and (making) Mathematical Connections. The remaining standards described content – but again, in a very broad way, covering four or five grades at a time. As a result, there was no single model curriculum that "fit" or "met" the *Standards;* one could imagine a wide range of very different curricula that had the same content and process emphases, and that achieved the same broad goals.

In 1989 no commercial publisher would undertake the creation and production of "standards-based" mathematics curricula, because doing so was too risky in financial terms. Mathematics curricula were typically sold in K-8 series, which were produced by large writing teams. This was done in something like production-line fashion, to meet the textbook adoption deadlines of major states such as California, Texas, and New York. In order not to lose huge chunks of the market, publishers made sure their books met the adoption criteria of those three states. Publishers claimed the cost of developing and marketing a K-8 series was on the order of $25 million – far too much to risk on an untried product.

In sum, the national context in 1989 looked like this. There was a perceived need for an upgrading of mathematics curricula. The NCTM *Standards* offered a set of criteria by which a new curriculum could be judged – but commercial publishers would not make the investment. The U.S. National Science Foundation (NSF) addressed this dilemma by serving as a catalyst for curriculum development. NSF began issuing requests for proposals (RFPs) for Standards-based curricula in the early 1990s[2]. A list of NSF-supported curriculum projects that can be found at <http://forum.swarthmore.edu/mathed/nsf.curric.html>. The NSF provided support for the development and implementation of standards-based curricula in a variety of ways.

For example, NSF sponsored a series of annual "Gateway" conferences from 1992 through 1998 at which curriculum developers came together to discuss issues of common concern. Later NSF established four national centers devoted to the support of standards-based curricula: the K-12 mathematics curriculum center, whose web site is <http://www.edc.org/mcc>, an elementary grades curriculum center at <http://www.arccenter.comap.com>, a middle grades center at <http://showmecenter.missouri.edu>, and a high school center at <http://www.ithaca.edu/compass>. In parallel, NCTM worked steadily to maintain support for standards-based instruction. NCTM's annual meetings focused on *Standards*-related activities, and NCTM produced a number of publications aimed at helping teachers to implement *Standards*-based instruction in their classrooms.

[2] It should be noted that a small number of "reform" efforts had begun (also with grant funding, independent of the major publishers) prior to the NSF call for proposals. For example, the University of Chicago School Mathematics Project (UCSMP) and the Interactive Mathematics Project (IMP) pre-dated the NSF curriculum RFPs.

When the previous version of this *Handbook* was published in 1996, there was scant evidence – either positive or negative – regarding the effectiveness of the new curricula. This stands to reason. The NSF curriculum RFPs were first issued in 1991, so in 1996 the various curriculum projects were just completing their first (alpha) round of development. Indeed, it was not until the turn of the 21st century that cohorts of students had worked through the full beta versions of many of these curricula. The current situation with regard to data evaluating curricular effectiveness is not much better. The current state will be discussed below. First, however, it is important to discuss the largest social confrontation over mathematics curricula since the controversies over the "new math."

Math Wars

While the "math wars" in the U.S. (and now in parts of the world as far distant as Israel) are in a sense outside the realm of educational psychology (though not social psychology!) and research in mathematics education, they are a critically important phenomenon that needs to be discussed. Researchers need to understand the contexts within which their work is done.

For a detailed history of the math wars in California (where they began), see Rosen (2000); see also Jackson (1997a, b) and Schoenfeld (2004). The 1985 California Mathematics *Framework* was considered a mathematically solid and progressive document. State Superintendent of Education Bill Honig supported educational reform, and the California Mathematics Council (the state affiliate of NCTM) actively supported *Standards*-based practices after the *Standards* were published. The 1992 California Mathematics *Framework* represented a next step in the change agenda. Publishers created texts in line with their view of *Standards*- and *Frameworks*-based mathematics. In 1994 the California State Board of Education approved instructional materials consistent with the Mathematics *Framework*.

The *Framework,* like the *Standards,* was a vision statement regarding the desired substance and character of instruction rather than a blueprint for them. Such documents invite curriculum designers to create innovative materials. But, there is a downside to opening the door to such creativity. Rosen (2000, p. 61) notes that "the new textbooks were radically different from the traditional texts' orderly, sequential presentation of formulas and pages of practice problems familiar to parents. New texts featured colorful illustrations, assignments with lively, fun names and sidebars discussing topics from the environment to Yoruba mathematics (prompting critics to dub new programs with names such as 'Rainforest Algebra' and 'MTV Math')." Sometimes frenetic in appearance, sometimes different in content, many of the new texts could be easily caricatured. Once the rhetorical battles heated up, they were.

In addition, many "reformers" and reform curricula called for new teaching practices, urging less dependence on teacher exposition and whole-class recitations, and increased dependence on small group work. Maintaining a focus on substantial mathematics while also fostering communication and collaboration in group work is quite difficult. Teachers who had themselves been taught in traditional ways were now being asked to teach in new ways. Many were not up to the task (see, e.g., Ferrini-Mundy & Schram, 1997.)

These new materials and practices raised concerns among some parents, some of whom viewed them as a repetition of the mistakes of the "new math." Parent groups organized, established websites hostile to "reform," and created a very effective anti-reform movement. Local oppositional movements soon coalesced into a state-wide (and then nation-wide) movement, supported by prominent conservatives such as California Governor Pete Wilson. The state legislature held highly contentious public hearings on the *Frameworks* in 1995 and 1996. Conservatives prevailed, and a new mathematics *Frameworks* writing team was convened ahead of schedule. The State legislature enacted AB 170, which "requires the State Board of Education to ensure that the basic instructional materials it adopts for reading and mathematics in grades 1 to 8, inclusive, are based on the fundamental skills required by these subjects, including… basic computational skills." (See http://www.cde.ca.gov/board/readingfirst/exhibit-i.pdf.)

The next major skirmish took place over the California Mathematics Standards. In line with traditional California practice, a draft had been developed by a committee and submitted to the State Board of Education. The orientation of the draft, which had been put together over a year and a half and had undergone a substantial amount of public review, was generally consistent with that of the NCTM *Standards*. The State Board summarily rejected the draft. Over a period of just a few weeks, the Board rewrote much of the elementary grades section itself. It commissioned a small number of mathematics faculty (who had negligible experience with K-12 classrooms or curricula) to rewrite the standards for the secondary grades. These acts elicited protests from highly visible scholars such as Hyman Bass, research mathematician and Director of the National Research Council's Mathematical Sciences Education Board, and William Schmidt, who had conducted curriculum content analyses for the Third International Mathematics and Science Study. The Board ignored the protests.

How one views these events depends on one's perspective. Here is how anti-reform activist David Klein described them:

> "Question: What would happen if California adopted the best, grade-by-grade mathematics achievement standards in the nation for its public schools?
>
> Answer: The education establishment would do everything in its power to make them disappear.

> In December 1997, the State Board of Education surprised the world by not accepting extremely bad, “fuzzy” math standards written by one of its advisory committees, the Academic Standards Commission. Instead, in a few short weeks and with the help of four Stanford University math professors, the state board developed and adopted a set of world-class mathematics standards of unprecedented quality for California's public schools.”

Klein’s rhetoric suggests the level of vitriol spewed in the math wars – for example, Maureen DiMarco, California State Secretary of Child Development and Education, referred to the new curricula as “fuzzy crap.” Acrimony reached the point where U. S. Secretary of Education Richard Riley felt compelled in January 1998 to address the annual Joint Mathematics Meetings[3], urging civility and respectful exchange in battles over mathematics curricula. His words went unheeded, and Riley soon found himself immersed in the math wars: anti-reform forces orchestrated the signature-gathering for an open letter to Riley, published in major newspapers nationwide, protesting the U. S. Department of Education’s listing of “exemplary” and “promising” instructional programs in mathematics education. In California, those who had power exercised it without restraint. For example, members of the State’s Curriculum Framework and Criteria Committee were barred from introducing research into the record or into the group’s deliberations. Interested readers should see Becker & Jacob (2000), Jacob (1999, 2001), and Jacob & Akers, (2003) for details.

The point here is that when educational/psychological issues enter the political arena, scholarly discourse and well-grounded research findings are often marginalized. The research community needs to think about how to deal responsibly with such issues. Research does little good if it can be ignored for purposes of political expediency.

[3] The meetings are sponsored by the American Mathematical Society, Mathematical Association of America, Association for Symbolic Logic, Association for Women in Mathematics, National Association of Mathematicians, and Society for Industrial and Applied Mathematics.

We now return to research issues and the question of evidence. Just what evidence was there, at the time of the math wars, of the efficacy of traditional or reform curricula? What evidence is there now?

Evidence, then and now: Issues of assessment

Simply put, the math wars were fought in an informational vacuum. As noted a few standards-based instructional programs were under development when NSF issued its curricular RFPs in the 1991. Preliminary development of most of the NSF-supported curricula took four or five years, and refinements (to the "beta level" took another few years. It was not until the late 1990s that full cohorts of students had worked their way through any of the new curricula. Perhaps surprisingly, there are even fewer detailed evaluations of "traditional" curricula than of the more recent standards-based curricula[4]. The traditional curriculum has existed for many years in various forms – over the years, mainstream textbook series came to resemble each other closely in content coverage. However, until recently – specifically, until the passage of the No Child Left Behind legislation (see http://www.ed.gov/offices/OESE/esea/) – textbook publishers had little or no incentive to gather data regarding student performance. Textbook marketing depends on focus groups (what do the consumers want?) and testimonial in advertising rather than on data. This stands in sharp contrast to the marketing of consumer items such a cellular phones, washing machines and cars, where marketing depends of necessity on focus groups, testimonials, *and* data. If one wants to buy a new car or washing machine, there are specialist magazines (e.g., *Car and Driver* or *Motor Trend*) and general consumer magazines (e.g., *Consumer Reports*) that evaluate the features of various models on their own or in comparison with others. If a product has a specific design flaw, a bad performance record, or some other

[4] There have been, of course, many studies of the impact of traditional curricula. The issue here is whether there have been evaluations along dimensions now considered appropriate for assessing students' mathematical competency.

problem, that problem will be made public and will need to be addressed. In contrast, until recently, there was no mechanism for curricular performance evaluations in education.[5] Since such assessments are costly, there was no reason for publishers to undertake them.

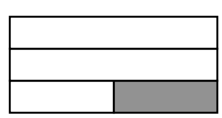
Second and equally important, there is the question of just what a test "covers." The standardized tests available until recently, including NAEP, were designed under the standard psychometric assumptions of trait and/or behaviorist psychology (see Glaser & Linn, 1997; Greeno, Pearson, & Schoenfeld, 1997). They focused largely on content mastery. In contrast, the more fine-grained analyses of proficiency developed over the past decade or so (for example the *New Standards* and *Balanced Assessment* tests) tend to be aligned with the content and process delineations found in NCTM's (2000) *Principles and Standards for School Mathematics.*

Let us consider a topic such as "understanding fractions" to see why test coverage is a critically important issue. A traditional assessment would look for students' ability to perform standard algorithms, for example

Task 1. Find (1/2)(3/5) + (1/2)(1/5).

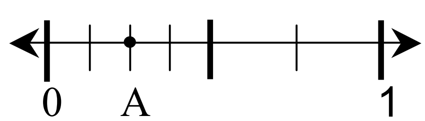
More contemporary assessments might seek to understand students' abilities to work with different representations of fractions, for example:

Task 2. Write a fraction for the shaded part of the region below



or

Task 3. Write a fraction for point A.



[5] That mechanism is the What Works Clearinghouse, which is discussed below.

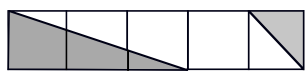
Each of Tasks 2 and 3 calls for understanding a particular representation of fractions. Many students will respond "1/4" to Task 2, neglecting the criterion that the four ostensible "fourths" of the figure must all have equal area. Similarly, many students will respond "2/6" to Task 3. Interestingly, students may answer correctly to one of those two tasks and not the other. Responding correctly depends on some familiarity with the representation.

