Making Mathematics Work for All Children: Issues of Standards, Testing, and Equity

by Alan H. Schoenfeld

"Mathematics Education is a civil rights issue," says civil rights leader Robert Moses, who argues that children who are not quantitatively literate may be doomed to second-class economic status in our increasingly technological society. The data have been clear for decades: poor children and children of color are consistently shortchanged when it comes to mathematics. More broadly, the type of mathematical sophistication championed in recent reform documents, such as the National Council of Teachers of Mathematics' (2000) *Principles and Standards for School Mathematics,* can be seen as a core component of intelligent decision making in everyday life, in the workplace, and in our democratic society. To fail children in mathematics, or to let mathematics fail them, is to close off an important means of access to society's resources. This article discusses the potential for providing high quality mathematics instruction for all students. It addresses four conditions necessary for achieving this goal: high quality curriculum; a stable, knowledgeable, and professional teaching community; high quality assessment that is aligned with curricular goals; and stability and mechanisms for the evolution of curricula, assessment, and professional development. The goal of this article is to catalyze conversations about how to achieve sustained, beneficial changes.

Robert Moses (2001) argues,

> 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. (p. 5)

This is a powerful statement. In the simplest terms, Moses' argument is about economics and technology. Factory jobs, once "pure" physical labor, now have technological components; even forklifts in warehouses have computer modules, and the people who use the machines must be able to use the computer controls. Most important, the technological divide is going to widen over the coming years. Moses (2001) argues that those who are technologically literate will have access to jobs and economic enfranchisement, while those without such skills will not:

Educational Researcher, Vol. 31. No. 1, pp. 13–25

> Sixty percent of new jobs will require skills possessed by only 22 percent of the young people entering the job market now. These jobs require the use of a computer and pay about 15 percent more than jobs that do not. And those jobs that do not are dwindling. Right now, the Department of Labor says, 70 percent of all jobs require technology literacy; by the year 2010 *all* jobs will require significant technical skills. And if that seems unimaginable, consider this: the Department of Labor says that 80 percent of those future jobs *do not yet exist.* (pp. 8–9)

There are counterarguments to this position, perhaps the most common one being that technology is consistently being adapted to provide jobs for those who have minimal mathematical skills. Years ago, for example, McDonald's "solved" the problem of cashiers who had difficulty making change by introducing cash registers that had pictures of the items being sold on its buttons and that calculated change automatically. But this line of argument winds up in much the same place as Moses': those students with negligible skills wind up having access to the lowest paying jobs. In purely functional terms, mathematics has long been recognized as a *critical filter* (Sells, 1975, 1978). Course work in mathematics has traditionally been a gateway to technological literacy and to higher education. On such grounds alone, one could argue that there is a national obligation to insure that all students have access to high quality mathematics instruction. The argument for economic enfranchisement is, however, only first among equals. As will be elaborated later in this article, making decisions in one's personal life, on the job, and in matters of public interest calls increasingly for quantitatively sophisticated reasoning. More than ever before, today's students need to learn to reason and communicate using mathematical ideas.

Mathematical literacy should be a goal for all students—so what makes it a civil rights issue? The answer becomes clear when one looks at the numbers. Disproportionate numbers of poor, African-American, Latino, and Native American students drop out of mathematics and perform below standard on tests of mathematical competency, and are thus denied both important skills and a particularly important pathway to economic and other enfranchisement (Madison & Hart, 1990; Miller, 1995; National Action Committee for Minorities in Engineering, 1997; National Commission on Mathematics and Science Teaching for the 21st Century, 2000; National Science Foundation [NSF], 2000). Hence conversations about the mathematical needs of American students must focus not only on what mathematics the students should learn, but also on how we as a nation can insure that all students have the opportunity to learn it.

This article addresses those issues. To set the stage it begins with a description of mathematics instruction over the second half of the 20th century, describing the nature of the mathematics studied and increasing perceptions through the 1980s of difficulties with the curriculum. There was unhappiness with the curriculum on a number of grounds: equity being first, a narrow content focus aimed at preparing students for college being second, and national security being third. Such discontent led to calls for reform, with the National Research Council (NRC) issuing *Everybody Counts* and the National Council of Teachers of Mathematics (NCTM) issuing *Curriculum and Evaluation Standards for School Mathematics* (the *Standards*) in 1989. On the basis of a decade's experience and research, NCTM issued a new vision statement, *Principles and Standards for School Mathematics* (*Principles and Standards*) in 2000.

In the years since 1989 there have been significant changes in the curriculum. Many of these changes have been controversial: Opponents of reform feared that an emphasis on process over content would result in weakening the curriculum and decrease American children's mathematical competencies. Until now, there has not been the opportunity to evaluate the potential effectiveness of reform efforts. Large-scale change is slow. It took some years after the 1989 *Standards* was issued before curricula aligned with its reform goals could be developed and implemented; it took more time still before data on their effectiveness would become available. Such data are now beginning to come in, allowing one to see if the new curricula provide a basis for achieving the content and equity goals of reform—enfranchising all students and having them learn more powerful and useful mathematics. This article presents some preliminary data indicating that there are grounds for optimism. In research on some of the first large-scale implementations of reform curricula, data indicate that reform students do as well on skills as students who study the traditional curricula, and that they do better on an understanding of concepts and problem solving. Moreover, traditional performance gaps between majority students and poor or underrepresented minorities are diminished, though not eliminated.

Those data indicate that the first few steps of reform seem to be going in the right directions. Given that, what kind of infrastructure is required to stabilize and build on this progress? Sustained, incremental progress calls for the availability of high quality curricula; for a stable, knowledgeable, and professional teaching community; for high quality assessment that is aligned with curricular goals; and for stability and mechanisms for the evolution of curricula, assessment, and professional development. This article briefly assesses the current state of each of these and the possibilities for improvement.

Status and Visions of Mathematics Instruction in the Late 20th Century

Cases can be made for the virtues and drawbacks of the traditional U.S. mathematics curriculum. On the plus side, the United States' stature as a world superpower has been attributed in part to our education system. Education has been viewed in our society as a democratic vehicle for advancement. Millions of citizens, from the children of immigrants to the soldiers who took advantage of the GI Bill after World War II, would give testimony to the power of education as a mechanism for social and economic advancement. More specifically, mathematics education has been viewed as providing some of the underpinnings of the nation's technological and scientific prowess. The K–12 mathematics pipeline produced large enough numbers of people with strong enough mathematical backgrounds to serve as the backbone of the nation's mathematical and scientific infrastructure.

There has also, however, been a significant downside. First, the adequacy of the mathematical content that American students learned has periodically come into question—typically at points of national crisis. The new math came about, for example, in the aftermath of Russia's successful launch of Sputnik in 1957. Another impetus for change—economic rather than military—came with the waxing and waning of the Japanese and American economies, respectively, in the 1980s (e.g., National Commission on Excellence in Education, 1983). A sense of economic jeopardy, combined with American children's poor showing on the Second International Mathematics and Science Study (McKight et al., 1987), led to concerns about the adequacy of the traditional curriculum.

While a key issue leading to reform was the sense that American mathematics curricula were not internationally competitive, there were other reasons. There was a serious concern that the high school mathematics curriculum was designed for those who planned to enter college, focusing largely on the skills that would ultimately enable students to study calculus. In K–12 mathematics there was little emphasis on the kinds of mathematics that would enable students to make sense of the world around them—neither statistics nor mathematical modeling was part of the traditional curriculum. There was also little or no emphasis on communicating and using mathematical ideas. Typically all one needed to earn full credit on a mathematics problem was to perform some computations, arrive at the correct answer, and write the answer in a box.

