A Coherent Curriculum

The Case of Mathematics

By William Schmidt, Richard Houang, and Leland Cogan

Consider the agricultural prospects of two countries: In Country A, the nation takes the best that's known about growing crops and translates it into clear, coherent, manageable guidelines for farming. These guidelines are distributed to all farmers in the country. Further, Country A makes available to all farmers up-to-date tools (tractors, balers, harvesters, etc.) and training on how to use these tools that allow them to implement the wisdom contained in the guidelines. Just as in any other country, some farmers have inherently greener thumbs than others; they find ways to surpass the guidelines and cultivate extra-rich crops. But the broad availability of the guidelines and tools puts a floor beneath farming quality. As a result, the gap between the most- and least-effective farmers is not very great, and the average quality of farming is quite good. Moreover, the average quality slowly increases as the knowledge of the best farmers is incorporated into the guidelines.

In Country B, the situation is very different. States, and sometimes towns, assemble a list of everybody's favorite ideas about farming. The list is available to any farmer who seeks it out, but it's up to the individual farmers to develop their own guidelines based on the list. The ideas are interesting, but there are too many ideas to make use of, no indications of which ideas are the best, and no pointers on which ideas fit together with other ideas. Plus, using the ideas requires tools—and training about how to use the tools. Few farmers have ready access to either.

The result: A few particularly skilled farmers in Country B figure out how to farm productively. They are mainly the farmers in more affluent areas—they have been able to attend great local agricultural schools and can afford the tools suggested by their training. A few additional farmers—those with a special knack—do fine anyway, despite their lack of training and use of poor tools. But most of Country B's farms aren't particularly efficient, certainly not in comparison with Country A's. In Country B, the gap between the most- and least-effective farms is huge, and the productivity of the average farm is far less than its Country A counterpart.

This analogy explains much of the difference between schooling and teaching in the highest achieving countries in the world and in the United States. Like the farmers in Country A, teachers in the highest achieving countries have coherent guidelines in the form of a national curriculum. They also have related tools and training—teacher's guides, student textbooks and workbooks, and preservice education—that prepare them to teach the curriculum and provide opportunities for curriculum-based professional development. In contrast, like the farmers in Country B, teachers in the U.S. have long lists of ideas about what should be taught (aka standards) and market-driven textbooks that include something for everyone but very little guidance, tools, or training.

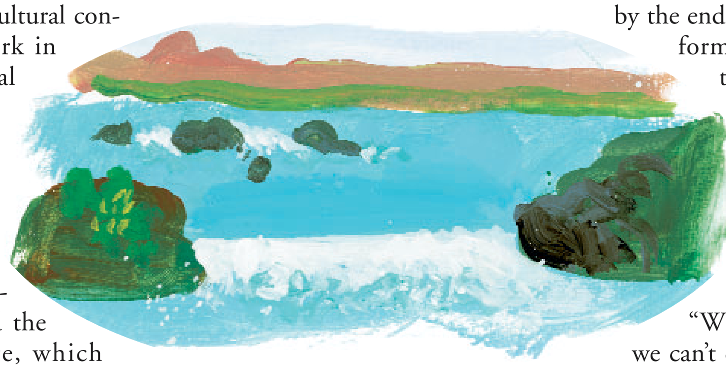
Why should we be concerned if teachers in the U.S. have to work a little harder to figure out what they are going to teach? A new analysis of data from the Third International Math and Science Study (TIMSS) provides evidence that American students and teachers are greatly disadvantaged by our country's lack of a common, coherent curriculum and the texts, materials, and training that match it.

Some people think that the purpose of an international comparison is to see which country is best and then get the U.S. to emulate its practices. That idea is naïve. You cannot

William Schmidt is the director of the U.S. National Research Center for the Third International Mathematics and Science Study (TIMSS), where Richard Houang is the associate director and Leland Cogan is a senior researcher. Other researchers who contributed to the work presented in this article are Curtis McKnight, Hsing Chi Wang, David Wiley, and Richard Wolfe. Sections of this article were adapted from "Curriculum Coherence: An Examination of U.S. Mathematics and Science Content Standards from an International Perspective," a research paper that is being prepared for publication, and from "The Implications of TIMSS for Teacher Quality," a speech delivered by Dr. Schmidt at the AFT/NEA Conference on Teacher Quality. The article is also based on research first described in Why Schools Matter, Facing the Consequences, *and* A Splintered Vision.

lift something from one cultural context and expect it to work in another. But international research can cause us to challenge some of our common assumptions about education and consider alternatives to what we are doing.