Different tasks may probe for conceptual understanding:

Task 4. Explain what happens to the value of a fraction when its denominator is doubled.

Or, they may call for a combination of representational use and problem solving. Consider the following problem, for example.

Task 5. Write a fraction for the shaded part of the region below.



There are various ways to solve this problem, but here is one:

> Each of the boxes is 1/5 of the whole region. The diagonal line on the left-hand side of the figure divides the first three boxes in half, so the shaded area on the left is (1/2)(3/5) = (3/10) of the region; the shaded area on the right is (1/2)(1/5) = (1/10) of the region; so the area of the two shaded regions combined is (3/10) + (1/10) = (4/10) = 2/5 of the region.

Note that this solution calls for problem solving in making use of a non-standard representation, "seeing" the rectangle formed by the first three boxes, knowing that the diagonal divides a rectangle in half, obtaining the areas of each of the shaded regions, and performing the numerical computations given above. It stands to reason that many students who could obtain the correct answer to task 1 would fail to obtain the correct answer to task 5 – which includes the computation in Task 1 as a sub-problem.

These few examples are just the tip of the proverbial iceberg. The point is that what one assesses matters. A student who does well on a test comprised of items like task 1 might do poorly on a test that included tasks 2 through 5; hence a test comprised of items like task 1might yield a "false positive" for the student, indicating more competency than was actually there. Or, consider two curricula, with Curriculum A focusing on procedural skills and Curriculum B giving attention to skills, conceptual understanding, and problem solving. If a test comprised of items like tasks 1 through 5 was used to assess learning outcomes, curriculum B would most likely outperform curriculum A. But if a test using only items like Task 1 was used, it is possible that both groups would perform comparably – an assessment "false negative," because the test did not capture a range of understandings possessed by students of curriculum B.

This example is not hypothetical. Ridgway, Crust, Burkhardt, Wilcox, Fisher and Foster (2000) compared students' performance at grades 3, 5, and 7 on a standardized high-stakes, skills-oriented test (the California STAR test, primarily the SAT-9 examination) with their performance on a much broader standards-based test (the Balanced Assessment tests produced by the Mathematics Assessment Resource Service, or MARS.). For purposes of simplicity, scores reported here are collapsed into two simple categories. A student is reported as being either "proficient" or "not proficient" as indicated by his or her scores on each of the examinations. The data are given below.

Grade 3 (N=6136):

		SAT-9	
		Not Proficient	Proficient
MARS	Not Proficient	27%	21%
	Proficient	6%	46%

Grade 5 (N=5247):

		SAT-9 Not Proficient	SAT-9 Proficient
MARS	Not Proficient	28%	18%
	Proficient	5%	49%

Grade 7 (N=5037):

		SAT-9 Not Proficient	SAT-9 Proficient
MARS	Not Proficient	32%	28%
	Proficient	2%	38%

These data suggest that more than 20% of the student population, who passed the SAT-9 but failed the MARS test, were actually "false positives." With this as context, I examine data that provide comparative evaluations of standards-based and more "traditional" curricula. I first describe the general knowledge base, then findings from the What Works Clearinghouse.

As noted above, there exist sparse data regarding the effectiveness of the newer curricula, and even less evidence regarding traditional curricula. This is partly a matter of timing, partly a matter of incentives. Because the first cohorts of students emerged from standards-based curricula just about the turn of the 21st century, there has not been time for extensive studies of those curricula. Most of the newer curricula have been examined, however – as proposed alternatives to the status quo, they had to prove themselves. In contrast, the mainstream texts, which dominated the marketplace, had little incentive to prove themselves until the passage of the No Child Left Behind act, known as NCLB. (See <http://www.ed.gov/nclb/landing.jhtml?src=pb> for the U. S. Department of Education's web site devoted to NCLB.) The accountability procedures specified in NCLB mandate the gathering of testing data, though they do not specify the content of the tests. Hence the issues discussed above remain central.

The most complete record of evaluations of standards-based curricula to date can be found in Senk & Thompson (2003). Putnam (2003) summarized the results of four elementary curriculum evaluations as follows:

> [These four curricula] … all focus in various ways on helping students develop conceptually powerful and useful knowledge of mathematics while avoiding the learning of computational procedures as rote symbolic manipulations… Students in these new curricula generally perform as well as other students on traditional measures of mathematical achievement, including computational skill, and generally do better on formal and informal assessments of conceptual understanding and ability to use mathematics to solve problems. These chapters demonstrate that "reform-based" mathematics curricula can work. (Putnam, 2003, p. 161).

Putnam notex common assumptions underlying the four programs: that there should be a focus on important mathematics, that instruction should build on students' current understanding, that learning should be grounded in settings that are meaningful to students, that curricula should build on students' informal knowledge and help them formalize what they learn; and that the curricula call for teachers' guiding students through learning experiences to achieve curricular goals. He also notes the difficulties of using widely available standardized measures: "A disadvantage in using standardized measures is that they often do not shed much light on the more complex mathematical understanding, reasoning, and problem solving emphasized in the new curricula. This is the "curricular false negative" problem mentioned above. Putnam observes that the authors of the curricula used a range of measures to get at competencies not revealed by standardized measures, including "evaluator-developed measures, with their concomitant strengths and weaknesses." (Putnam, 2003, p. 166).

The story was much the same at the middle school level. Like the elementary curricula, the NSF-supported middle school curricula discussed are grounded in NCTM's 1989 *Standards;* they too focus on meaningful and engaging problem-based mathematics. There are differences in emphases and style, but the similarities in character and outcomes outweigh the differences. Chappell (2003) summarizes the results as follows:

> Collectively, the evaluation results provide converging evidence that *Standards*-based curricula may positively affect middle-school students' mathematical achievement, both in conceptual and procedural understanding…. They reveal that the curricula can indeed push students beyond the 'basics' to more in-depth problem-oriented mathematical thinking without jeopardizing their thinking in either area (Chappell, 2003, pp. 290-291).

One sees similar results at the high school level. On standardized tests such as the CTBS or PSAT, there were no significant differences between student performance on three reform curricula; on two other reform curricula students "performed as well, if not better, than students from a traditional curriculum on standardized or state assessments." (Swafford, 2003, p.459.) When one turns to evaluator-developed measures of conceptual understanding or problem solving, students in all of these curricula outperformed students from traditional curricula. In sum:

> Taken as a group, these studies offer overwhelming evidence that the reform curricula can have a positive impact on high school mathematics achievement. It is not that students in these curricul[a] learn traditional content better but that they develop other skills and understandings while not falling behind on traditional content. (Swafford, 2003, p.468.)

A series of studies in Pittsburgh, PA (See Briars, 2001, Briars & Resnick, 2000, Schoenfeld, 2002b) documents what may be the most positive "local" implementation of a standards-based curriculum. Diane Briars, mathematics specialist for the Pittsburgh schools, had been conducting *Standards*-based professional development activities for some years when the *Everyday Mathematics* curriculum became available. Uneven implementation of the new curriculum in Pittsburgh provided the opportunity for a natural experiment – a comparison of student performance in "strong implementation schools" where teachers implemented the curricula with strong fidelity with student performance in demographically matched schools where the new curricula were essentially ignored, and teachers implemented the prior, "traditional" curriculum. The Pittsburgh data document an across-the-boards improvement in test scores for the new curriculum on tests on sub-scores of skills, conceptual understanding, and problem solving. Indeed, some "racial performance gaps" were overcome with the new curricula.

Because of the context, this is a somewhat idealized case – but it indicates that when:

- a district's mathematics program is grounded in a rich, connected set of standards;
- the mathematics curricula used are consistent with the standards;
- assessments are consistent with the standards;
- teachers' professional development is consistent with the standards; and
- there is enough systemic stability for sustained growth and change,

there is the potential for substantial improvement in student performance, and for the reduction of racial performance gaps.

Riordan & Noyce (2001) report on a series of comparison studies in Massachusetts. These studies, which used the statewide assessment (the MCAS) as the measure of performance, show that fourth and eighth graders using standards-based texts "outperformed matched comparison groups who were using a range of textbooks commonly used in Massachusetts…. These performance gains… remained consistent for different groups of students, across mathematical topics and different types of questions on the state test" (Riordan & Noyce, 2001, pp. 392-393).

Reys, Reys, Lappan, Holliday &Wasman (2003) examined the impact of standards-based curricula on the performance of more than 2000 eighth grade students in three matched pairs of school districts in Missouri. The assessment used for the comparison was the state-mandated mathematics portion of the Missouri Assessment Program (MAP), which is administered annually to all Missouri eighth graders. The MAP assesses mathematical skills, concepts, and problem solving abilities as delineated in the State Mathematics Framework for. Students who had used standards-based materials for at least two years scored significantly higher than students from the districts that used non-NSF curricular materials.

In the largest study conducted to date, the ARC Center (see <http://www.comap.com/elementary/projects/arc/>) an NSF-funded project, examined reform mathematics programs in elementary schools in Massachusetts, Illinois, and Washington. The study included more than 100,000 students. It compared the mathematics performance of students from schools implementing *Standards*-based curricula with a matched sample of comparison schools that used more "traditional curricula. (Criteria for matching included reading level and socioeconomic status.)

> Results show that the average scores of students in the reform schools are significantly higher than the average scores of students in the matched comparison schools. These results hold across all racial and income subgroups. The results also hold across the different state-mandated tests, including the Iowa Test of Basic Skills, and across topics ranging from computation, measurement, and geometry to algebra, problem solving, and making

connections. The study compared the scores on all the topics tested at all the grade levels tested (Grades 3-5) in each of the three states. Of 34 comparisons across five state-grade combinations, 28 favor the reform students, six show no statistically significant difference, and none favor the comparison students. (See <http://www.comap.com/elementary/projects/arc/tri-state%20achievement.htm>.)

There are a few additional studies comparing standards-based and "traditional" curricula, which will be considered in the discussion of the What Works Clearinghouse. But let us take stock at this point. Overall, there are a quite small number of studies that compare the two kinds of curricula. It is the case that the "score sheet" is uniformly in favor of the Standards-based curricula: virtually every study in the literature shows either no significant differences or an advantage to the standards-based curricula on measures of skills, and most show significant advantages to the standards-based curricula on measures of conceptual understanding and problem solving. However, the evidence base is embarrassingly weak. The vast majority of studies, like the majority of those reported in Senk & Thompson (2003), employed evaluator-developed measures of conceptual understanding and problem solving. These could be considered biased toward the curricula they evaluated. Little was said about the comparative conditions of implementation. Because beta versions of the standards-based curricula were generally being tested, one can assume that there was some implementation fidelity in the case of those curricula. In contrast, one knows little about implementation fidelity, or the overall quality of instruction, in the comparison classrooms. Hence the comparative studies can best be considered existence proofs about what the newer curricula can do rather than definitive comparative evidence. This point is made in very clear terms by a recent report of the National Research Council (2005).

The What Works Clearinghouse

Let us now turn to the evaluation effort conducted by the What Works Clearinghouse (WWC, at <http://www.whatworks.ed.gov/>). To quote from the front page of the WWC web site:

> On an ongoing basis, the What Works Clearinghouse (WWC) collects, screens, and identifies studies of the effectiveness of educational interventions (programs, products, practices, and policies). We review the studies that have the strongest design, and report on the strengths and weaknesses of those studies against the WWC Evidence Standards so that you know what the best scientific evidence has to say.

WWC was established to address the kinds of issues discussed above – the fact that there is a paucity of rigorous studies assessing the effectiveness of various kinds of interventions, ranging from mathematics and reading curricula to programs for dropout prevention. WWC's efforts are largely modeled on the Cochrane Collaboration's efforts to develop "evidence-based health care." (See http://www.cochrane.org/index0.htm.) The idea is not for WWC to conduct new studies, but to comb the literature for studies that meet its very stringent methodological criteria, to report on those studies, and ultimately to conduct meta-analyses regarding the effects of educational (and other) interventions.