If a large proportion of K–12 students had been successful in the traditional curriculum, the impetus for change might have been muted. But that was not the case. Large numbers of students failed or left mathematics, and a disproportionate number of those who left were students of color. A 1989 report from the NRC, *Everybody Counts,* made the case as follows:

> More than any other subject, mathematics filters students out of programs leading to scientific and professional careers. From high school through graduate school, the half-life of students in the mathematics pipeline is about one year; on average, we lose half the students from mathematics each year, although various requirements hold some students in class temporarily for an extra term or a year. Mathematics is the worst curricular villain in driving students to failure in school. When mathematics acts as a filter, it not only filters students out of careers, but frequently out of school itself. (p. 7)

These negative effects were not distributed equally. *Everybody Counts* (NRC, 1989) argued the importance of mathematics education as a civil rights issue:

> Low expectations and limited opportunity to learn have helped drive dropout rates among Blacks and Hispanics *much* higher—unacceptably high for a society committed to equality of opportunity. It is vitally important for society that *all* citizens benefit equally from high quality mathematics education. (p. 7)

Myriad data document disproportionate dropout and low performance rates for students of color (e.g., Madison & Hart, 1990;

Miller, 1995; National Action Committee for Minorities in Engineering, 1997; National Commission on Mathematics and Science Teaching for the 21st Century, 2000; NSF, 2000). As one example, consider the following breakdown by race of mathematics scores on the National Assessment of Educational Progress (NAEP).[1] The critical scores are at age 17, when students are about to graduate. In recent years more than two thirds of the White 17-year-olds sampled by NAEP performed at benchmark levels—that is, were deemed to know the appropriate level of mathematics. Only about 40% of the Latino 17-year-olds, and less than one third of the African-American 17-year-olds, met benchmark performance levels. A comparison of scores for 9-, 13- and 17-year-olds shows that the gap in scores between Whites and underrepresented minorities increases as students get older.

Data from the Third International Mathematics and Science Study (TIMSS) indicate that socioeconomic status (SES) correlates with performance: "The data . . . show a relationship between the relative wealth of districts and student achievement in the subject areas. . . . Those that 'have' get more" (Schmidt, 2001). Mathematics scores on the SAT also correlate astoundingly well with parental income (National Action Committee for Minorities in Engineering, 1997). And one hardly needs to provide evidence of the correlations between race and SES or race and opportunity to learn. Powerful arguments and data are given in Miller (1995). (For compelling descriptions of the reality behind the figures, see Kozol, 1992.)

Recognition of the content and equity issues highlighted above contributed to calls for reform. The NRC issued *Everybody Counts* in early 1989 and NCTM issued the *Standards* later that year. Rather than focusing simply on mathematical content, the *Standards* addressed "what it means to be mathematically literate in a world that relies on calculators and computers to carry out mathematical procedures and in a world where mathematics is rapidly growing and is extensively being applied in diverse fields" (NCTM, 1989, p. 1). In addition to identifying the content students should know, the *Standards* focused on *process*: there was to be a focus at all grade levels on problem solving, reasoning, connections (between mathematical topics and to real world applications), and the communication of mathematical ideas in written and oral form. Indeed, the very goals of mathematics education were reconceptualized. After characterizing the then-current structures and content of schooling as a product of the industrial age, the *Standards* went on to say the following:

> The education system of the new industrial age does not meet the economic needs of today. New societal goals for education include (1) mathematically literate workers, (2) lifelong learning, (3) opportunity for all, and (4) an informed electorate. (p. 3)

To the surprise of its authors and just about everyone else, the *Standards* and two subsequent volumes that focused on teaching and assessment (NCTM, 1991, 1995) catalyzed a national standards movement.

Like Horace Greeley's famous injunction "Go west, young man," the *Standards* was long on direction and short on detail—it was a vision statement rather than a blueprint. In hindsight, that was a good thing, for both political and intellectual reasons. Given the American context of curricular pluralism, no set of specific curricular recommendations would have been politically viable. More important, however, is that the publication of the *Standards* in 1989 set the mathematical community in motion. Over the following decade, a number of curricula that took very different approaches to achieving the goals of the *Standards* were developed and refined. The result has been a collection of diverse models of curricular change, and significantly enriched curricular discourse.

Similarly, NCTM's (2000) *Principles and Standards* is a vision statement for mathematics education designed to reflect a decade's experience since the publication of the original *Standards* volumes. In terms of context, there are new curricular possibilities. In 1989 a call for all students to use scientific calculators was revolutionary; today web access makes huge databases available, and students have easy access to statistical packages to examine such data. As discussed below, the original *Standards* documents were inspired by research but not grounded in large-scale empirical testing (materials for such testing did not exist in 1989). A decade later, there is a much more solid research base. There is thus strong continuity between *Principles and Standards* and its antecedents, but also some bold new perspectives enabled by a decade's experience and reflection.

The new document rests on a set of principles that make the social and intellectual commitments of its authors clear: equity (high expectations and strong support for all students); coherent curricula rather than disconnected sets of activities; teacher professionalism, including knowledge of curricula and learning; and the effective use of assessment and technology in the service of mathematics learning. The document refines the curricular focus of earlier documents, focusing on five content standards (number, algebra, geometry, measurement, data analysis, and probability) interwoven with five process standards (problem solving, reasoning and proof, connections, communication, and representation). In a deliberate break with the traditional assumption that only some students will enroll in mathematics courses that prepare them for attending college, *Principles and Standards* calls for the development of a core curriculum that prepares all students with the mathematical background for quantitative literacy, for the workplace, and for study at the college level.

Like its antecedent, *Principles and Standards* can (despite its nearly 400 pages of densely packed text) be accused of being long on vision and somewhat short on detail. It identifies some essential goals, but does not provide a blueprint for achieving them. Its authors (among whom is the author of this article) hope that, like its antecedent, the vision will come to life over the decade following its publication as the mathematical community labors to make it a reality.

Can it Work? Hopeful Signs

The 1989 *Standards* were grounded in contemporary research on mathematical thinking and problem solving. At the same time, however, the recommendations found therein were speculative. In 1989 no curricula designed to meet the goals delineated in the *Standards* were widely available, so there had been no opportunity to test the large-scale implementation of its ideas. Thus, opponents of *reform* (as the *Standards*-based movement came to be known) could and did complain that an untested approach was being forced upon American school children. They raised concerns that the reform approach would cause a precipitous decline in students' knowledge of mathematical facts and procedures. For reasons of space, the curricula and the controversies surrounding them will not be reviewed here.[2]

The large-scale implementation and evaluation of curricula embodying the ideas in the *Standards* took some years—curricula needed to be constructed, published, adopted, and assessed. The National Science Foundation (NSF) supported the development of a number of reform curricula in the early 1990s. By the mid-1990s, some of these were available for adoption. Insofar as these were 3-, 4-, or 5- year curricular packages (typically covering one of the elementary, middle, or high school grade bands), it was not until the turn of the century that significant numbers of students had made their way through a full reform curriculum package. Now, more than a decade after the publication of the *Standards,* hard data on large-scale implementations of these curricula are beginning to come in. To briefly summarize the current state, a converging body of data indicates the following:

1. On tests of basic skills, there are no significant performance differences between students who learn from traditional or reform curricula.
2. On tests of conceptual understanding and problem solving, students who learn from reform curricula consistently outperform students who learn from traditional curricula by a wide margin.
3. There is some encouraging evidence that reform curricula can narrow the performance gap between Whites and underrepresented minorities.

The remainder of this section reviews such data, focusing largely on data from Pittsburgh, Pennsylvania. The reason for devoting so much space to that review is that Pittsburgh offers an early and unusually well-documented set of results concerning the large-scale implementation of reform—including extensive data regarding the impact of reform curricula on the racial performance gap discussed above.

Pittsburgh serves about 40,000 students in 97 public schools (59 elementary, 19 middle, 11 high, and eight other). Roughly 56% of the student population is African American and 44% is White/Other; more than 60% of the students qualify for free or reduced-price lunches.