First, let us briefly review what TIMSS is and the TIMSS findings to date, which have been published in a series of previous reports. Then we will turn to our more recent findings in grades one through eight mathematics curricula, in which we can see that high-performing countries teach a very similar, very coherent, core math curriculum to all of their students—and we, decidedly and clearly, do not. Lastly we will look at the importance of this finding by examining the cascade of benefits that flow from attaining a coherent, common curriculum.



I. The Early TIMSS Findings

TIMSS is the most extensive and far-reaching cross-national comparative study ever attempted. It was conducted in 1995, with 42 countries participating in at least some part of the study. TIMSS tested three student populations: those who were mostly nine years old (grades three and four in the U.S.); those who were mostly 13 years old (grades seven and eight in the U.S.); and students in the last year of secondary school (12th grade in the U.S.). In addition to the student tests, the study included a great deal of other data collection, including extensive studies of curriculum. Findings from the curriculum study are the heart of this article; but first, let's review what's already been reported in the general press about TIMSS.

The Horse Race

The horse race—who comes in first, second, and third—is not particularly important in and of itself. In fact, the ranking of nations is simply the two-by-four by which to get people's attention.

At the fourth-grade level, the U.S. did reasonably well on the TIMSS exam. Our students scored above the international average in both math and science. In science, in fact, we came very close to being number one in the world; our fourth-graders were second only to the South Koreans. In mathematics, on the other hand, our performance was only decent; it was above average, though not in the top tier of countries. (Detailed findings, including tables and graphs, can be found on our Web site, **http://ustimss.msu.edu**, or at the U.S. Department of Education's TIMSS Web site, **http://nces.ed.gov/timss**).

By eighth grade, however, the U.S. dropped to the international average, slightly above average in science and slightly below average in mathematics. In other words, just four years along in our educational system, our scores fell to average or even below average. The decline continues so that by the end of secondary school our performance is near the bottom of the international distribution. In both math and science, our typical graduating senior outperformed students in only two other countries: Cyprus and South Africa.

Some people might ask, "What difference does it make if we can't do fancy math problems?" It does make a difference. A typical item on the TIMSS 12th-grade math test shows a rectangular wrapped present, provides its height, width, and length, as well as the amount of ribbon needed to tie a bow, and asks how much total ribbon would be needed to wrap the present and include a bow. Students simply need to trace logically around the package, adding the separate lengths so as to go around in two directions and then add the length needed for the bow. Only one-third of U.S. graduating seniors can do this problem, however. This is serious.

Another part of the 12th-grade TIMSS study involved advanced students, those taking courses like calculus or college-preparatory physics. The results are quite startling: We are near the bottom of this international distribution also. In the past, when international results have been reported, many people have suggested, "It's really not a problem because our best students are doing okay." That's simply not true. In fact, a comparison of mathematics scores in 22 countries revealed that U.S. eighth-graders who scored at the 75th percentile were actually far below the 75th percentile in 19 of the other countries. The most dramatic results were in comparison to Singapore—a score at the 75th percentile in the U.S. was below the 25th percentile in Singapore. The problems we must address affect not only our average students, but even those who are above average.

Curriculum Matters: What You Teach is What You Get

Now these horse race results are interesting and disquieting. But they hide important results that we think help with understanding our poor performance and give us the keys to fixing it. To really understand the TIMSS results, you have to examine student achievement in different areas of the curriculum within math and science.

When you look at the performance of eighth-grade students in different math and science content areas, you will find that U.S. performance is remarkably different on different topics. And, the same is true for virtually every other country. For example, Singapore was number one in science at eighth grade, but students there were not number one in all of the different science areas.