> The WWC does not endorse any interventions nor does it conduct field studies. The WWC releases study, intervention, and topic reports. A study report rates individual studies and designs to give you a sense of how much you can rely on research findings for that individual study. An intervention report provides all findings that meet WWC Evidence Standards for a particular intervention. Each topic report briefly describes the topic and each intervention that the WWC reviewed. (<http://www.whatworks.ed.gov/>)

WWC considers for review only studies that employ randomized controlled trials, regression discontinuity designs, and quasi-experimental designs with equating. If a study is of one of those types, it is then examined for possible flaws such as lack of implementation fidelity, differential dropout rates between groups, and such. (See <http://www.whatworks.ed.gov/reviewprocess/standards.html> for a list of criteria.) One can, legitimately, complain that the paradigmatic choices made by WWC are far too narrow: there are many ways to conduct informative studies of mathematics curricula, and the three kinds of studies potentially acceptable to WWC represent only a small part of that universe. (See the framework developed by the National Research Council (2004), discussed below). That critique notwithstanding, one can take the comparison studies on their own terms. The question is, what kinds of information does WWC, which is intended as a sort of "consumer's guide to curricula," actually provide?

The first topic report produced by WWC, concerning middle school mathematics curricula, was produced in December 2004. Here is WWC's summary of the evidence base.

> From a systematic search of published and unpublished research, the What Works Clearinghouse (WWC) identified 10 studies of 5 curriculum-based interventions for improving mathematics achievement for middle school students. These include all studies conducted in the past 20 years that met WWC standards for evidence… The WWC identified 66 other studies that included evaluations of 15 additional interventions. Because none meets the WWC standards for evidence, we cannot draw any conclusions about the effectiveness of these other 15 interventions. The WWC also identified an additional 24 interventions that did not appear to have any evaluations. (What Works Clearinghouse, 2004a)

These are indeed slim pickings. Moreover, they are controversial, and for good reason. As noted above, there are a wide range of studies other than randomized controlled trials, regression discontinuity designs, and quasi-experimental designs with equating that can provide valuable information about curricular effectiveness. Thus WWC can be accused of being far too narrow in its criteria for what counts as documented evidence of effectiveness. But even if one restricts one's attention to just the types of studies examined by WWC, the way in which WWC has chosen to report the studies that do meet its methodological criteria raises serious issues. Viadero (2004) described the controversy over one of those studies (What Works Clearinghouse 2004b), as follows:

> James J. Baker, the developer of a middle school mathematics program known as Expert Mathematician, is also dismayed at the way his research on the program is reported. His study—the only one that fully met the criteria for this topic—used a random assignment strategy to test whether students could learn as much with this student-driven, computer-based program as they could from a traditional teacher-directed curriculum known as Transition Mathematics. The problem, he argues, is that the [WWC] web site said his program had no effect without explaining that students made learning gains in both groups. (p. 32)

This issue is important in applied terms, for a study that says there were "no significant differences" due to the use of a particular curriculum may be taken by readers to mean that the curriculum in question had no beneficial impact. This problem can, of course, be can be resolved in a straightforward way. The summaries provided by WWC can be re-written to indicate the size of the gains made through the use of each curriculum.

Another study report, however, WWC's (2004c) detailed study report of C. Kerstyn's (2001) evaluation of the I CAN LEARN Mathematics Classroom, demonstrates what may be a fatal flaw in the nature of current WWC reports. Kerstyn's study employed a quasi-experimental design with matching. It met the WWC standards "with reservations"; there were concerns regarding the implementation fidelity of the curriculum, some sampling issues, and issues regarding which subgroups of students were tested. However, the statistical reporting in the study fully meets WWC criteria. In discussing the measures used, the WWC report says:

> The fifth outcome is the Florida Comprehensive Assessment Test (FCAT), which was administered in February 2001. The author does not present the reliability information for this test; however, this information is available in a technical report written by the Florida Department of Education (2002). This WWC Study Report focuses only on the FCAT measures, because this assessment was taken by all students and is the only assessment with independently documented reliability and validity information.

This is deeply problematic. Reliability and validity scores represent psychometric properties of the FCAT; they say *nothing* about the actual mathematical content of the examination. In other words, the report provides no information about the mathematics actually covered by the measure. As a result, the report cannot be interpreted in meaningful ways – and it could be seriously misinterpreted. Consider the data given earlier in this section regarding students' differential performance on the SAT-9 and the MARS examinations. Is the FCAT more like the former or the latter? What content does it emphasize? Does it focus on procedural knowledge, or does it demand some relatively sophisticated problem solving skills? Unless WWC provides an independent auditing of the examination's contents, it is impossible to say what students actually learned. This raises the possibility of individual "false positives" and curricular "false negatives," as described earlier in this section. Moreover, not knowing what the outcome measures actually test makes it impossible to conduct meaningful meta-analyses of studies of the same curriculum. The author has urged WWC to re-do its analyses and to include content analyses of all assessment measures in its mathematics studies, so that readers of its study reports can determine what students actually learned. Unless and until the reports are re-done, the value of the entire enterprise is in question.

How should one study curricular effectiveness?

Determining whether and in what ways something as complex as a curriculum "works" is a complex matter. One view of the subject, rather different from that put forth by WWC, is offered by the National Research Council's (2004) committee for the review of the evaluation data of the effectiveness of NSF-supported and commercially generated mathematics curriculum materials. The committee's report, which urges methodological pluralism, presents a broad-based framework for evaluating curricular effectiveness.

The NRC framework is bipartite. First, it articulates a *program theory* – in essence, a description of what needs to be examined in the evaluation of an instructional program. Foci for examination include program components (including the mathematical content of the program and curricular design elements), implementation components (including resources, processes, and contextual influences), and student outcomes (including multiple assessment, enrollment patterns, and attitudes).

The second part of the framework elaborates the kinds of decision-making to be made by program evaluators. In any evaluation, there are methodological choices: What does one decide to look at, and how? The NRC panel points to three intellectually robust ways to examine curricula: content analyses, comparative studies, and case studies. It notes that there is a wide range of things to look at, and a wide range of rigorous ways in which such evaluations can be conducted. It points to the ways in which all of these kinds of studies can contribute to the field's collective understanding of curricular impact. The methodological pluralism of the NRC report stands in sharp contrast to the narrowly defined criteria employed by the What Works Clearinghouse.

A proposal by Burkhardt & Schoenfeld (2003) stakes out a middle ground in terms of breadth and focus. The authors argue that consumers of educational materials would profit from having access to reports that describe the conditions under which curricula can be successfully implemented, and on the kinds of results one might expect under those conditions; they should also be warned about conditions that make it unlikely for a particular curriculum to succeed. Thus, both comparative studies and benchmarking studies (using a stable and rigorous set of standards and outcome measures) would help inform those who are faced with curricular choices.

Burkhardt & Schoenfeld argue that the leap to experimental studies in education is premature. In both engineering and medical studies, research proceeds in stages. The first sets of studies typically include the design of prototypes (whether products or treatments) and the close observation of their effects under very controlled (and narrow) circumstances. This corresponds to design experiments (see, e.g., Brown, 1992; Collins, 1992; Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003; Schoenfeld, in press) or "alpha testing" of curricula. The goals of such work are to understand what is happening: to develop theory and instruction in dialectic with each other, under "greenhouse" conditions. Once the phenomena are understood, it is time to broaden the range of conditions of implementation. Here the issue becomes: in what conditions does the "treatment" function, in what ways? Which factors shape the implementation, with what results?

One engineering analogy among many is that some equipment is appropriate for rough terrain, while other equipment will function well on smooth but not rough terrain. Obviously, it is important to know this. Medical analogies are that one needs to discover how a medicine will be affected if it is taken with or without food, and whether the medicine interacts with other medicines or with particular conditions. Presumably there are factors that can affect the success of a curriculum in similar ways. What degree and kind of teacher professional development is necessary? How well is the curriculum suited for second language learners, and how can it be modified to make it more accessible if need be? What kinds of prerequisites are necessary? And more. The "beta stage" of curriculum study would consist of a planned series of observational evaluations of curricula in a carefully chose range of contexts: urban, suburban, and rural schools with a range of demographic factors, including the competency of the teacher corps, the amount of curriculum-specific professional development obtained by the teachers, the demographics of the student body, the availability and character of backup support for students, and so on. When these conditions are met – that is, when one is in a position to say "in a school or school district that looks like *this,* you should expect to provide *these* specific curricular supports, and then you can expect the following spectrum of results" – then one is ready to proceed to the most informative kinds of wide-spread experimental or comparative ("gamma stage") testing.

To sum up, the state of the art is still somewhat primitive. Much needs to be done along the lines of instrumentation (the development of robust outcome measures covering a wide range of expected mathematical content and processes), the creation of observational protocols, and the conduct of the wide range of studies described in *On Evaluating Curricular Effectiveness* (National Research Council, 2004). The challenges are not necessarily theoretical, although some theoretical work will need to be done; rather, this is in large measure a challenge of instrumentation, incentives, and implementation (Burkhardt & Schoenfeld, 2003).

Equity and Diversity in Mathematics Education

This topic area and the next, "Learning in Context(s)," are deeply intertwined, in that they both cross borders of classroom and culture. The division between the two is thus somewhat arbitrary. The same is the case for any separation between mathematics education and education writ large concerning these issues. For example, while the conditions of poverty described in Kozol (1992) unquestionably contribute to racial performance gaps in mathematics (see, e.g., J. Lee, 2002; Schoenfeld, 2002b; Tate, 1997), they contribute to performance differences in other fields as well. Hence in this section and the next, I will sketch the larger surround and then point to particular pieces within mathematics education.

The issues here are especially complex, because they cross traditional disciplinary boundaries as well. Many fields have *something* to say that informs our collective understanding of, for example, why different ethnic, racial, socioeconomic, linguistic, and gender subgroups of the population perform differently on a wide range of measures. However, while each casts some light on the phenomena, the illumination is partial and the underlying theoretical perspectives are often different.

For example Willis (1977/1981) makes it clear that issues of identity are central to one's participation (or not) in school practices. The students Willis's examined defined their personal affiliations along class lines. Those affiliations shaped their interactions with schooling – and thus, in large measure, the outcomes. Nothing in Willis's book is specific to mathematics classrooms. Yet, it clearly applies, at least in broad-brush terms: the students' *identities* shape their participation in all classrooms, including mathematics. Consider as well the econometric analyses of the contributions of schooling and parental economic status to people's economic success found in Bowles and Gintis classic (1976) volume *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. This says *something*, at least in correlational terms, about mathematical performance. The grand theoretical issue is how to meld such theoretical perspectives, and other powerful perspectives, into or with the sociocultural and cognitive perspectives that now predominate in discipline-oriented fields such as mathematics education. This is not merely a matter of one perspective subsuming another, or of foregrounding and backgrounding. The challenge is to build a theoretical and empirical program that provides leverage for the examination and explanation of the phenomena that are considered central to each of the constituent perspectives.

This section begins with a brief reprise of data indicating the reasons that "equity" and "diversity" have been, and remain, major concerns. It then discusses a series of efforts within mathematics education to redress some of the inequities documented by the literature. It concludes with a discussion of a theoretical reconceptualization of these issues offered by Cobb and Hodge.