What makes the Pittsburgh Public Schools unusual is that Pittsburgh has, since the early 1990s, made a coherent systemic effort to implement standards-based education in mathematics and other subject areas. Efforts have included the delineation of content and performance standards in line with the *Standards* and the 1996 New Standards mathematics exam (New Standards, 1996); the use of standards-based assessments (specifically, the New Standards Reference Examinations, Harcourt Educational Measurement, 1996–1999) for purposes of aligning curriculum and assessment; the use of traditional assessments (specifically, the Iowa Test of Basic Skills) for comparison purposes; and standards-based instructional materials and professional development (Briars, 2001; Briars & Resnick, 2000).

The overall results of Pittsburgh's systemic efforts are seen in Figures 1 and 2. Note that the 1996 and 1997 cohorts had studied the traditional curriculum; the 1998–2000 cohorts had studied one of the major reform curricula.

As Figure 1 indicates, scores on concepts and problem solving increased with the implementation of the new curriculum, and continued to rise as the teachers became increasingly familiar with it. In 1997 roughly 10% of the traditional students met or exceeded the standards for concepts or problem solving; in 2000 roughly 25% of Pittsburgh's (now reform) students met or exceeded those standards. Although there is still huge room for improvement, the fact that two and a half times as many students than before met or exceeded those standards is impressive. Of course, problem solving was not a focus of the traditional curriculum. It is in the area of skills that the most surprising data emerge. Traditional measures of skills such as the Iowa Test of Basic Skills (ITBS) show that the reform curricula more than hold their own against traditional curricula with regard to skills. Even in the first year of new curriculum implementation, 1998, greater percentages of students scored above the 50th and 70th percentiles on the ITBS than in previous years, and fewer students scored below the 25th percentile. Skill scores on the New Standards examinations are more dramatic. Under the traditional curriculum, less than a third of the students met or exceeded the skill standard on the New Standards reference exam. With the new curriculum significantly more than 50% (and in 2000, 60%) do. That is, with the switch from traditional to reform curriculum the proportion of students performing well in terms of skills doubled from 1997 to 2000. In short, the fears of anti-reform groups that reform curricula would cause a decrease in student skill levels appear to be unwarranted.

Data such as these indicate that coherent approaches to teaching mathematics for conceptual understanding produce significant improvements across the board—not only in concepts and problem solving, but in skills as well. Moreover, Figure 2 shows that such curricula do not cream off the best students and help them do better: not only do many more students do well, as seen in Figure 1, but fewer students sink to the bottom of the barrel. There seems to be some generality to such findings.

As one would expect in any district that contains nearly 100 schools, some Pittsburgh teachers embraced the proposed reforms and some did not. This raises the question of whether there were performance differences between students who experienced the curriculum as implemented (more or less) faithfully, and those that did not. Observers identified "strong implementation" teachers as those in whose classrooms students were familiar with activities and procedures specific to the *Everyday Mathematics* curriculum; curricular artifacts such as visual aids and manipulative materials were accessible and showed clear signs of use; students had frequent opportunities to work together and explain their work to each other when appropriate; displays of student work showed curriculum-specific projects and activities; and there was no evidence of the use of other programs. "Weak implementers were either not using the curriculum at all, or using it so little that the overall instruction in the classroom was hardly distinguishable from traditional mathematics instruction" (Briars & Resnick, 2000, p. 6). Strong implementation schools were identified as those in which there was strong implementation by all Grade 3 and 4 teachers in the 1996–1997 and 1997–1998 years, respectively. Weak implementation schools were identified as those in which all but one or two teachers were identified as weak implementers. Briars and Resnick (2000) created a matched sample of

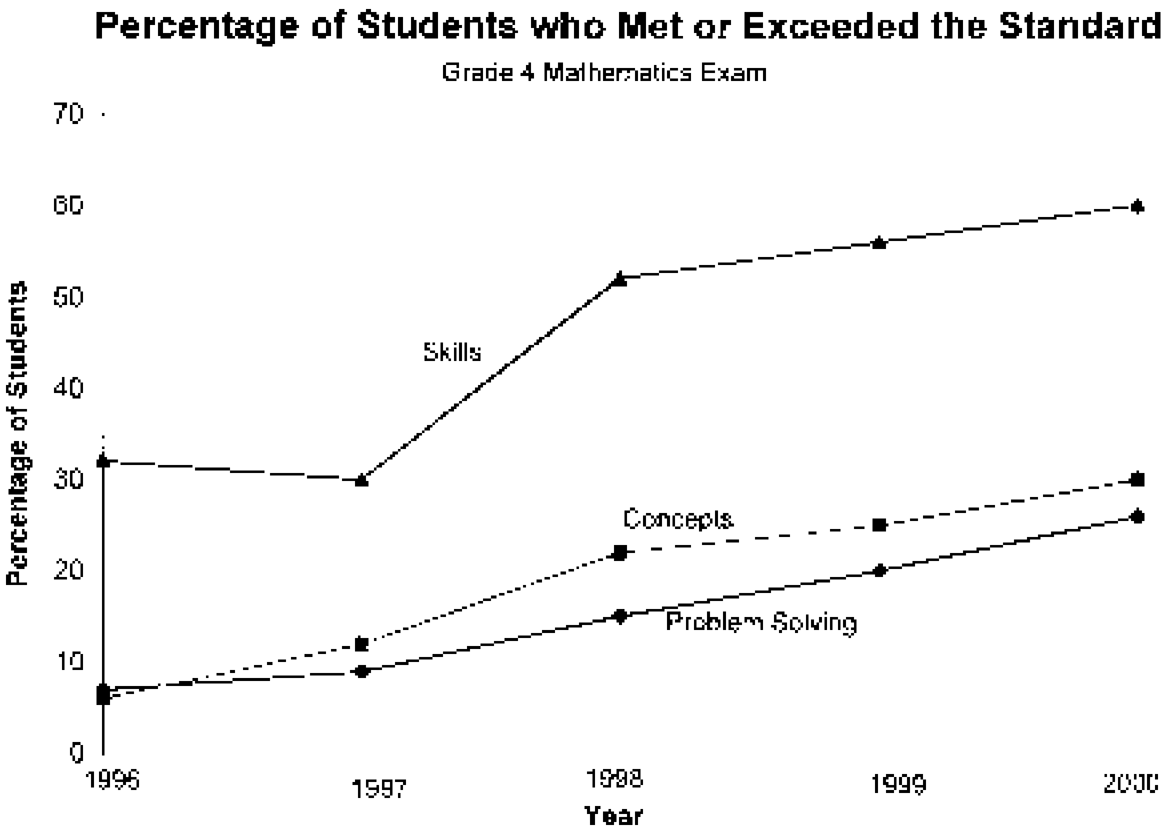

FIGURE 1. *(Adapted from Briars, 2001) Percentage of Grade 4 students meeting exceeding standards of New Standards Mathematics exam, 1996–2000, 1996, and 1997. Cohort studied traditional curriculum 1998–2000. Other cohorts studied Everyday Mathematics.*

strong and weak implementation schools on socioeconomic grounds. They also verified that students at matched schools had "virtually identical" ITBS scores in 1995, the year before implementation of the new curriculum. Figure 3 represents a composite of the 1998 fourth-grade comparative student data reported in Briars and Resnick (2000). The data show that students at the strong implementation schools outperformed students at the weak implementation schools by a very wide margin. In other words, the more consistently the reform curriculum was implemented, the better students did.

Perhaps most important with regard to equity, not only did average scores on skills, concepts, and problem solving go up, but racial differences in performance diminished substantially. Figure 3 shows the percentage of Pittsburgh's White and African-American fourth graders in weak implementation schools and in demographically similar strong implementation schools who met the standards in skills, problem solving, and concepts on the 1998 New Standards exam. Both White and African-American students at the strong implementation schools far outperformed their counterparts at the weak implementation schools. At the strong implementation schools the performance of White and African-American students on the skills component of the test was roughly comparable.