One of the most important findings from TIMSS is that the differences in achievement from country to country are related to what is taught in different countries. In other words, this is not primarily a matter of demographic variables or other variables that are not greatly affected by schooling. What we can see in TIMSS is that schooling makes a difference. Specifically, we can see that *the curricu-*

lum itself—what is taught—makes a huge difference.

Consider the performance of Bulgarian students in science. They were tops in the world in the area of the structure of matter, but almost dead last in the area of physical changes. Consider, too, the remarkable variations in U.S. performance in mathematics. Our eighth-grade students did their very best math work in the area of rounding. Our kids are among the world's best rounders. We obviously teach it thoroughly. But based on the TIMSS results, we are obviously not doing an adequate job of teaching measurement; perimeter, area and volume; and geometry.

These findings emerged from a substantial line of research within TIMSS that examined what is taught in 37 countries. To get a rich picture of math and science instruction in each country, we looked at the "intended" content—that is, what officials intended for teachers to teach; and "enacted" content—that is, what teachers actually taught in their classrooms. In most countries, the intended content was simply the national curriculum. But in the handful of countries without a national curriculum, we sought out other formal statements of intended content at the regional or local level. For example, in the U.S. we examined state and district standards. In all of the countries we determined the enacted content by surveying teachers about what they believed they had covered. Additional information on what is taught came from a review of several major textbooks in each country and, in a few countries, classroom observations.

Based on these studies of the "intended" and "enacted" content in mathematics, we can make some general claims. We know that in most countries studied, the intended content that is formally promulgated (at the national, regional, or state level) is essentially replicated in the nation's textbooks. We can also say that in most countries studied, teachers "follow" the textbook. By this we mean that they cover the content of the textbook and are guided by the depth and duration of each topic in the textbook. From this knowledge, we can say with statistical confidence that what is stated in the intended content (be it a national curriculum or state standards) and in the textbooks is, by and large, taught in the classrooms of most TIMSS countries. Knowing all of this, we can often trace the strengths and weaknesses that a nation's students display on given topics to comparable strengths and weaknesses in the intended content. In short, our study shows clearly that curriculum matters. If a nation asks teachers to teach a particular set of topics in a particular grade, that is what teachers will likely teach—and, in the aggregate, it is what students will likely learn. This was true even after we controlled for students' socioeconomic status.[1]

Curricula in the U.S.: A Mile Wide, an Inch Deep

Based on these early analyses of TIMSS data, we can characterize the intended math and science content (as stated in sets of standards and textbooks) in the U.S., relative to others in the world, in four ways:

1. Our intended content is not focused. If you look at state standards, you'll find more topics at each grade level than in any other nation. If you look at U.S. textbooks, you'll find there is no textbook in the world that has as many topics as our mathematics textbooks, bar none. In fact, according to TIMSS data, eighth-grade mathematics textbooks in Japan have around 10 topics, but U.S. eighth-grade textbooks have over 30 topics. (See photo on page 20.) And finally, if you look in the classroom, you'll find that U.S. teachers cover more topics than teachers in any other country.

2. Our intended content is highly repetitive. We introduce topics early and then repeat them year after year. To make matters worse, very little depth is added each time the topic is addressed because each year we devote much of the time to reviewing the topic.

3. Our intended content is not very demanding by international standards. This is especially true in the middle-school years, when the relative performance of U.S. students declines. During these years, the rest of the world shifts its attention from the basics of arithmetic and elementary science to beginning concepts in algebra, geometry, chemistry, and physics.

4. Our intended content is incoherent. Math, for example, is really a handful of basic ideas; but in the United States, mathematics standards are long laundry lists of seemingly unrelated, separate topics. Our most recent analysis has more to say about this and we will return to it in the next section.

As a result of these poorly designed standards and textbooks, the curriculum that is enacted in the U.S. (compared to the rest of the world) is highly repetitive, unfocused, unchallenging, and incoherent, especially during the middle-school years. There is an important implication here. Our teachers work in a context that demands that they teach a lot of things, but nothing in-depth. We truly have standards, and thus enacted curricula, that are a "mile wide and an inch deep."

One popular response to a study like TIMSS is to blame the teachers. But the teachers in our country are simply doing what we have asked them to do: "Teach everything you can. Don't worry about depth. Your goal is to teach 35 things briefly, not 10 things well."