Statistics on what have come to be known as "racial performance gaps" can be found in J. Lee (2002), National Science Foundation (2000), and Tate (1997). As is well known, when data on in mathematics course-taking, course grades, high school graduation rates, scores on national examinations such as the SAT, and college enrollments are disaggregated by racial or ethnic groups or by socioeconomic status, one sees persistent and substantial differences – to the disadvantage of Latinos, African Americans, Native Americans, and poor children. The differences are consequential. These are the words of noted civil rights leader Robert Moses:

> Today … the most urgent social issue affecting poor people and people of color is economic access. In today's world, economic access and full citizenship depend crucially on math and science literacy. I believe that the absence of math literacy in urban and rural communities throughout this country is an issue as urgent as the lack of Black voters in Mississippi was in 1961. (Moses & Cobb, 2001, p. 5)

In brief, Moses' argument is that a lack of mathematical and scientific literacy leads to economic disenfranchisement. Moses' approach to the problem, called the "Algebra Project" and described in Moses & Cobb (2001), is focused on providing mechanisms of mathematical enfranchisement for students. The project began by providing students with a set of empirical experiences that served as a basis for internalizing certain mathematical notions – for making them personally meaningful, so that mathematical formalization (via a process called the "regimentation of ordinary discourse" motivated by the ideas of the philosopher and mathematician Willard Van Orman Quine) became the codification of personally meaningful experiences rather than a set of instructions for operating on abstract symbolic structures. This idea of rooting mathematics in personally meaningful experiences lies at the core of work done in the Algebra Project. But, the project goes far beyond that. Moses' goals are ultimately those of the civil rights movement.

> It was when sharecroppers, day laborers, and domestic workers found their voice, stood up and demanded change, that the Mississippi political game was really over. When these folk, people for whom others had traditionally spoken and advocated, stood up and said, "We demand the right to vote," refuting by their voices and actions the idea that they were uninterested in doing so, they could not be refused… To understand the Algebra Project you must begin with the idea of our targeted young people finding their voice as sharecroppers,

> day laborers, maids, farmers, and workers of all sorts found theirs in the 1960s. (Moses, 2001, p. 20)

Thus, in Moses' view (but in this author's interpretation and phrasing), issues of voice, issues of entitlement, issues of responsibility, and issues of identity are all central concerns when considering enfranchisement in mathematics. This perspective is shared in various ways by a number of authors who view mathematics through a social justice lens. Martin (2000; in press), for example, identifies an aspect of identity that he calls *mathematics identity.*

> Mathematics identity refers to the dispositions and deeply held beliefs that individuals develop about their ability to participate and perform effectively in mathematical contexts (i.e., perceived self-efficacy in mathematical contexts) and to use mathematics to change the conditions of their lives. A mathematics identity therefore encompasses how a person sees himself or herself in the context of doing mathematics (i.e. usually a choice between a competent performer who is able to do mathematics or as incompetent and unable to do mathematics). (Martin, in press, p. 10)

Martin (2000) presents case studies of under-represented minority students who succeed in mathematics at school while the vast majority of their peers do not. His work indicates that those students tend to have a sense of personal agency (closely related to their mathematical identities) that has them act in ways that, at times, defy the norms and expectations of their peers and others. Martin's research also indicates, in the same ways that Shirley Brice Heath's (1983) study of literacy patterns does, that membership in different subpopulations (or perhaps subcultures) of the population at large tends to provide very different affordances for participation in school practices. Martin (in press) expands upon these ideas by considering, in the case of African Americans, the potential conflicts between one's mathematics identity and one's racial identity: "for [the subject of his study] and other African Americans like him, there is often a struggle to maintain and merge positive identities in the contexts of being African American and being a learner of mathematics. This struggle is brought on by a number of forces that racialize the life and mathematical experiences of African Americans."

Martin's language, while somewhat different from Wenger's (Wenger, 1998), is entirely consistent in theoretical terms. Wenger writes in terms of an individual's (unitary) identity – a "work in progress," shaped by both individual and collective efforts "to create a coherence in time that threads together successive forms of participation in the definition of a person" (p. 158). Wenger stresses that:

> We all belong to many communities of practice, some past, some current; some in more peripheral ways. Some may be central to our identities while others are more incidental. Whatever their nature, all these various forms of participation contribute in some way to the production of our identities. As a consequence, the very notion of identity entails
>
> 1) an experience of multimembership
> 2) the work of reconciliation necessary to maintain one identity across boundaries. (Wenger, 1998, p. 158)

If one reads Martin's "mathematics identity" and "racial identity" as aspects of one larger identity, then one sees Martin's case studies as cases in point for Wenger's theoretical claims. And, to be explicit: just as participation in communities of practice has ramifications for one's construction of identity, one's identity shapes patterns of participation. Hence the issue of "multimembership" in different communities of practices, and the affordances each community offers for the individual, are central. Are the practices that constitute or signal membership in one community (say the mathematics classroom) consistent with those of another community that is central to one's identity (say one's home life, or peer group)? Do they build on one's perceived strengths, or do they negate or undermine them?

Some of the most promising empirical and theoretical work is grounded in these underlying perspectives. There is, for example, the idea that students have "funds of knowledge" (González, Andrade, Civil, & Moll, 2001; Moll, Amanti, Neff, & González, 1992; Moll & González, 2004) upon which school knowledge can be built and expanded, rather than having "deficits" that need to be remediated. This perspective is also powerfully demonstrated in work by Carol Lee (C. Lee, 1995) and explicated by K. Gutiérrez and Rogoff (2003). There is the related idea that culturally responsive pedagogy (Ladson-Billings, 1994, 1995) meets the "whole child" and creates a classroom community of students who feel and are empowered to learn. In various ways, these underlying perspectives are represented in analyses by Brenner (1994; Brenner & Moschkovich, 2002), R. Gutierrez (1996; 2002), Khisty (1995; 2002), and Moschkovich (1999; 2002; in press). Empirically, they are embodied in a number of powerful attempts at teaching mathematics or science for social justice. One of the longest established and best-known programs is Chèche Konnen. The project (see, e.g., Rosebery, Warren, & Conant, 1992; Rosebery, Warren, Ogonowski & Ballenger, 2005; Warren, Ballenger, Ogonowski, Rosebery, & Hudicourt-Barnes, 2001; Warren & Rosebery, 1995) emphasizes the sense-making resources that children from ethnically and linguistically diverse backgrounds bring to the study of science, and how instruction can build on them.

A second series of papers with a social justice focus in mathematics comes from Gutstein and colleagues (Gutstein, 2003, 2005; Gutstein, Lipman, Hernández & de los Reyes, 1997). Gutstein takes direct aim at issues of social justice and at empowering students as advocates for themselves and their communities by assigning projects that use mathematics to address social injustices. For example, Gutstein asks his students to consider the cost of a B-2 bomber and asks his students to compute how many students could be given fellowships to a major university if the money were used for that purpose instead.

Gutstein and others teaching for social justice recognize that they, like anyone teaching a somewhat non-standard mathematics class, are the servants of at least two masters. They will be held accountable for their students' performance on standard mathematical content, and then for whatever additional goals they have for instruction. Thus, Gutstein (2003) provides evidence of his students' mathematics learning according to traditional, standardized measures. He also provides evidence of his students' empowerment – of their eagerness after his course to use their mathematical knowledge to address issues of social justice.

A third example, "Railside School," is discussed in the next section. But, it is worth noting here as an exemplar of a coherent attempt on the part of a high school mathematics department to create a culture, both for staff and for students, that supports a strong equity agenda. Boaler (in press) documents the disappearance of racial performance gaps in mathematics scores at Railside. The claim is that such results were possible because of a department-wide effort that focused on: a curriculum that allowed all students to engage with meaningful mathematics; a pedagogy grounded in the assumption that all students are capable of grappling meaningfully with mathematically rich problems, and have something to contribute to their solution; as part of that pedagogy, a set of classroom accountability structures that hold students accountable to each other and to the teacher for very high standards of mathematics; and a number of mechanisms for supporting the teachers in implementing the pedagogy just described. (Boaler, in press; Horn, 2003.)

Each of the efforts described in this section, in different ways, considers individual learners as members of a number of different communities and as participants in different Discourses (in the sense of Gee, 1996) associated with each of those communities; each is concerned with continuities and discontinuities between those communities and the Discourses in them. In various ways, the educational efforts try to bridge the discontinuities. This interpretation of these efforts is consistent with a re-framing of the issues of diversity and equity put forth by Paul Cobb and Linn Liao Hodge in a seminal piece entitled "A relational perspective on issues of cultural diversity and equity as they play out in the mathematics classroom" (Cobb & Hodge, 2002).

One version of the standard notion of diversity might be as follows: the diversity of a group is related to the number of members of the group who come from different ethnic/racial/socioeconomic/gender/sexually-oriented/other subgroups of the general population. Cobb and Hodge offer the following alternative framing:

> We propose to conceptualize diversity relatively broadly in terms of students' participation in the practices of either local, home communities or broader groups or communities within wider society. … Equity as we view it is concerned with how continuities and discontinuities between out-of-school and classroom practices play out in terms of access (Cobb & Hodge, 2002, p. 252).

Cobb and Hodge's perspective and the standard notion can result in different views of the diversity of a particular group. A recent example of the difference in the two characterizations and their entailments came in series of news articles about the town of Cupertino, a wealthy enclave in Silicon Valley that has very high-performing schools – and real-estate values to match. The town was characterized in the media as "diverse." It is diverse, in the standard sense:

the town is populated by Whites and a wide range of (mostly Asian) ethnic minorities. Cobb and Hodge's characterization provides an alternative view. Part of the reason Cupertino's schools do as well as they do is that parents with particular sets of values and practices (and the incomes to back them up) comprise a huge majority of the town's inhabitants. In that sense there are relatively few discontinuities between some important in-school and out-of school practices for the children who attend school the district – i.e., not tremendously much diversity along some of the main dimensions that "count" in this context.

A shift to Cobb and Hodge's characterization of diversity has some slight drawbacks but also some potentially significant advantages. One disadvantage of the definition is that it obscures the traditional focus on the differential treatment of different subgroups of the general population – e.g., the data on racial performance gaps that make so tangibly clear the inequities of our educational system. On the flip side, however, this kind of definition takes a significant step away from the stereotyping and essentializing that often come as an entailment of classification – the idea that members of any particular ethnic, racial, socioeconomic, gender, or other group all share important attributes by virtue of their membership in that group, and therefore can be "pigeonholed" and treated in ways appropriate for members of those groups.

The main advantage of the shift in perspective, however, may be its shift in focus. Learning is a concomitant of the practices in which we engage. Hence a focus on the opportunities that various contexts provide to individuals to engage in particular kinds of practices provides a central lens on learning. The opportunities available to each individual are clearly a function of the continuities and discontinuities between the practices of the different communities in which that individual is a member. This framing, thus, is a potentially useful lens with which to view all of learning. It leads us to the next section of this chapter.

Learning in Context(s)

This section focuses directly on the issue of learning. Somewhat more than a quarter century ago, the field of cognitive science began to coalesce. Its emergence as an interdisciplinary enterprise was motivated in large part because its varied constituent disciplines – among them anthropology, artificial intelligence, education, linguistics, neurobiology, philosophy, and psychology – all offered partial views of a phenomenon (cognition) that was too big for any one of them to grasp individually. As a result of that coming together, tremendous progress has been made over the past few decades. That progress can be seen in the flowering of understandings and results described in the two editions of this *Handbook.* The outlines of individual pieces of the puzzle of learning are beginning to become clear. But, how they fit together is still at issue. This section describes some of those pieces, and some of what remains before they can be put together. It begins with some brief additional commentary on the character of the puzzle, and then works its way down from the big picture to more fine-grained issues of mechanism, and an evaluation of the state of the art.