There were still significant disparities in scores on problem solving and concepts. In the weak implementation schools, a negligible fraction (approximately 4%) of African-American students met the standards in problem solving and concepts. White students did better—approximately 18% met each standard—but the percentage is still low, and the ratio of White students to African-American students who met the standard is more than four to one. In the strong implementation schools, the percentage of African-American students meeting each of the concepts and problem solving standards is 30% or more, a more than sevenfold increase over their counterparts at the weak implementation schools. Moreover, the ratio of White students to African-American students who met those standards dropped from more than four to one to about three to two. Although this falls far short of the goal of parity, it represents very significant progress toward it.

The bottom line is that standards-based reform appears to work when it is implemented as part of a coherent systemic effort in which curriculum, assessment, and professional development are aligned. Not only do many more students do well, but the racial performance gap diminishes substantially.

The data described represent one of the first evaluations of coherent systemic efforts at mathematics reform, but they represent the tip of an emerging iceberg. Not long after the Pittsburgh data became available, the Mathematics in Michigan Convocation released data regarding Michigan students' performance on an international benchmarking study (Mullis et al., 2001). The Michigan mathematics standards are closely aligned with the *Standards*, as are the three main curricula highlighted in the Mathematics in Michigan Convocation: *Everyday Mathematics* at the elementary level, *Connected Mathematics Project* at the middle school level,

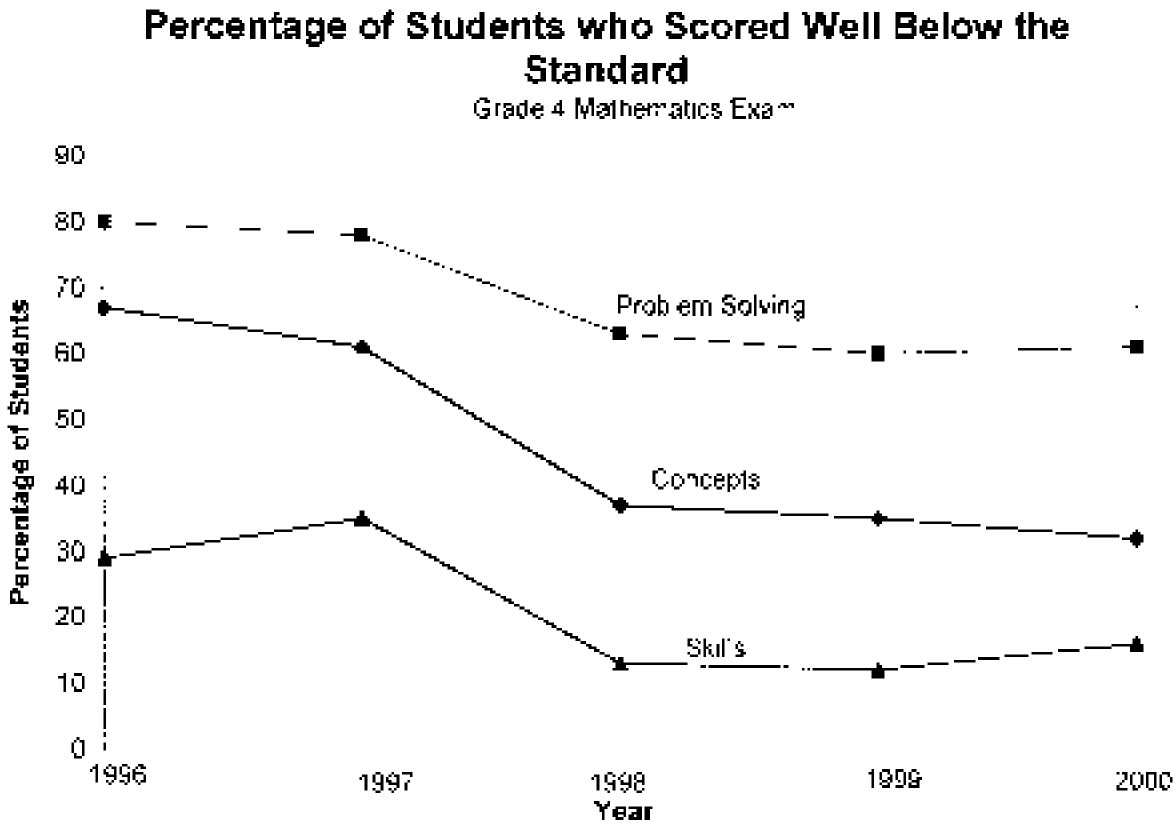

FIGURE 2. *(Adapted from Briars, 2001) Percentage of Grade 4 students who scored well below the standard of New Standards Mathematics exam, 1996–2000, 1996, and 1997. Cohort studied traditional curriculum 1998–2000. Other cohorts studied Everyday Mathematics.*

and *Contemporary Mathematics in Context (Core-Plus Mathematics)* at the high school level.

The *Michigan Government News* noted on April 4, 2001, "eighth graders in the state of Michigan scored best in math and science compared to other states that participated in the Third International Mathematics and Science Study–Repeat (TIMSS-R). An 'invitational' group made up of 21 Michigan schools using criteria that stress a curriculum that aligns to the state's high standards scored even better."

It should be stressed that the invitational group contained a substantial number of schools in low-SES districts. What distinguished them, according to the article, were the curricular resources available and the alignment of curricula, standards, and assessment.

Additional data regarding large-scale implementations of reform curricula are now becoming available. For example, in the July 2001 issue of the *Journal for Research in Mathematics Education,* Riordan and Noyce describe the results of a matched comparison of traditional and reform curricula in Massachusetts, with a statewide assessment as the measure of performance:

> Fourth-grade students using *Everyday Mathematics* and eighth-grade students using *Connected Mathematics* outperformed matched comparison groups who were using a range of textbooks commonly used in Massachusetts. The gain in student performance was greater in schools farther along in their implementation of the standards-based programs. These performance gains, which were moderate in size, remained consistent for different groups of students, across mathematical topics and different types of questions on the state test. This study supports the notion held by proponents of standards-based curriculum, that curriculum itself can make a significant contribution to improving student learning. (pp. 392–393)

Over the next few years, much more data regarding the implementation of reform curricula will become available. For example, the first volume of comprehensive evaluations of *Standards*-based instructional materials (Senk & Thompson, in press) is nearing publication. In a commentary following four chapters evaluating reform curricula at the elementary level, Ralph Putnam (in press) observes, "The first striking thing to note about the chapters is the overall similarity in their findings. 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." Commenting on the analyses of three middle school reform curricula, Chappell notes, "With respect to long-term benefits, the findings identified are convincing. 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" (in press). Reviewing performance data from five high school reform curricula, Swafford (in press) writes,