II. The Coherent Curriculum

Discussion of the TIMSS achievement results has prompted policymakers in the U.S. and elsewhere to wonder just what it might mean to have a world-class mathematics or science curriculum. In response to this interest, we investigated the top achieving TIMSS countries' curricula in mathematics and science to distill what they considered essential content for virtually all students[2] over the different grades of schooling. With this new analysis, we can go beyond the critique of our "mile-wide-inch-deep curricula" and look at the character and content of a world-class curriculum.[3] Although we conducted this analysis for both math and science, in this article we will only address the math findings.

After identifying the top achieving (or A+) countries and devising a methodology to determine the topics that were common to their curricula, we developed a composite set of topics consisting of the topics that at least two-thirds of the A+ countries included in their curricula. This A+ composite is displayed in Figure 1. Next, composites for U.S. mathematics standards from 21 states (Figure 2) and 50 districts (Figure 3, page 11) were also developed and compared to the A+ composite. (For more details on the methodology, please see page 47.)

While examining the A+ composite, it is important to keep in mind that this figure represents a "core" curriculum, not a complete curriculum. Our goal in developing the composite was to find out which topics at least two-thirds of A+ countries believed to be essential. Not surprisingly, these countries' points of agreement resulted in a smaller set of topics in our composite than any one of these countries includes in its national curriculum.[4]

To represent the full scope of a complete mathematics curriculum in a typical A+ country, roughly three topics would have to be added at each grade level in addition to those listed in Figure 1. As noted in the last line of Figure 1, the average number of topics that would have to be added range from one (in grades four and five) to as many as six (in grades two and seven). This is important information for Americans who understand that there is a need for a com-

FIGURE 1

A+ Composite: Mathematics topics intended at each grade by at least two-thirds of A+ countries.

Note that topics are introduced and sustained in a coherent fashion, producing a clear upper-triangular structure.

TOPIC ***GRADE:***	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***
Whole Number Meaning	■	■	■	▒	▒			
Whole Number Operations	■	■	■	■	▒			
Measurement Units	□	■	■	■	■	■	▒	
Common Fractions			□	■	■	▒		
Equations & Formulas			□	▒	▒	▒	■	■
Data Representation & Analysis			□	□	▒	▒		□
2-D Geometry: Basics			□	▒	▒	▒	■	■
Polygons & Circles				▒	▒	▒	■	■
Perimeter, Area & Volume				▒	▒	▒	▒	□
Rounding & Significant Figures				▒	▒			
Estimating Computations				▒	▒	▒		
Properties of Whole Number Operations				□	▒			
Estimating Quantity & Size				□	□			
Decimal Fractions				▒	■	▒		
Relationship of Common & Decimal Fractions				▒	■	▒		
Properties of Common & Decimal Fractions					▒	▒		
Percentages					▒	▒		
Proportionality Concepts					▒	▒	▒	□
Proportionality Problems					▒	▒	■	■
2-D Coordinate Geometry					□	□	▒	▒
Geometry: Transformations						▒	▒	▒
Negative Numbers, Integers & Their Properties						□	▒	
Number Theory							▒	□
Exponents, Roots & Radicals							▒	▒
Exponents & Orders of Magnitude							□	□
Measurement Estimation & Errors							□	
Constructions w/ Straightedge & Compass							■	□
3-D Geometry							▒	■
Congruence & Similarity								■
Rational Numbers & Their Properties								□
Patterns, Relations & Functions								□
Slope & Trigonometry								□
Number of topics covered by at least 67% of the A+ countries	3	3	7	15	20	17	16	18
Number of additional topics intended by A+ countries to complete a typical curriculum at each grade level	2	6	5	1	1	3	6	3

□ – intended by 67% of the A+ countries ▒ – intended by 83% of the A+ countries ■ – intended by 100% of the A+ countries mon, prescribed curricular core, but also believe some local discretion must be accommodated. The A+ composite shows that, at least in math, it is eminently sensible and doable to think of some math topics as part of a required core taught in particular grades and others as topics that can float according to, say, state or district discretion.