The issue is how to put things together – how to see everything connected to an individual (both "internally" in the sense of knowledge, identity, etc., and "externally" in terms of that person's relationship to various communities) and the communities to which the individual belongs as a coherent whole. Ideally, one would like to be able to understand the evolution of individuals and communities as well. Obviously, there are issues of grain size: some phenomena are "macro" and some "macro," and different explanatory lenses (local theories) will be appropriate for focusing in on different levels. But, the linkages should be smooth, in the same sense that (for example) a "big picture" theory of ecosystems should frame the discussion of the ecology or a particular region, establishing the context for a discussion of the evolving state of classes of organisms in that region. A more micro view of a specific class of organisms in that region describes how they live in interaction with that region, and yet more micro analyses describe the anatomy and physiology of individual organisms (and so on).

I consider one broad metaphor before returning to mathematics education. It is in the tradition of an early paper by Greeno (1991), which compared learning in a content domain (and the affordances of various symbolic and physical tools therein) with learning to make one's way comfortably around a physical environment.

> Learning the domain . . . is analogous to learning to live in an environment: learning your way around, learning what resources are available, and learning how to use those resources in conducting your activities productively and enjoyably." (p. 175). . . "In [pursuing] the metaphor of an environment such as a kitchen or a workshop, this section is about knowing how to make things with materials that are in the environment. . . ." (p. 177)

Motivated by Greeno's metaphor and a passion for food, the author (Schoenfeld, 1998b) pursued some parallels between learning to cook and learning mathematics, e.g., the development of skills and the ability to perceived and take advantage of affordances in the environment, and the character of memory and representation in the two domains. Here I would like to pick up the metaphor, in terms of the big themes introduced in this chapter – e.g., the relationships between identity, knowledge, and community, and how each can be foregrounded and explored in ways that lead continuously from one to the other. To put things simply: being a "foodie," like being a mathematician and being an educational researcher, is part of the author's identity. Manifestations of that aspect of the author's identity are easy to spot, both materially and personally: materially in his well-equipped kitchen and in a large collection of cookbooks and of restaurant guidebooks, personally in his ongoing practices (taking pains to prepare meals, stopping off at specialty stores to buy provisions for dinner, using food metaphors in his research group, and occasionally mixing his food and work identities by writing about both). If one were interested in understanding these aspects of the author's identity, his personal history would be clearly important; so would membership in various (sometimes distributed, but clearly defined) communities and the role they played not only in the development of identity but in the development of skills and understandings. Deeply intertwined with identity is a body of skills and knowledge – his knowledge that particular dishes are best made in particular kinds of pans, that a particular preparation calls for a blazingly hot pan while another calls for gentle heat; his knowing (by sight, or "feel," or other input) when a particular stage in the cooking process is done; and more. Call this a "knowledge inventory" if you will; the fact is that no description of the author as cook is complete without a thorough categorization of the set of skills, practices, and understandings he possesses (cf. Hillman, 1981; McGee, 2004). This can be done, more or less in standard cognitive science tradition. What is called for in theoretical terms is specifying the linkage between the author's identity and knowledge base. One can imagine ways to specify the linkage: narrative stories of the protagonist's enjoyment of food; descriptions of familial and other practices that enhanced that enjoyment; a characterization of the support he had in developing various culinary practices, on his own and in interaction with others; and the details of that support structure and those interactions, which gave rise to the entries in the knowledge inventory and the coherence among them. This set of issues – explaining and linking aspects of identity, participation, and knowledge – is by analogy the set of issues one confronts when trying to paint the big picture in (mathematics) learning as well.

A recent paper by Saxe and Esmonde (in press) takes on some of the large-scale issues related to the dialectic shaping, over time, of relationships between individual and community. In a distillation of research that includes field studies conducted in 1978, 1980, and 2001 and a historical analysis that covers 1938 to the present, Saxe and Esmonde examine the micro- and macro-changes in the counting systems used by inhabitants of the Oksapmin valleys in the highlands of central New Guinea. In the mid-20th century, the Oksapmin people used a body part counting system, counting digits and pointing to different places on their bodies. By the time of Saxe's first visit in 1978, commercial incursions from the West had put pressure on the indigenous people to switch (at least in some interactions) from a trade-based economy to a cash-based economy involving Western currency such as pounds and shillings. The effects of those changes were recorded in Saxe's early work. When he returned in 2001, he saw further changes in the body count system. Saxe and Esmonde (in press) present an analytic framework to guide their analyses of the interplay between the social history of the Oksapmin and the development, over time, of new forms of mathematical represent5ation and thought. The framework is fundamentally *cultural;* it is also fundamentally *developmental.* The authors present intertwined arguments at three levels: microgenetic, sociogenetic, and ontogenetic. Microgenetic analyses show the ways in which individuals "turn cultural forms like the body system into means for accomplishing representational and strategic goals" (p. 65). Sociogenetic analyses focus on the ways in which such changes become part of a community. The argument is that (as with some theories of language development) at first the new developments are synchronic, developing locally in a variety of locations (in this case, in trade stores, where the Western currency began to displace the body count system). Later the process of change is diachronic (taking place over a longer time interval), as interlocutors from different sites encounter each other and need to negotiate shared meanings. Finally, there is the issue of ontogenesis, used here to mean changes in the organization of cognition over the course of an individual's life span. Interviews with individuals at least suggested a longitudinal progression of conceptual growth.

The work by Saxe and Esmonde addresses some of the same issues, though from a somewhat different perspective, that are addressed by Engeström (1987, 1993, 1999). Engeström's activity-theoretic framing of the issues situates individual activities amidst a nexus of complex social structures, highlighting the tensions negotiated by individuals and communities over time. This theoretical structure was used by K. Gutiérrez, Baquedano-Lopez, & Tejeda (1999) to explore what they call "third spaces," zones of development that can open up within classrooms to accommodate productive activities by diverse sets of learners. As noted in the previous section, Wenger (1998) provides a theoretical framework that focuses on issues of practice(s) and (aspects of) identity, and the dialectic relationships between them.

Issues of identity are explored in interesting ways by Nasir and Saxe (Nasir, 2002; Nasir & Saxe, 2003). Drawing upon Wenger's (1998) and Saxe's (1999) frameworks, Nasir (2002) illustrates the ways in which individuals change as they engage in the practices of playing dominos and basketball. Her analyses indicate that as individuals become more accomplished, their goals change; their relationships to the communities of practice (domino and basketball players) change; and their own definitions of self relative to the practices change. Nasir illustrates the bidirectional character of relations between identity, learning, and goals. (Learning creates identity, and identity creates learning; and so on.) Nasir & Saxe (2003) examine different facets of identity – ethnic identity and academic identity – and point to circumstances in which the two may be in conflict.

As noted in the previous section, Cobb and Hodge (2002) offer a theoretical framing of issues related to the continuities and discontinuities between practices in which individuals engage, inside the classroom and outside, that serves as a useful lens with which to examine all learning. Much of the work in the previous section of this paper is grounded in, or at least consistent with, this perspective. Let us now focus more directly on the mathematics classroom.

At the broad level of linking practices to outcomes, Boaler (2002; in press; Boaler & Greeno, 2000) has identified the characteristics of different communities of classroom mathematical practice and their impact both on student performance and on aspects of identity. In her book *Experiencing School Mathematics,* Boaler (2002) describes two very different environments. “Amber Hill” was the very embodiment of exemplary “traditional” mathematics instruction. It had hard-working and professional teachers, a clearly specified curriculum (the English National Curriculum), and a straightforward, department-wide approach to instruction. There was a very high rate of “time on task” as Amber Hill students watched teachers model the solutions to problems at the blackboard, and then worked collections of problems on worksheets.

“Phoenix Park” school had a population similar to Amber Hill in terms of demographics, but a radically different approach to mathematics instruction. The curriculum was “problem-based,” with little emphasis on drill. One problem, for example, was for students to find as many shapes as they could whose volume was 216. Once the problem was assigned, teachers then worked with individual students, tailoring the problem to the students’ needs and skills. Students had a great deal of autonomy, and time on task was very low at times.

Boaler used multiple measures to determine the outcomes, in the aggregate and by way of individual descriptions. In the aggregate, there were few differences on skills-oriented tasks between students at the two schools – but there were differences in perspective. Students at Amber Hill felt qualified only to solve problems that were nearly identical to problems they had worked, and they were uncomfortable at times even with that.

Similar patterns were found in American schools studied by Boaler and Greeno (2000). In the "ecologies of didactic teaching," students view their roles vis-à-vis mathematics as passive memorizers; in the "ecologies of discussion-based teaching," students are active collaborators and co-constructors of knowledge:

> J: The teacher gives us something and has us work on a work sheet, because if I understand something, then I can explain it to the group members or if I don't understand it the group members may explain it to me. Whereas if she teaches the lesson and sends us home with it, I'm not really that confident because I haven't put like things together (Boaler & Greeno, 2000, p. 178).

Boaler (in press) pursues this issue in more fine-grained detail, examining the accountability structures by which the teachers at "Railside School" hold students, and the students hold each other, accountable for producing clear and cogent explanations of the mathematics under consideration. In a lesson described by Boaler (in press), one member of a group is asked a question by the teacher. When the student does not produce a viable explanation, the teacher simply says "I'll be back." The group knows the teacher will return to ask the same student the same question again. One member of the group gives the student a quick tutorial, saying "answer it this way." But the student resists. She argues that the teacher will not be satisfied with a pat answer – that the teacher will probe until the student produces an explanation that is mathematically correct and stands up to robust questioning. It is the group's responsibility to make sure every member of the group understands and can explain the material. The group takes on that responsibility, with the result that the student does come to understand the material – and, after demonstrating her understanding of it (withstanding tough questioning from the teacher), she is clearly more self-assured as well. In the classroom videotapes from

Railside one sees, at a micro-level, the ways in which classroom practices interact with issues of individual identity and knowledge.

The examples given above are cases in point for an argument by Engle and Conant (2002), that there tend to be the following substantial consistencies in some of the most productive learning environments for students:

- *Problematizing*: students are encouraged to take on intellectual problems
- *Authority*: Students are given authority in addressing such problems
- *Accountability*: Students' intellectual work is made accountable to others and to disciplinary norms
- *Resources*: Students are provided with sufficient resources to do all of the above. (Engle and Conant (2002, pp. 400-401).

The discussion of the Railside example also brings us explicitly to the issue of mechanism – the means by which the dialectic between individual and collective is worked out, with each being shaped as a result. Central to the study of classroom practices, of course, is the study of patterns of classroom discourse.

There are at least two useful theoretical notions involved in the discussion of the Railside classroom above. The first is *sociomathematical norms*. Erna Yackel and Paul Cobb (Cobb & Yackel, 1996; Yackel & Cobb, 1996) adapted the concept of general social norms to describe situations that specifically involve patterns of *taken-as-shared mathematical behavior*: "Normative understandings of what counts as mathematically different, mathematically sophisticated, mathematically efficient, and mathematically elegant are sociomathematical norms. Similarly, what counts as an acceptable mathematical explanation and justification is a sociomathematical norm." (Cobb & Yackel, 1996, p. 461)

The second closely related notion is that of *accountability structures.* In the Railside classroom studied by Boaler, the students are accountable to the teacher, to each other (in that the group is responsible for making sure that all of its members understand the mathematics), and to the mathematics – their discussions and explanations are expected to be rigorous and to meet high mathematical standards. Ball and Bass (2003b) provide examples of ways in which a third grade class, over the course of the school year, comes to grapple with such issues. Horn (in preparation) provides a fine-grained analysis of the accountability structures in Deborah Ball's well known January 19, 1990 class. Some of the critical aspects of what Horn calls *accountable argumentation* are that: it uses terms from the mathematical and academic registers (e.g. "proof," "conjecture"); discussions have a slow and measured pace; and, disagreements are important and may not (and need not) be resolved. Horn argues that

> In this classroom, accountable argumentation brings the often hidden practices of mathematical reasoning into the visible world of classroom interactions…
>
> During the class session, these thinking activities are focused on a particular set of ideas. Accountable argumentation supports engagement with specific ideas, particularly through the expectations that (a) students attend to whole class discussions and (b) students take a justified position in a discussion which they will act on or defend. In addition, once they are engaged in a disagreement, the stakes for engagement increase. Dissenters are obliged to ask questions or otherwise substantiate their position to their peers. Principals, on the other hand, must articulate their thinking to the whole class.
>
> Effectively, these thinking activities and the engagement in particular ideas support both the learning and creation of mathematics. (Horn, 2005, p. 25)

At an equally fine level of grain size, and also focusing on issues of mechanism, are studies of teachers' discourse moves such as those conducted by O'Connor (1998) and O'Connor and Michaels (1993; 1996). O'Connor and Michaels characterize a teacher move they call *revoicing.* In revoicing, a teacher picks up on a comment made by a student and draws attention to it – sometimes paraphrasing, sometimes clarifying, sometimes commenting on its relevance or importance. This act can legitimate and give status to a student; it can bring his or her ideas to center stage; it can position the student as author of the comment and place the student at the center of a dialogue, to which other students are expected to participate. All of these moves can contribute to the creation of a classroom discourse community in which students are given authority to work on consequential problems, positioned as knowledgeable members of the community, and attributed ownership of important ideas. This kind of discourse move on the part of teachers stands in stark contrast to traditional classroom discourse communities in which the teacher typically initiates discussion with a "short answer" question, a student responds, and the teachers evaluates the response ("IRE sequences"; see Mehan, 1979).