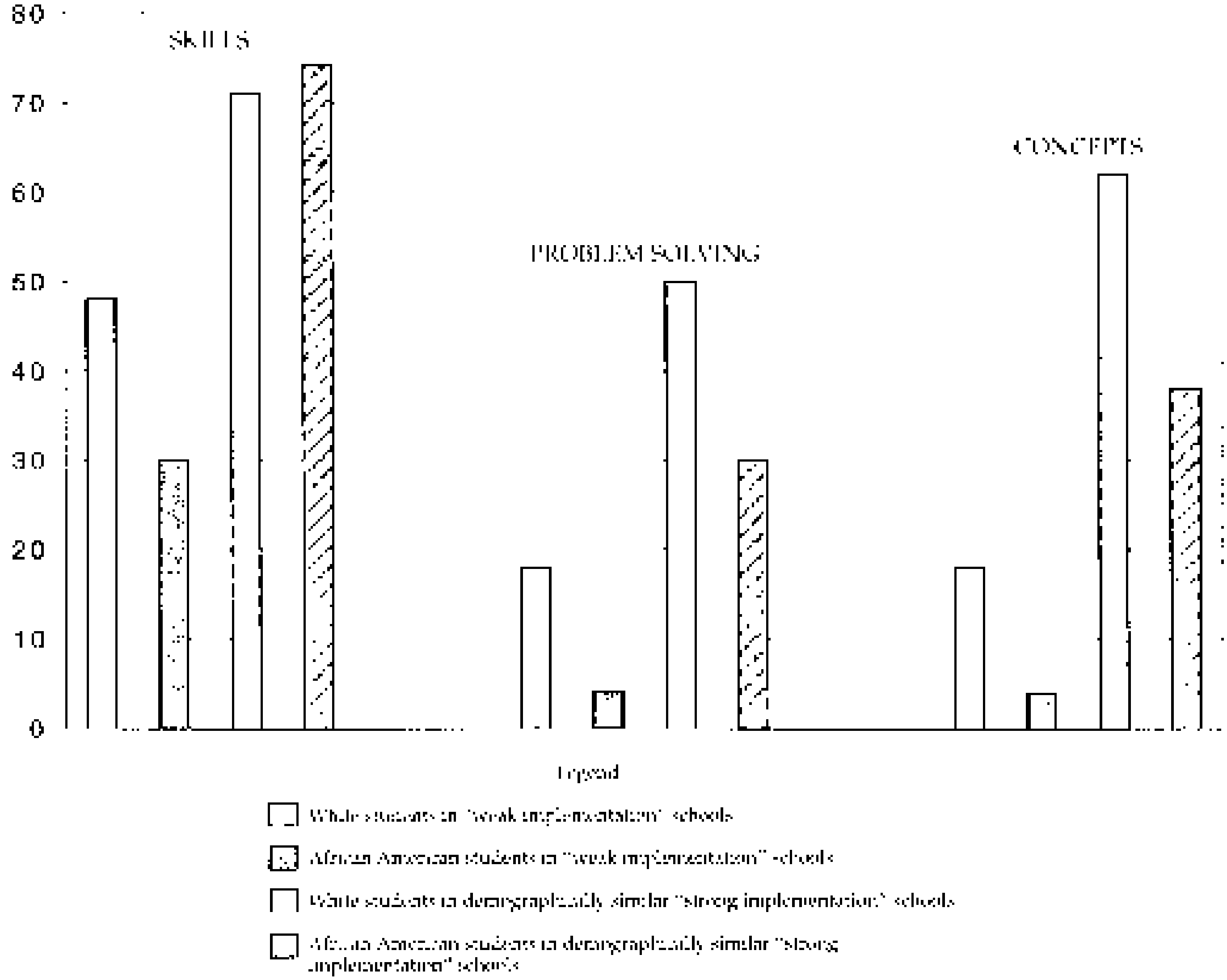

FIGURE 3. *(Derived from Briars and Resnick, 2000) Percentage of 4th-grade students in demographically matched "weak implementation" and "strong implementation" schools who achieved the skills, problem solving, and concepts standard in 1998.*

> What has been presented in this section is not only extensive research on five reform curricula, but also considerable evidence that the promises of reform mathematics are real and the fears of the anti-reformers not justified. . . . [T]he preceding five chapters present ample evidence that students in reform curricula are experiencing and profiting from a broader, richer curriculum. Many would argue that it is precisely these problem solving and reasoning skills as well as a knowledge of statistics, probability and discrete mathematics that are needed, not only for future success in mathematics, but for life. (pp. 1–6)

These summaries tell a clear and consistent story: with well-designed curricula, it is possible to teach for understanding without sacrificing procedural skill. Moreover, in those cases where authors conducted studies to address the issue, the data suggest that the new curricula, though hardly eradicating performance differences between Whites and underrepresented or linguistic minorities, do tend to mitigate those differences.

There has been a significant amount of controversy, called the "math wars," regarding the implementation of reform curricula. Much of the controversy has been based on opinion and anecdote, rather than grounded in reliable data. One hopes that the nature of the dialogue can change as the evidence comes in. The fact is that reform curricula (curricula aligned with the *Standards*) can be made to work as hoped. When teachers are well supported in teaching for understanding and have good curricular materials to use, children really do learn, and racial differences in performance diminish. Given this, the policy issue that needs to be addressed is what kinds of systemic support structures will promote the successful implementation of such curricula and their progressive refinement over time? That question is the topic of the next section.

Issues and Obstacles: What Does it Take to Do it Right?

The following fundamental issues, among others, must be addressed in order to insure the sustained improvement of (mathematics) instruction in the United States: (a) high quality curriculum; (b) a stable, knowledgeable, and professional teaching community; (c) high quality assessment that is aligned with curricular goals; and (d) stability and mechanisms for evolution. This section evaluates the current state of each.

Issue 1: Curriculum

For perhaps the first time in American curricular history, research-based curricula grounded in an understanding of mathematical thinking, teaching, and learning are readily available. As the data in the previous section indicate, those curricula can make a substantial positive difference when properly implemented.

The current reform curricula are rudimentary in many ways and they can stand significant improvement—they represent first attempts at implementing a new set of goals for mathematics instruction. Moreover, times change and new visions emerge: *Principles and Standards* (NCTM, 2000) moves significantly beyond the 1989 *Standards*. But that vision is also evolutionary, grounded in what has been learned over the past decade. Therein lies the key. There is now a solid curricular base from which to work. Thus, two high priority goals should be to make sure that (a) the high quality curricula currently available are actually implemented, and (b) there are mechanisms by which those curricula can be improved and updated, by means of iterative improvements in response to feedback and evolving visions of mathematical proficiency for all.

Issue 2: Teaching as a Profession

Teaching for mathematical understanding is difficult. It requires a deep understanding of the mathematics involved (see, e.g., Ma, 1999) and of how to create instructional contexts that lead students to engage with mathematics in meaningful ways. The vast majority of today's American mathematics teachers learned the traditional mathematics curriculum in the traditional way. Hence they have neither models nor experience teaching in the ways that would best facilitate their students' development of mathematical understanding.

As the data discussed in the previous section indicate, systemic alignment and sustained professional development can make a difference. When teachers are treated like professionals and they are given the opportunity to develop their skills and understandings over time, the results can be significant improvements in students' mathematical performance.

The sad fact, however, is that places where teachers are treated like professionals and are given the opportunity to develop their skills and understandings over time are few and far between. As any number of reports make clear (see, e.g., National Commission on Mathematics and Science Teaching for the 21st Century, 2000; National Commission on Teaching and America's Future, 1996), teaching is one of the most demanding and least understood or rewarded occupations in the United States. The knowledge base required for effective teaching is substantial. There are issues of subject matter knowledge, knowledge of student understanding of subject matter (pedagogical content knowledge), of understanding curricular goals, of classroom management, and more. Yet in our system, the typical expectation is that 1 year of teacher training (sometimes post-baccalaureate, sometimes not) will prepare candidate teachers to take on full responsibilities in the classroom. This is a gross underestimate of the knowledge and skills required. Worse yet, once they enter the field the vast majority of teachers have minimal opportunities for professional growth. For these reasons among others, teaching is a profession more in name than in reality.

This is a national outrage and a national pathology. This situation is not present in other nations: consider the high regard in which teachers are held in many nations around the world, and perhaps more importantly, the specifics of lesson study in nations such as Japan and China (see, e.g., Stigler & Hiebert, 1999). The expectation in those countries is that even with support, talented beginners will take a decade to evolve into fully accomplished professionals. Teachers' work is structured so that opportunities for professional growth—for example, the collaborative study, observation, and refinement of lessons and curricula that take place in lesson study—are a part of their ongoing responsibilities. There is no reason for teacher preparation and professional development to be taken any less seriously in this country. We owe it to our children.