Also at the level of mechanism, there is the question of what actually takes place in extended (and not always productive) interactions between students. Sfard and Kieran (2001) present a detailed analysis of a series of interactions between two 13-year-old boys learning algebra over a two month period. They introduce the notions of *focal* and *preoccupational* analyses as analytical tools – the former for giving direct attention to the mathematical content contained in students' interactions, and focusing on communication and mis-communication between the two students, the latter focusing on meta-messages and engagement, providing tentative explanations for some of the students' communication failures. The authors note their conclusions as follows:

> We realized that the merits of learning-by-talking cannot be taken for granted. Because of the ineffectiveness of the students' communication, the collaboration we had a chance to observe seemed unhelpful and lacking the expected synergetic quality. Second, on the meta-level, we concluded that research which tries to isolate cognitive processes from all the other kinds of communicative activities is simply wrongheaded… For us, thinking became an act of communication in itself. This re-conceptualization led to the disappearance of several traditional dichotomies that initially barred our insights: the dichotomy between "contents of mind" and the things people say or do; the split between cognition and affect, and the distinction between individual and social research perspectives. (Sfard & Kieran, 2001, p. 42)

This latter issue is also pursued in Sfard (2001).

Finally, there is a need to understand classroom discourse practices and the use of artifacts – and the development of shared meanings over both. A fascinating exercise in multiple interpretation was carried out in Sfard & McLain (2003), in which a series of authors with related but different theoretical perspectives examine the same set of video-recorded classroom data from an experiment in teaching statistics. The juxtaposition of theoretical perspectives in the issue shows the progress the field has made in untangling social-cognitive phenomena. See also Sfard (2000), which provides a detailed examination of how symbols come to take on meanings.

To sum up, all of the studies referenced in this section offer advances over the perspectives and tools available to the field when the previous edition of the *Handbook of Educational Psychology* was published. As the field has matured, it has begun to grapple with complex issues of learning in context(s). The studies referenced here represent points of light in territory that, not long ago, was largely uncharted. As such, there is progress. There is not enough light to illuminate the terrain; but there may be enough points of light to allow one to get a sense of the landscape.

The state of the field.

The main substance of this chapter has delineated thematic progress in a number of domains central to mathematics teaching and learning: research focusing on issues of teacher knowledge and aspects of professional development; issues of curriculum development, implementation, and assessment; issues of equity and diversity; and issues of learning in context(s). This brief concluding section takes a step back from the details to examine the contextual surround within which researchers in mathematics education do their work.

As this chapter indicates, there has been a fair amount of theoretical progress – with the theory being grounded in, and tested by, empirical findings. However, this steady progress has not been met with recognition or support outside the field; and research has not had nearly the impact on practice that it might. Moreover, the external context (including the funding environment) for high quality research in mathematics education is as hostile as it has been for at least a quarter century. Some of this is undoubtedly the control of the field, but (mathematics) educators may have contributed to some of it them/ourselves.

The political context

Educational research as a whole in the United States is under attack. Consider, for example, the following language from the U.S. Department of Education's Strategic Plan for 2002-2007:

> Unlike medicine, agriculture and industrial production, the field of education operates largely on the basis of ideology and professional consensus. As such, it is subject to fads and is incapable of the cumulative progress that follows from the application of the scientific

method and from the systematic collection and use of objective information in policy making. We will change education to make it an evidence-based field. <http://www.ed.gov/pubs/stratplan2002-07/index.html, p. 48>

That language does not represent an empty threat. "Evidence-based" has been taken to mean "quantitative," and a narrow band of quantitative at that; sources of funding for anything other than a narrow, quantitative research agenda are drying up (see below).

Funding

Funding for educational research has always been ridiculously low. In 1998 the U. S. House Committee on Science wrote, "currently, the U.S. spends approximately $300 billion a year on education and less than $30 million, 0.01 percent of the overall education budget, on education research…. This minuscule investment suggests a feeble long-term commitment to improving our educational system" (p. 46).

The vast majority of funding for research in science and mathematics education in the U.S. in recent years has come from the Education and Human Resources (EHR) Directorate of the National Science Foundation, with the lion's share of funding for basic research coming from EHR's Division of Research, Evaluation, and Communication (REC). The March 25, 2005 issue of Science Magazine contains the following information on a new $120 Million funding initiative focusing on the use of randomized controlled trials to test the effectiveness of mathematics curricula:

> The initiative comes at the same time the Administration has requested a $107 million cut in NSF's $840 million Education and Human Resources (EHR) directorate. The cuts include … a 43% decrease for the foundation's division that assesses the impact of education reform efforts (Science, 11 February, p. 832). [Assistant secretary for vocational and adult education

> at the Department of Education Susan] Sclafani says this "reallocation of education dollars" reflects the Administration's eagerness for clear answers on how to improve math and science learning across the country. That's OK with NSF Director Arden Bement, who says ED is in a better position than NSF to implement reforms nationwide. (Bhattacharjee, 2005, p. 1863)

The Division of EHR sustaining the 43% budget cut identified by Bhattacharjee is the Division of Research, Evaluation, and Communication. Given that the REC has ongoing funding obligations, the proposed cuts essentially bring to a halt the funding for new research projects within the Division. Of course, it remains to be seen what the value of the new project at the Department of Education will be. But, given the track record of the What Works Clearinghouse to date (see the discussion above), there is some reason for concern.

Impact on practice

As discussed in the section on curriculum, the educational R&D community lacks robust mechanisms for taking ideas from the laboratory into "engineering design" and then large-scale implementation. This is partly for fiscal reasons. (The design refinement process outlined in this chapter is costly. As long as publishers can sell books based on the results of focus groups and avoid the expenses of that design refinement process, they will.) It is also partly a result of academic value systems. In promotion and tenure committees, theory and new academic papers tend to be valued over applications; new ideas tend to be valued over replications, extensions, and refinements; single authored work is valued more than work in teams (shares of "credit" are notoriously difficult to assign). Those who would systematize the R&D process thus face an incentive system that devalues teamwork, applications, and iterative design. Here too, some changes would help the field to have greater impact.

It should be noted that a focus on the engineering model described here does not represent an endorsement of the "linear model" of research-into-practice. A substantial proportion of the work done under this aegis can and should be done in "Pasteur's Quadrant" (Stokes, 1997), contributing both to theory and to the solution of practical problems. Both design experiments (the initial phases of design) and contextual studies (explorations of the ways in which instructional interventions work) can and should contribute as much to theory development as they do to the creation of improved instructional materials and practices.

Some final words

These are, in Dickens' words, the best of times and the worst of times; an age of wisdom and an age of foolishness. This chapter documents substantial progress in research on mathematical teaching and learning over the decade since the publication of the first *Handbook of Educational Psychology.* There have been significant theoretical advances in many areas, and practical advances as well (although the data in substantiation of those advances are less robust than one would like). There is, in sum, good reason for optimism on the intellectual front. At the same time, the larger climate is remarkably hostile to the scholarly enterprise, and there appears to be little prospect of improvement in the short run. A short-term view of the situation would be pessimistic. However a sense of history suggests that support for and hostility to the research enterprise seem to come in cycles and that in the long run, intellectual progress is sustained. It will be interesting to see what progress is reflected in the next edition of this *Handbook.*

ACKNOWLEDGMENT

I am indebted to Jim Greeno and Lani Horn, Abraham Arcavi, Hugh Burkhardt, Mari Campbell, Charles Hammond, Vicki Hand, Markku Hannula, Manya Raman, Miriam Sherin, and Natasha Speer for their incisive comments on a draft version of this manuscript. This chapter is much improved for their help. The flaws that remain are all my responsibility. I would also like to express my thanks to Patricia Alexander and Lane Akers for their graciousness and support from the beginning to the end of the process of producing the chapter.

References

Arcavi, A. A., & Schoenfeld, A. H. (1992). Mathematics Tutoring Through A Constructivist Lens: The Challenges of Sense-Making. *Journal of Mathematical behavior,* 11(4), 321-336.

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics In J. Boaler (Ed.), *Multiple perspectives on the teaching and learning of mathematics* (pp. 83-104). Westport, CT: Ablex.

Ball, D. L., & Bass, H. (2003a). Making mathematics reasonable in school. In J. Kilpatrick, W. G. Martin, & D. Schifter, D. (Eds.). *A research companion to Principles and Standards for School Mathematics* (pp. 27-44). Reston, VA: National Council of Teachers of Mathematics.

Ball, D. L., & Bass, H. (2003b). Toward a practice-based theory of mathematical knowledge for teaching. In B. Davis & E. Simmt (Eds.), *Proceedings of the 2002 Annual Meeting of the Canadian Mathematics Education Study Group* (pp. 3-14). Edmonton, AB: CMESG/GCEDM.

Bhattacharjee, Y. (2005). Can randomized trials answer the question of what works? *Science, 307,* 1861-1863.

Beaton, A., Mullis, I., Martin, M., Gonzalez, E., Kelly, D., & Smith, T. (1997). *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Report.* Boston: The International Association for the Evaluation of Educational Achievement, at Boston College.

Berliner, D., & Calfee, R. (Eds.) (1996). *Handbook of Educational Psychology*. New York: MacMillan.

Boaler, J. (2002) *Experiencing School Mathematics* (Revised and expanded edition). Mahwah, NJ: Erlbaum

Boaler, J. (in press.) Promoting Relational Equity in Mathematics Classrooms – Important Teaching Practices and their impact on Student Learning. Text of a 'regular lecture' given at the 10th International Congress of Mathematics Education (ICME X), 2004, Copenhagen. To appear in the ICME X *Proceedings*.

Boaler, J. and Greeno, J. (2000). Identity, agency and knowing in mathematical worlds. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 171-200). Westport, CT: Ablex Publishing.

Bowles, S., & Gintis, H. (1976). *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life.* New York: Basic Books.

Brenner, M. (1994). A communication framework for mathematics: Exemplary instruction for culturally and linguistically diverse students. In B. McLeod (Ed.), *Language and learning: Educating linguistically diverse students* (pp. 233-268). Albany, NY: SUNY Press.

Brenner, M., & Moschkovich, J. (Eds.), (2002). *Everyday and academic mathematics in the classroom*. *JRME* Monograph Number 11. Reston, VA: NCTM.

Briars, D., & Resnick, L. (2000). *Standards, assessments – and what else? The essential elements of standards-based school improvement*. Unpublished manuscript.