This article begins with a discussion of mathematics education as a civil rights issue. What follows may stretch Bob Moses' ideas further than he would like, but I would argue that those of us who are interested in teacher professionalism have a great deal to learn from the civil rights movement.

Moses (2001) argued that disenfranchised voters in the American South in the 1960s were not taken seriously until *they* demanded the vote, and that by analogy, mathematics for underrepresented minorities will not be taken seriously until students demand the mathematical preparation they need.

> Constructing the Mississippi Freedom Democratic Party so that sharecroppers and day workers could have a voice was radical. . . . There had been advocates for civil rights long before. . . . Nonetheless, 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.
>
> So to understand the Algebra Project you must begin with the idea of our targeted young people finding their voice as sharecroppers and day laborers, maids, farmers, and workers of all sorts found theirs in the 1960s. (pp. 18–20)

So it is with teaching. Real and sustained change will take place when teachers demand to be treated as professionals and welcome the concomitant responsibilities. Of course, teachers cannot go it alone. We need to help redefine the contexts of teaching, teachers' responsibilities, and accountability measures so that professionalism becomes a meaningful possibility.

Issue 3: Assessment

Depending on its nature and the relationship between assessment and the curriculum, assessment can be a positive or a negative force. Let us begin with the downside. In a research symposium devoted to examining the effects of high-stakes testing, Shepard (2001) reviewed the "effects of high-stakes accountability pressures: (a) inflated test score gains, (b) curriculum distortion, and (c) loss of intrinsic motivation to learn" (p. 1). Regarding inflated test score gains, Shepard cited numerous studies indicating that when a particular high-stakes examination is put in place, scores start out low (because students have not yet been prepared for the examination) but then rise substantially as students are prepared for the test. However, a large part of the test score gains appears to be an artifact of preparation for the particular test and not necessarily an indication that the students have learned the concepts that the tests are supposedly measuring. When other forms of tests assessing the same material in different formats are used, student performance drops substantially.

Regarding Shepard's second point, she observed that high-stakes testing affects both subject matter content and the nature of instruction. Subjects for which students are not held accountable by the assessment system tend to get decreased attention. For example, the November 21, 2001 San Francisco *Chronicle* reported that fourth- and eighth-grade Californian students' performance on the 2000 NAEP science examination was the worst in the nation. One cannot prove causality for the precipitous drop in California students' science scores compared to the rest of the nation. However, the article suggests that the drop is attributable to California's recent imposition of high-stakes testing in mathematics and English. Because there are not high-stakes exams in science, teachers' attention has been diverted away from science instruction. Perhaps more importantly, even in high priority content areas those forms of instruction that do not mesh with the assessments tend to receive decreased attention. Teachers assign fewer essay questions, for example, when understandings are to be measured by multiple-choice tests. This deprives students of important learning activities.

Shepard's (2001) third point drew from Stipek's (1996) research review of motivation and instruction. Her conclusions were,

> When teachers emphasize evaluation there is a corresponding decrease in students' intrinsic motivation and interest in the material for its own sake. When students focus on how they are doing or how they will be evaluated, they become only superficially involved in learning tasks and are much less likely to persist in trying to solve difficult problems. (p. 3)

Potential abuses and misuses of assessment were part of the American Educational Research Association's (2000) motivation for issuing its *Position Statement Concerning High-Stakes Testing in PreK–12 Education*. Indeed, numerous professional organizations, including NCTM (November 2000), have taken strong stances with regard to mathematics and the responsible use of high-stakes assessments. These include consistent recommendations that decisions affecting individual students' life chances or educational opportunities should not be made on the basis of test scores alone, and that alternative assessments should be provided where test results may not provide accurate reflections of students' abilities; that assessments should cover the broad spectrum of content and thought processes represented in the curriculum, not simply those that are easily measured; and that tests must provide appropriate accommodations for students with special needs or limited English proficiency. Unfortunately, many political jurisdictions are, in the name of "standards," violating many of these recommendations.

Having raised a cautionary flag, this discussion will turn to the positive side of assessment.

> Assessment should be a means of fostering growth toward high expectations and should support high levels of student learning. When assessments are used in thoughtful and meaningful ways, students' scores provide important information that, when combined with information from other sources, can lead to decisions that promote student learning and equality of opportunity. (NCTM, November 2000)

There are indeed positive cases. Pittsburgh, where there has been a decade-long effort to align curriculum, professional development, and assessment, provides clear and dramatic evidence of improvement. It also shows that racial differences in performance can be minimized by high quality instruction for all. The Pittsburgh data show, and every data analysis chapter in Senk and Thompson (in press) confirms, that when students are taught for understanding their scores on skills go up. In addition to these findings, the Michigan data indicate that when standards, assessment, curriculum, and professional development are appropriately aligned, low-SES districts can perform as well on meaningful assessments as other much more wealthy districts.

Indeed, the use of well-designed assessments that are aligned with curricula can demonstrate the limitations of less comprehensive assessments. Describing recent work with the Mathematics Assessment Collaborative in California, Ridgway et al. (2000) report on a comparison of test scores using a standards-based assessment developed by the Mathematics Assessment Resource Service (MARS) and the Stanford Achievement Test (9th ed.), a traditional standardized test that is known as the SAT–9. Correlations between the two tests at various grade levels were typically around $r = 0.65$. Yet, there were consistent and important differences between the two tests.

Ridgway et al. (2000) examined the percentage of students who scored high or low on each test. Not surprisingly, students who scored high or low on one test tended to score similarly on the other: the percentage of students who scored high on both tests or low on both tests ranged between 70% and 77% in Grades 3, 5, and 7. The data regarding the remaining students tell an interesting story. At each grade, at least 3.5 times as many students scored high on the SAT–9 and low on MARS as the number of students who scored high on MARS and low on the SAT–9.

One possible explanation of these consistent differences in scores is that the MARS tests, which are explicitly designed to be aligned with *Principles and Standards* (NCTM, 2000), are much broader in conception than the SAT–9. Both assessments test for procedural mastery and for some degree of conceptual understanding. In addition, however, the MARS tests also call for having students work extended problems, communicate their results coherently, etc. These are core competencies that go untested on the SAT–9, so the SAT–9 would not identify students who are weak in these areas. In that sense, a high score on the SAT–9 runs the risk of being a *false positive,* certifying a student as competent when the student is unable to meet some very important mathematical standards. Once again, it is essential to use assessments that examine the full range of desired competencies.

Issue 4: Stability and Mechanisms for Evolution

The mathematics curriculum in the United States has seen any number of pendulum swings over the past 50 years. At mid-20th century, a very traditional curriculum focused on facts and procedures. Post-Sputnik, things got turned upside-down, and the new math introduced elementary students to sophisticated notions such as set theory. A backlash in the 1970s produced the back-to-basics movement. Standardized tests a decade later showed that students were no better at the basics than the students from the new math years; thus problem solving became the theme of the 1980s. This was superseded in the 1990s by standards-based mathematics, which, when it became highly visible, gave rise to the math wars and catalyzed the existence of what is in essence a neo-conservative back-to–back-to-basics movement. This way lies madness.

In much of the rest of the world there is evolutionary change, grounded in the assumption that if professionals keep working at something, they can make continual improvements. In China and Japan, for example, curricula change much less frequently and much more slowly than in the United States (Ma, 1999; National Institute on Student Achievement, Curriculum, and Assessment, 1998). To begin with, these curricula are carefully conceived and known to be reasonably effective. These curricula are refined on the basis of classroom observations and student performance. Teachers make the curriculum a collaborative object of study, working to find better ways to teach lessons or to improve them. In that way, gradual and sustained improvements are made.

Progress demands more than change for the sake of change, more than political rallying cries and radically new goals and tests every few years. Some stability in the system is necessary to make real progress. There is now the base for such sustained evolution. The content and process standards described in *Principles and Standards* (NCTM, 2000) represent the evolution of ideas developed in the 1989 *Standards*. A number of reform curricula, which have now been shown to produce results significantly better than the traditional curriculum, are aligned with the 1989 *Standards* and thus reasonably aligned with *Principles and Standards.* There exist at least two large-scale assessments (New Standards and MARS) aligned with the *Principles and Standards*. Thus, any educational system that desires has access to solid and well-aligned standards, curriculum, and assessments. An aligned program of professional development can contribute to the gradual strengthening of the teaching force and the iterative improvement of curricula and assessments.

Conclusion and Discussion

Curriculum

There now exist a number of well-designed, coherent mathematics curricula at various grade levels. There is substantial and mounting evidence that when teachers are adequately prepared to help students work through these curricula, the students learn not only skills and procedures, but also concepts and problem solving as well. Moreover, these curricula appear to represent a significant step toward equitable instruction. Performance gaps between Whites and underrepresented minorities, and between low- and high-SES students, though not eradicated, are far less dramatic than those typical of traditional curricula.

None of the extant curricula represent "the solution." All can stand improvement and as our vision gets clearer and more ambitious, changes will have to be made. But there is a solid base on which to build.

Teaching as a Profession

As indicated previously in this article, the nation is in big trouble here. In the current climate of accountability, teachers are increasingly being de-professionalized. Many of the current high-stakes accountability measures focus on skills. Given the stakes, many teachers feel that they deviate from skills-based instruction at their (and their students') peril. Partly because there are (real and perceived) weaknesses in the teaching force, a number of widely used skills-oriented curricula (in reading as well as in mathematics) are so prescriptive that little teacher discretion is allowed. This can lead to a downward spiral, since neither the curricula nor the work conditions under which most teachers operate provide opportunities for professional growth. It may contribute to high attrition rates, which contribute to teacher shortages, which result in the hiring of under-prepared teachers, who (in this way of thinking) would then need even more prescriptive teaching materials.

This cycle needs to be broken. In the ideal, new teachers would enter the profession with much more solid preparation for teaching than is now the case. Even then, however, it must be recognized that they are beginners who will take years to evolve into full-fledged professionals. As Garet, Porter, Desimone, Birman, and Yoon (in press) report, sustained professional development makes a difference:

> Sustained and intensive professional development is more likely to have an impact, as reported by teachers, than is shorter professional development. . . . Professional development that focuses on academic subject matter (content), gives teachers opportunities for "hands-on" work (active learning), and is integrated into the daily life of the school (coherence), is more likely to produce enhanced knowledge and skills.
>
> Activities that are linked to teachers' other experiences, aligned with other reform efforts, and encouraging of professional communication among teachers appear to support change in teaching practice, even after the effects of enhanced knowledge and skills are taken into account. . . . The collective participation of groups of teachers from the same school, subject or grade is related both to coherence and active learning opportunities, which in turn are related to improvements in teacher knowledge and skill and changes in classroom practice.

The National Commission on Teaching and America's Future (1996) made its recommendation in blunt terms: "Reinvent teacher preparation and professional development" (p. 76). Specifically, it recommends "states and districts need to . . . make ongoing professional development part of teachers' daily work through joint planning, research, curriculum and assessment groups, and peer coaching" (pp. 84–86). We need to redefine teachers' work lives so that meaningful opportunities for professional growth are embedded in and seen as part of their ongoing work.

Assessment

The good news is that some high quality standards-based assessments do exist, and that such assessments can be used productively as vehicles for positive change. The bad news is that the vast majority of high-stakes testing do not use such measures. The good tests tend to be more expensive than the others are, and they are harder for the general public to understand. We face a problem of public education and public relations.

The issue of public education is critically important. As long as the public continues to believe in the value and meaningfulness of traditional skills-based tests, reform faces a major uphill battle. There are at least three dimensions to this problem. First, there is the need to inform parents and others about the broad spectrum of mathematical understandings that is appropriate for students to learn—for example that problem solving, reasoning, and communication are essential goals of the curriculum, and that they need to be assessed.

Second, one has to counteract the very common misunderstanding that in mathematics students have to master skills before using them for applications and problem solving. People who believe this will focus first on skills, thinking that applications and problem solving can come later. The result is the traditional curriculum (and traditional skills-oriented assessments). An underlying assumption of reform is that students can develop mastery of skills through problem solving. As noted repeatedly in this article, the data bear out this hypothesis: students in reform curricula do just fine on skills, and much better than students in traditional curricula on concepts and problem solving. But the adoption of reform curricula is likely to be slow and controversial unless parents and other stakeholders come to understand the evidence and the fact that the skills first position is based on false assumptions.

Third, stakeholders in the educational system have to understand the great variability inherent in testing. People put great faith in the stability of test scores. The common metaphors of tests as thermometers or yardsticks point to the problem. When a child is sick, a difference of one or two degrees in a temperature reading is significant. Similarly, a few inches of height gain, or a few pounds' difference in weight, are taken as meaningful and important changes. Likewise, most people attribute significant meaning to particular test scores and to minor variations in them. In fact, test–retest differences on almost all standardized tests can be substantial—perhaps reflecting the fact that the student was having a good or bad day, that alternative forms of the same test did not really measure the same thing, or that the student was particularly lucky (or unlucky) in being familiar (or unfamiliar) with some aspect of the test.

Fourth, the public needs to understand the points raised in the previous section about high-stakes testing—that gains in test scores are often illusory or artifactual, and that high-stakes testing can result both in curricular deformation and in loss of intrinsic motivation for students. We stand a much better chance of having meaningful assessments if people understand these points.

Stability

Stability is essential if the nation is to move systemically toward meeting the vision of documents such as *Principles and Standards* (NCTM, 2000). There is now proof that real progress can be made: school districts such as Pittsburgh, which worked over the long term with a set of consistent goals and a systemic approach, get results. But stability is hard to come by. A decade is an eternity in the lives of those who have the most influence over what happens in schools; for most politicians, alas, the only things that count are the ones that produce results before the next election. As above, the educational research community confronts a problem of public education and public relations. In order to have a context where productive evolutionary change is possible, people have to be convinced that there is a solid base and that there are mechanisms for steady improvement. Absent either of these perceptions, there is the temptation for the quick fix—out with the old, in with the new—that typically results in change but not in progress. The issue of stability raises one last issue for discussion.

The Life Cycle of Ideas, From Research Into Practice

For those of us in educational research who hope to make a difference—and for policy makers and politicians who want to go about their business intelligently—it is essential to understand the amount of time it takes for ideas to make their way from basic research to successful large-scale implementation. Research on mathematics thinking, teaching, and learning provides an excellent case in point.

Contemporary research on mathematical thinking and problem solving, which represented a significant shift in both content and methods from previous research, began to take hold in the mid-1970s.[3] By the mid-1980s, the main dimensions of mathematical cognition—the nature and contents of the knowledge base, problem solving strategies, metacognition, and beliefs—were well established (see, e.g., Schoenfeld, 1985). Such contemporary understandings were at the foundations of the 1989 *Standards,* which, in a sharp break with tradition, gave significant emphasis to mathematical processes as well as to content; at each grade level, the first four standards concerned problem solving, reasoning, communication, and connections (across mathematical content and to real-world applications). In the early 1990s, the National Science Foundation issued a request for proposals for the development of new curricula consistent with the *Standards.* Experimental curricula were developed and published in the mid-1990s. Wide-scale implementation of those curricula took place in the late 1990s. As indicated in this article, the results of some of the first evaluations of those curricular implementations are just beginning to come in. The ideas work, and they work on a large scale. This really is an educational success story of major proportions.