Briars, D. (March, 2001). Mathematics performance in the Pittsburgh public schools. Presentation at a Mathematics Assessment Resource Service conference on tools for systemic improvement, San Diego, CA.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences,* 2(2), 141-178.

Burkhardt, G. H., & Schoenfeld, A. H. (2003). Improving educational research: toward a more useful, more influential, and better funded enterprise. *Educational Researcher 32(9), 3-14.*

Carpenter, T. P. (1985). Learning to Add and Subtract: An Exercise in Problem Solving. In E. A. Silver (Ed.), *Teaching and Learning Mathematical Problem Solving: Multiple Research Perspectives*, pp. 17-40. Hillsdale, NJ: Erlbaum.

Carpenter, T. P., Fennema, E., & Franke, M. L. (1996). Cognitively Guided Instruction: A knowledge base for reform in primary mathematics instruction. *Elementary School Journal,* 97, (1), 1-20.

Cobb, P., Confrey, J. diSessa, A., Lehrer, R., & Schauble, L. (2003). Design Experiments in Educational Research. *Educational Researcher*, 32(1), 9-13.

Cobb, P., & Hodge, L. L. (2002). A relational perspective on issues of cultural diversity and equity as they play out in the mathematics classroom. *Mathematical Thinking and Learning, 4*(2&3), 249-284.

Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist* 31(3-4), 175-190.

Cohen, D., Raudenbush, S., & Ball, D. (2003). *Resources, instruction, and research.* Educational Evaluation and Policy Analysis, 25 (2), 1-24.

Cohen, S. (2004). *Teachers' professional development and the lementary mathematics classroom.* Mahwah, NJ: Erlbaum.

Collins, A. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15-22). Berlin: Springer.

deCorte, E., Greer, B., & Verschaffel, L. (1996). Mathematics teaching and learning. In D. Berliner & R. Calfee (Eds.), *Handbook of Educational Psychology*, 491-549. New York: MacMillan.

Educational Researcher (2002, November). Theme Issue on Scientific Research and Education, 31(8).

Engeström, Y. (1987). *Learning by expanding*. Helsinki, Finland: Orienta-Konsultit Oy.

Engeström, Y. (1993). Developmental studies of work as a test bench of activity theory: The case of primary care medical practice. In S. Chaiklin & J. Lave (Eds.), *Understanding practice: Perspective on activity and context* (pp. 64Ð103). Cambridge, UK: Cambridge University Press.

Engeström, Y. (1999). Activity theory and individual and social transformation In. Y Engeström, R. Miettinen, & R. Punamaki (Eds.), *Perspectives on activity theory* (pp. 19-38). Cambridge, England: Cambridge University Press.

Engle, R., & Conant, F. (2002) Guiding principles for fostering productive disciplinary engagement: Explaining emerging argument in a community of learners classroom. *Cognition and Instruction 20*(4), 399-483.

Fernandez, C., & Yoshida, M. (2004). *Lesson study: A Japanese approach to improving mathematics teaching and learning.* Mahwah, NJ: Erlbaum.

Ferrini-Mundy, J., & Schram, T. (Eds.). (1997). The Recognizing and Recording Reform in Mathematics Education Project: Insights, issues, and implications. *Journal for Research in Mathematics Education, Monograph Number 8*. Reston, VA: National Council of Teachers of Mathematics.

Schifter, D. (Eds.). *A research companion to Principles and Standards for School Mathematics* (pp. 333-352). Reston, VA: National Council of Teachers of Mathematics.

Franke, M., Carpenter, T. P., Levi, L., & Fennema, E. (2001). Capturing teachers' generative change: A follow-up study of professional development in mathematics. *American Educational Research Journal, 38(3),* 653-689.

Gee, J. (1996). *Social linguistics and literacies: Ideology in discourses (second edition).* Philadelphia: Routledge/Falmer.

Goodwin, C. (1994). Professional Vision. *American Anthropologist, 96,* 606-633.

Gonzalez, N., Andrade, R., Civil, M., & Moll, L. (2001). Bridging funds of distributed knowledge: Creating zones of practices in mathematics. *Journal of Education for Students Placed at Risk (JESPAR), 6(1-2),* 115-132.

Greeno, J. G. (1991). Number sense as situated knowing in a conceptual domain. *Journal for research in mathematics education, 22(3)*, 170-218.

Greeno, J.G., Collins, A., & Resnick, L. (1996). Cognition and learning. In D. Berliner & R. Calfee (Eds.), *Handbook of Educational Psychology*, 15-46.

Greeno, J. G., Pearson, P. D., & Schoenfeld, A. H. (1997). Implications for the National Assessment of Educational Progress of Research on Learning and Cognition. In: *Assessment in Transition: Monitoring the Nation's Educational Progress, Background Studies*, pp. 152-215. Stanford, CA: National Academy of Education.

Gutiérrez, K., Baquedano-Lopez, P., & Tejeda, C. (1999). Rethinking diversity: Hybridity and hybrid language practices in the third space. *Mind, Culture, & Activity*, *6*, 286–303.

Gutiérrez, K, & Rogoff, B. (2003). Cultural Ways of Learning: Individual Traits or Repertoires of Practice. *Educational Researcher* 32(5), 19-25.

Gutierrez, R. (1996). Practices, beliefs, and cultures of high school mathematics departments: Understanding their influences on student advancement. *Journal of Curriculum Studies, 28,* 495-466.

Gutierrez, R. (2002). Beyond essentialism: The complexity of language in teaching mathematics to Latina/o students. *American Educational Research Journal,* 39(4), 1047-1088.

Gutstein, E. (2003). Teaching and learning mathematics for social justice in an urban, Latino school. *Journal for Research in Mathematics Education, 34,* 37-73.

Gutstein, E. (in press). "And That's Just How It Starts": Teaching Mathematics and Developing Student Agency. To appear in N. S. Nasir & P. Cobb (Eds.), *Diversity, equity, and access to mathematical ideas.* New York: Teachers College Press. Manuscript in preparation

Gutstein, E., Lipman, P., Hernández, P., & de los Reyes, R. (1997). Culturally relevant mathematics teaching in a Mexican American context. *Journal for Research in Mathematics Education, 28*, 709-737.

Heath, S. B. (1983). *Ways with Words : Language, Life and Work in Communities and Classrooms.* Cambridge: Cambridge University Press.

Hiebert. J., Gallimore, R., Garnier, H., Givvin, K., Hollingsworth, H., Jacobs, J., Chui, A., Wearne, D., Smith, M., Kersting, N., Manaster, A., Tseng, E., Etterbeek, W., Manaster, C., Gonzales, P., & Stigler, J. *Teaching Mathematics in Seven Countries: Results from the TIMSS 1999 Video Study.* Washington, DC: National Center for Education Statistics.

Hill, H. C., and Ball, D. L. (2004). Learning mathematics for teaching: Results from California's mathematics professional development institutes. *Journal for Research in Mathematics Education*, 35. (5). 330-351.

Hill, H. C., Rowan, B., & Ball, D. (in press). *Effects of teachers' mathematical knowledge for teaching on student achievement*. American Educational Research Journal.

Hillman, H. (1981). *Kitchen Science.* Boston: Houghton Mifflin

Horn, I. (2003). *Learning on the job: Mathematics teachers' professional development in the contexts of high school reform.* Doctoral Dissertation. Berkeley, CA: University of California, Berkeley.

Horn, I. (2005). Accountable argumentation as a participant structure to support learning through disagreement. Manuscript submitted for publication.

Kelley, D., Mullis, I., & Martin, M. (2000). *Profiles of Student Achievement in Mathematics at the TIMSS International Benchmarks: U.S. Performance and Standards in an International Context.* Boston: The International Association for the Evaluation of Educational Achievement, at Boston College.

Khisty, L. L. (1995). Making inequality: Issues of language and meanings in mathematics teaching with Hispanic students. In W. G. Secada, E. Fennema, & L. B. Adajian (Eds.), *New directions for equity in mathematics education* (pp. 279-297). New York: Cambridge University Press.

Kirshner, D., & Whitson, J. A. (Eds.). (1997). *Situated Cognition*. Mahwah, NJ: Erlbaum.

Klein, D. (2003). A brief history of American K-12 mathematics education in the 20th century. Retrieved on July 1, 2003 from: <http://www.csun.edu/~vcmth00m/AHistory.html>

Kozol, J. (1992). *Savage inequalities.* New York: Harper Perennial.

Ladson-Billings, G. (1994). *The dreamkeepers*. San Francisco: Jossey-Bass.

Ladson-Billings, G. (1995). But that's just good teaching! The case for culturally relevant pedagogy. *Theory into Practice*, 34(3), 159-165.

Lampert, M. (2001). *Teaching Problems and the Problem of Teaching*. New Haven: Yale University Press.

Lee, C. D. (1995). Signifying as a Scaffold for Literary Interpretation. *Journal of Black Psychology, 21*(4), 357-381.

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational researcher*, 31(1), 3-12.

Ma, L. (1999). *Knowing and teaching elementary mathematics.* Mahwah, NJ: Erlbaum.

Martin, D. (2000). Mathematics success and failure among African-American youth: The roles of sociohistorical context, community forces, school influence, and individual agency. Mahwah, NJ: Lawrence Erlbaum Associates.

Martin, D. (in press). Mathematics Learning and Participation in African American Context: The Co-Construction of Identity in Two Intersecting Realms of Experience. To appear in N. S. Nasir & P. Cobb (Eds.), *Diversity, equity, and access to mathematical ideas.* New York: Teachers College Press.

McGee, H. (2004). *On Food and Cooking: The Science and Lore of the Kitchen.* New York: Scribner.

Mehan, H. (1979). *Learning lessons: Social organization in the classroom.* Cambridge, MA: Harvard University Press.

Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory Into Practice*, 31(2), 132-141.

Moll, L., & González, N. (2004). Engaging life: A funds-of-knowledge approach to multicultural education. In James Banks & Cherry Banks (Eds.), *Handbook of research on multicultural education,* pp. 699-715. San Francisco: Jossey-Bass.

Moschkovich, J. N. (1999). Supporting the participation of English language learners in mathematical discussions. *For the Learning of Mathematics*, 19(1), 11-19.

Moschkovich, J. N. (2002). "A Situated and Sociocultural Perspective on Bilingual Mathematics Learners." *Mathematical Thinking and Learning*, 4(2&3), 189-212.

Moschkovich, J. N. (in press). Bilingual Mathematics Learners: How Views of Language, Bilingual Learners, and Mathematical Communication Impact Instruction. To appear in N. S. Nasir & P. Cobb (Eds.), *Diversity, equity, and access to mathematical ideas.* New York: Teachers College Press.

Moses, R. P., & Cobb, C. E. (2001). *Radical equations: Math literacy and civil rights*. Boston MA: Beacon Press.

Mullis, I. Martin, M., Beaton, A., Gonzalez E., ,Kelly, D., & Smith, T. (1998). *Mathematics and Science Achievement in the Final Year of Secondary School.* Boston: The International Association for the Evaluation of Educational Achievement, at Boston College.

Mullis, I., Martin, M., Gonzalez, E., Gregory, K., Garden, R., O'Connor, K., Chrostowski, S., & Smith, T. (2000). TIMSS 1999. *Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade. International Mathematics Report. Boston:* The International Association for the Evaluation of Educational Achievement, at Boston College.

Nasir, N. S. (2002). Identity, Goals, and Learning: Mathematics in Cultural Practice. *Mathematical Thinking and Learning*, 4(2&3), 213-248.

Nasir N. S., & Saxe, G. (2003). Ethnic and Academic Identities: A Cultural Practice Perspective on Emerging Tensions and Their Management in the Lives of Minority Students. *Educational Researcher,* 32(5), 14-18.

National Academy of Education, Committee on Teacher Education. (2005). *Preparing teachers for a changing world.* San Francisco, CA: Jossey-Bass.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Washington, DC: U.S. Government printing office.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

National Research Council. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

National Research Council. (1989). *Everybody counts: A report to the nation on the future of mathematics education*. Washington, DC: National Academy Press.

National Research Council. (2002a). *How people learn (Expanded edition).* Washington, DC: National Academy Press.

National Research Council. (2002b). *Scientific research in education.* Washington, DC: National Academy Press.

National Research Council. (2005). *On evaluating curricular effectiveness: judging the quality of K-12 mathematics evaluations.* Washington, DC: National Academy Press.

National Science Foundation. (2000). *Science and engineering indicators*. Washington, DC: National Science Foundation.

O'Connor, M.C.: 1998, Language socialization in the mathematics classroom: Discourse practices and mathematical thinking. In M. Lampert & M. Blunk (Eds.), *Talking mathematics: Studies of teaching and learning in school* (pp. 17-55). NY: Cambridge University Press.

O'Connor. M. C., & Michaels, S. (1993). Aligning academic task and participation status through revoicing: Analysis of a classroom discourse strategy. *Anthropology and Education Quarterly, 24,* 318-335.

O'Connor. M. C., & Michaels, S. (1996). Shifting participant frameworks: Orchestrating thinking practices in group discussion. In D. Hicks (Ed.), *Discourse, learning, and schooling* (pp. 63-103). New York: Cambridge University Press.

Pólya, G. (1945). *How to solve it.* Princeton, NJ: Princeton University Press.

Putnam, R. (2003). Commentary on Four elementary mathematics curricula. In S. Senk & D. Thompson (Eds.), *Standards-oriented school mathematics curricula: What does the research say about student outcomes?* (pp. 161-178). Mahwah, NJ: Erlbaum.

RAND mathematics study panel. (2002). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education.* Santa Monica, CA: RAND Foundation.

Ridgway, J., Crust, R., Burkhardt, H., Wilcox, S., Fisher, L., and Foster, D. (2000). MARS Report on the 2000 Tests. Mathematics Assessment Collaborative, San Jose, CA. pp 120.

Reys, R. E., Reys, B. J., Lappan, R., Holliday, G., & Wasman, D. (2003). Assessing the impact of standards-based middle grades mathematics curriculum materials on student achievement. *Journal for Research in Mathematics Education, 34*(1), 74-95.

Ridgway, J, Zawojewski, J., Hoover, M., and Lambdin, D. (2003). Student Attainment in the Connected Mathematics Curriculum. In Sharon Senk & Denisse Thompson (Eds.), *Standards-oriented school mathematics curricula: What does the research say about student outcomes?* (pp. 193-224). Mahwah, NJ: Erlbaum.

Riordan, J., & Noyce, P. (2001). The impact of two standards-based mathematics curricula on student achievement in Massachusetts. *Journal for Research in Mathematics Education*, 32(4), 368-398.

Rosebery, A., Warren, B., & Conant, F. (1992). Appropriating scientific discourse: Findings from language minority classrooms. *Journal of the Learning Sciences* 2: 61-94.

Rosebery, A., Warren, B., Ogonowski, M. & Ballenger, C. (2005). The generative potential of students' everyday knowledge in learning science. In T. Carpenter and T. Romberg (Eds.), *Understanding matters: Improving student learning in mathematics and science* (pp. 55-80). Mahwah, NJ: Erlbaum.

Rosen, L. (2000) Calculating concerns: The politics or representation in California's "Math Wars." Unpublished doctoral dissertation. University of California, San Diego.

Saxe, G. (1999). Cognition, development, and cultural practices. In E. Turiel (Ed.), *Development and cultural change: Reciprocal processes* (pp. 19-35). San Francisco: Jossey-Bass.

Schifter, D. (1993). Mathematics process as mathematics content: A course for teachers. *The Journal of Mathematical Behavior, 12(3),* 271-283.

Schifter, D. (1998). Learning mathematics for teaching: From the teachers' seminar to the classroom. *Journal for Mathematics Teacher Education, 1,* 55-87.

Schifter, D. (2001). Learning to See the Invisible: What skills and knowledge are needed to engage with students' mathematical ideas? In T. Wood, B. S. Nelson, & J. Warfield (Eds.), *Beyond classical pedagogy: Teaching elementary school mathematics* (pp. 109–134). Hillsdale, NJ: Erlbaum.

Schifter, D., Bastable, V., & Russell, S.J. (1999) *Developing mathematical ideas.* Parsippany, NJ: Dale Seymour Publications.

Schifter, D. & Fosnot, C.T. (1993). *Reconstructing mathematics education: Stories of teachers meeting the challenge of reform.* New York: Teachers College Press.

Schifter, D., Russell, S.J., & Bastable, V. (1999). *Teaching to the big ideas. In M. Solomon (Ed.), The diagnostic teacher* (pp. 22-47). New York: Teachers College Press.

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

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334-370). New York: MacMillan.

Schoenfeld, A. H. (1998a). Toward a theory of teaching-in-context. *Issues in Education*, Volume 4, Number 1, pp. 1-94.

Schoenfeld, A. H. (1998b). Making mathematics and making pasta: From cookbook procedures to really cooking. In J. G. Greeno & S. V. Goldman (Eds.), *Thinking practices in mathematics and science learning* (pp. 299-319). Mahwah, NJ: Erlbaum.

Schoenfeld, A. H. (1999) (Special Issue Editor). *Examining the Complexity of Teaching.* Special issue of the *Journal of Mathematical Behavior*, 18 (3).

Schoenfeld, A. H. (2000). Models of the teaching process. *Journal of Mathematical Behavior*, 18 (3), 243-261.

Schoenfeld, A. H. (2002a) A highly interactive discourse structure. In J. Brophy (Ed.), *Social Constructivist Teaching: Its Affordances and Constraints* (Volume 9 of the series *Advances in Research on Teaching*) (pp. 131-170). New York: Elsevier.

Schoenfeld, A. H. (2002b). Making mathematics work for all children: Issues of standards, testing, and equity." *Educational researcher*, 31(1), 13-25.

Schoenfeld, A. H., Minstrell, J., and van Zee, E. (2000).The detailed analysis of an established teacher carrying out a non-traditional lesson. *Journal of Mathematical Behavior*, 18 (3), 281-325.

Schoenfeld, A. H. (2004). The math wars. *Educational Policy, 18(1)*, 253-286.

Schoenfeld, A. H. (2005). Dilemmas/Decisions: Can we model teachers' on-line decision-making? Manuscript submitted for publication.

Schoenfeld, A. H. (in press.) Design Experiments. In P. B. Elmore, G. Camilli, & J. Green (Eds.), *Complementary Methods for Research in Education.* Washington, DC: American Educational Research Association

Schoenfeld, A. H. (in preparation). Educational Research and Practice. Manuscript prepared for the Macarthur Research Network on Teaching and Learning.

Schoenfeld, A. H., Gamoran, M., Kessel, C., Leonard, M., Orbach, R., & Arcavi, A. (1992). Toward a comprehensive model of human tutoring in complex subject matter domains. *Journal of Mathematical behavior,* 11(4), 293-320.

Schoenfeld, A. H., Minstrell, Jim, and van Zee, Emily. (2000).The detailed analysis of an established teacher carrying out a non-traditional lesson. *Journal of Mathematical Behavior*, 18 (3), 281-325.

Senk, S., & Thompson, D. (Eds.). (2003). *Standards-oriented school mathematics curricula: What does the research say about student outcomes?* Mahwah, NJ: Erlbaum.

Sfard, A. (2000). Symbolizing mathematical reality into being: How mathematical discourse and mathematical objects create each other. In P. Cobb, E. Yackel, & K. McClain (Eds), *Symbolizing and communicating: Perspectives on mathematical discourse, tools, and instructional design* (pp. 37-98). Mahwah, NJ: Erlbaum.

Sfard, A. (2001e). There is More to Discourse than Meets the Ears: Learning from mathematical communication things that we have not known before. *Educational Studies in Mathematics*, 46(1/3), 13-57.

Sfard, A., & Kieran, C. (2001). Cognition as Communication: Rethinking Learning-by-Talking Through Multi-Faceted Analysis of Students' Mathematical Interactions. *Mind, Culture, and Activity*, 8(1), 42-76.

Sfard, A., & McLain, K. (2003). *Analyzing tools: Perspectives on the role of designed artifacts in mathematics learning.* Special issue of the *Journal of the learning Sciences, 11*(2&3).

Sherin, M. G. (2001). Developing a professional vision of classroom events. In T. Wood, B. S. Nelson, & J. Warfield (Eds.), *Beyond classical pedagogy: Teaching elementary school mathematics* (pp. 75–93). Hillsdale, NJ: Erlbaum.

Sherin, M. G. (2002). When teaching becomes learning. *Cognition and Instruction*, 20(2), 119–150.

Sherin, M. G. (2004). New perspectives on the role of video in teacher education. In J. Brophy (Ed.), *Using video in teacher education* (pp. 1–27). NY: Elsevier Science.

Sherin, M. G., & Han, S. Y. (2004). Teacher learning in the context of a video club. *Teaching and Teacher Education 20*, 163–183.

Shulman, Lee S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 17(1), 4-14.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57, 1-22.

Stigler, J.W., Gonzales, P., Kawanaka, T., Knoll, S., and Serrano, A. (1999). *The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States. (NCES 1999-074).* U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technical innovation*. Washington, DC: Brookings.

Study of Instructional Improvement. (2002) *Measuring teachers' content knowledge for teaching: Elementary mathematics release items*. Ann Arbor: University of Michigan. Downloaded February 6, 2005, at <http://www.sii.soe.umich.edu/instruments.html>

Swafford, Jane. (2003). Reaction to high school curriculum projects' research. In Sharon Senk & Denisse Thompson (Eds.), *Standards-oriented school mathematics curricula: What does the research say about student outcomes?* (pp. 457-468). Mahwah, NJ: Erlbaum.

Tate, William. (1997). Race-ethnicity, SES, Gender, and language proficiency trends in mathematics achievement: an update. *Journal for Research in Mathematics Education, 28(6)*, 652-679.

Thorndike, E. L. (1931). *Human Learning*. New York: Century

U. S. House Committee on Science. (1998). *Unlocking our future: Toward a new national science policy. A report to Congress by the House Committee on Science*. Washington, DC: Author. See also <http://www.house.gov/science/science_policy_report.htm>.

Viadero, Debra. (2004, August 11). Researchers question clearinghouse choices. *Education Week,* pp. 30–32.

Warren, B., Ballenger, C., Ogonowski, M., Rosebery, A., & Hudicourt-Barnes, J. (2001). Rethinking diversity in learning science: The logic of everyday sense-making. *Journal of Research in Science Teaching* 38: 1-24.

Warren, B. & Rosebery, A. (1995). Equity in the future tense: Redefining relationships among teachers, students, and science in linguistic minority classrooms. In W. Secada, E. Fennema and L. Adajian (Eds.), *New directions for equity in mathematics education* (pp. 298-328). NY: Cambridge University Press.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity.* Cambridge: Cambridge University Press.

What Works Clearinghouse (2004a). Curriculum-based interventions for improving K-12 mathematics achievement – middle school. Downloaded from <http://www.whatworks.ed.gov/>, February 27, 2005.

What Works Clearinghouse (2004b). Detailed study report: Baker, J. J. (1997). Effects of a generative instructional design strategy on learning mathematics and on attitudes towards achievement. Unpublished doctoral dissertation, University of Minnesota. Downloaded from <http://www.whatworks.ed.gov/>, February 27, 2005.

What Works Clearinghouse (2004c). Detailed study report: Kerstyn, C. (2001). Evaluation of the I CAN LEARN Mathematics Classroom. First year of implementation (2000–2001 school year). Unpublished manuscript. Downloaded from <http://www.whatworks.ed.gov/>, February 27, 2005.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education,* 27(4), 458-477.