It is important to note that the process just described, from basic research to large-scale implementation, took roughly 25 years. In industry that kind of research and development cycle is understood, but for the public at large and for most policy makers, that kind of time frame is beyond the imaginable. The point is not that one should expect to wait 25 years before good ideas become implemented in practice. There is a much shorter cycle for some ideas, and incremental change can be made. But, America is the land of the quick fix, at least with regard to education: whatever the problem is, we believe we can solve it, fast. This is a mistake, which leads to the kinds of pendulum swings discussed above. People need to understand that simple-minded solutions to complex problems will not work, and that progress is best made by building carefully on a well-established base. At the same time, researchers need to work consistently to bring good ideas into the world of practice. The best proof of the importance of research is documentation of its effects.

This article begins by pointing out the great importance of quantitative literacy in our society, and the history of inequities associated with mathematics education; it will end on an appropriately positive note. In recent years we have made substantial progress. Not only do the new mathematics curricula enable more students to do better, but also they decrease the traditional performance gaps between majority and traditionally underrepresented minority students, and between low- and high-SES students. Such curricula, aligned with robust standards and a solid assessment system, provide a solid base on which to stand. It is time to stay the course, build on what we know, and work in evolutionary fashion toward the improvement of (mathematics) education for all students.

NOTES

The author is grateful to Diane Briars for providing information about the Pittsburgh public schools. Thanks to Julia Aguirre, Claire Bove, Sally Goldstein, Ilana Horn, Cathy Kessel, Sue Magidson, Andy Porter, Ann Ryu, Natasha Speer, Elizabeth Stage, Joe Wagner, and Dan Zimmerlin for comments on earlier versions of this manuscript. Thanks also to the anonymous reviewers of this article and the editors of *Educational Researcher*, whose comments resulted in a significant re-framing of the article.

[1] NAEP trend data for 9-, 13-, and 17-year-old students covering the time period from 1978 to 1996 are available at www.nsf.ed.gov/nationsreportcard/mathematics/results/scale-ethnic.asp.

[2] The Mathematics Forum at http://forum.swarthmore.edu offers numerous links to resources in mathematics education, including web sites from both pro- and anti-reform groups. A list of NSF-supported curriculum projects that can be found at http://forum.swarthmore.edu/mathed/nsf.curric.html. The NSF has established four centers devoted to the support of standards-based curricula: the K–12 mathematics curriculum center available at www.edc.org/mcc, an elementary grades curriculum center at www.arccenter.comap.com, a middle grades center at http://showmecenter.missouri.edu, and a high school center at www.ithaca.edu/compass. Starting in 1998, the U.S. Department of Education commissioned panels of experts to identify "exemplary" and "promising" programs in mathematics education. Information regarding these programs and the review process by which they were selected can be found at www.ed.gov.offices/OERI/ORAD/KAD/expert_panel/math-science.html. (These awards were not uncontroversial. A large group of mathematicians and scientists opposed to reform wrote an open letter to then-Secretary of Education Richard Riley asking that the awards be rescinded.) Standards and other related information can be downloaded from NCTM at www.NCTM.org. The most prominent and anti-reform web site, with links to many others, can be found at www.mathematicallycorrect.com.

[3] Any one-sentence summary is an oversimplification, of course. For a more nuanced view of the history, see Schoenfeld, 2001.

REFERENCES

American Educational Research Association. (2000, July). *Position statement concerning high-stakes testing in preK–12 education*. Retrieved from http://www.aera.net/about/policy/stakes.htm

Briars, D. (March, 2001). Mathematics performance in the Pittsburgh public schools. Paper presented at a conference of the Mathematics Assessment Resource Service, San Diego, CA.

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

Chappell, M. (in press). Keeping mathematics front and center: Reaction to middle grades curriculum projects' research. In S. Senk & D. Thompson (Eds.), *Standards-oriented school mathematics curricula: What does the research say about student outcomes?* Mahwah, NJ: Erlbaum.

Garet, M., Porter, A., Desimone, L., Birman, B., & Yoon, K. (in press). What makes professional development effective: Results from a national sample of teachers. *American Education Research Journal.*

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

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

Madison, B. L., & Hart, T. A. (1990). *A challenge of numbers: People in the mathematical sciences.* Washington, DC: National Academy Press.

McKight, C., et al. (1987). *The underachieving curriculum: Assessing U.S. school mathematics from an international perspective.* Champaign, IL: Stipes Publishing Company.

Michigan Government News. (2001, April 4). Retrieved from http://www.migov.state.mi.us/gov/PressReleases/

Miller, S. (1995). *An American imperative.* New Haven, CT: Yale University Press.

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

Mullis, I., et al. (2001). *Mathematics benchmarking report, TIMSS 1999—Eighth grade: Achievement for U.S. states and districts in an international context.* Chestnut Hill, MA: Boston College. Retrieved from http://timss.bc.edu

National Action Committee for Minorities in Engineering. (1997). Engineering and affirmative action: Crisis in the making. New York: Author.

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

National Commission on Mathematics and Science Teaching for the 21st Century. (2000). *Before it's too late: A report to the nation.* Washington, DC: U.S. Department of Education.

National Commission on Teaching and America's Future. (1996). *What matters most: Teaching for America's future.* New York: Author.

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

National Council of Teachers of Mathematics. (1991). *Professional teaching standards for school mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics.* Reston, VA: Author.

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

National Council of Teachers of Mathematics. (2000, November). Position statement on high stakes testing. Retrieved from http://www.nctm.org/about/position_statements/highstakes.htm

National Institute on Student Achievement, Curriculum, and Assessment. (1998). *The educational system in Japan: Case study findings.* Washington, DC: U.S. Department of Education.

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

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

New Standards. (1996). New Standards mathematics reference examination technical summary. Pittsburgh, PA: Learning Research and Development Center, University of Pittsburgh, National Center for Education and the Economy.

Putnam, R. (in press). 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?* Mahwah, NJ: Erlbaum.

Ridgway, J., Crust, R., Burkhardt, H., Wilcox, S., Fisher, L., & Foster, D. (2000). *MARS report on the 2000 tests*. San Jose, CA: Mathematics Assessment Collaborative.

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.

Schmidt, W. (2001, April 4). TIMSS press release. Retrieved from http://USTIMSS.msu.edu/whatsnew.html

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

Schoenfeld, A. H. (2001). Mathematics education in the 20th century. In L. Corno (Ed.), *Education across a century: The centennial volume* (pp. 239–278). Chicago: National Society for the Study of Education.

Sells, L. W. (1975). Sex, ethnic, and field differences in doctoral outcomes. Unpublished doctoral dissertation, University of California, Berkeley.

Sells, L. W. (1978). Mathematics: A critical filter. *Science Teacher, 45,* 28–29.

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

Shepard, L. (2001, April). Protecting learning from the harmful effects of high-stakes testing. Paper presented at the 2001 annual meeting of the American Educational Research Association, Seattle, WA.

Stigler, J., & Hiebert, J. (1999). *The teaching gap*. New York: Free Press.

Stipek, D. J. (1996). Motivation and instruction. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 85–113). New York: Macmillan.

Swafford, J. (in press). Reaction to high school curriculum projects' research. In S. Senk & D. Thompson (Eds.), *Standards-oriented school mathematics curricula: What does the research say about student outcomes?* Mahwah, NJ: Erlbaum.

AUTHOR

ALAN H. SCHOENFELD is the Elizabeth and Edward Conner Professor of Education in the Graduate School of Education, University of California, Berkeley, 94720-1670; alans@socrates.berkeley.edu. His areas of specialization include the study of mathematical thinking, teaching, and learning.

Overseas Members!

The 2003 Annual Meeting Call for Proposals will be available in May 2002 on the AERA website, **www.aera.net,** and in the May *ER*.

Overseas membersmay request that a copy of the Call be sent to them by airmail (to arrive in advance of the May *ER*) by sending an e-mail message to AERA through the website or by sending a self-addressed envelope to

AERA
Annual Meeting 2003
1230 17th Street, NW
Washington, DC 20036-3078