The A+ Composite

Figure 1 presents the A+ composite for mathematics by topic and grade. The 32 topics listed are those that are in the national curricula at a given grade in at least two-thirds of the A+ countries. As evidenced by the "upper-triangular" shape of the data, the A+ composite reflects an evolution from an early emphasis on arithmetic in grades one through four to more advanced algebra and geometry beginning in grades seven and eight. Grades five and six serve as a transitional stage in which topics such as proportionality and coordinate geometry are taught, providing a bridge to the formal study of algebra and geometry.

More specifically, these data suggest a three-tier pattern of

FIGURE 2

State Composite: Mathematics topics intended at each grade by at least two-thirds of 21 U.S. states.

Note that topics are introduced and sustained in a way that produces no visible structure.

TOPIC ***GRADE:***	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***
Whole Number Meaning	■	■	■	■	▨	□		
Whole Number Operations	■	■	■	■	▨	□		
Measurement Units	▨	■	■	▨	■	■	■	▨
Common Fractions	□	▨	▨	▨	■	▨	□	□
Equations & Formulas	□	□	▨	▨	▨	■	■	■
Data Representation & Analysis	■	■	■	■	■	■	■	▨
2-D Geometry: Basics	▨	▨	▨	▨	▨	▨	▨	▨
Polygons & Circles	■	■	■	■	■	■	■	▨
Perimeter, Area & Volume		□	□	□	▨	■	■	▨
Rounding & Significant Figures								
Estimating Computations	□	□	▨	▨	▨	▨	▨	▨
Properties of Whole Number Operations	□	□	□	□				
Estimating Quantity & Size			□					
Decimal Fractions			□	▨	▨	▨	□	□
Relationship of Common & Decimal Fractions				□	□	□		
Properties of Common & Decimal Fractions								
Percentages					□	▨	▨	□
Proportionality Concepts						▨	□	
Proportionality Problems						▨	▨	▨
2-D Coordinate Geometry			□	▨	□	□	□	▨
Geometry: Transformations	▨	▨	▨	▨	▨	▨	▨	▨
Negative Numbers, Integers & Their Properties						□	▨	□
Number Theory					▨	□	□	□
Exponents, Roots & Radicals						□	□	■
Exponents & Orders of Magnitude							□	□
Measurement Estimation & Errors	□	□	▨	□	▨	▨	▨	□
Constructions w/ Straightedge & Compass								
3-D Geometry	■	■	■	▨	■	▨	■	▨
Congruence & Similarity					□	▨	▨	□
Rational Numbers & Their Properties						▨	▨	□
Patterns, Relations & Functions	▨	■	■	■	▨	■	■	■
Slope & Trigonometry								
Number of topics covered by at least 67% of the states	14	15	18	18	20	25	23	22
Number of additional topics intended by states to complete a typical curriculum at each grade level	8	8	7	8	8	5	6	6

□ – intended by 67% of the states ▨ – intended by 83% of the states ■ – intended by 100% of the states increasing mathematical complexity. The first tier includes an emphasis primarily on arithmetic, including common and decimal fractions, rounding, and estimation. It is covered in grades one through four. The third tier, covered in grades seven and eight, consists primarily of advanced number topics such as number theory (including primes and factorization, exponents, roots, radicals, orders of magnitude, and rational numbers and their properties), algebra (including functions and slope), and geometry (including congruence and similarity, and 3-dimensional geometry). Grades five and six appear to serve as an overlapping transitional tier with continuing attention to a few arithmetic topics, but also with an introduction to more advanced topics such as percentages; negative numbers, integers and their properties; proportional concepts and problems; two-dimensional coordinate geometry; and geometric transformations.

The curriculum structure also includes a small number of topics that provide a form of continuity across all three tiers. These continuing topics (such as measurement units, which are covered in grades one through seven, and equations and formulas, which are covered in grades three through eight) seem to support the overall curriculum structure. These topics have an

The Benefit to Equity

By E.D. Hirsch, Jr.

When children share a common base of knowledge, their classroom instruction can be far more effective. Why is this? Anyone who has ever taught a class knows that explaining a new subject will induce smiles of recognition in some students, but looks of puzzlement in others. Every teacher who reads exams has said or thought, "Well, I *taught* them that, even if some of them didn't *learn* it." What makes the click of understanding occur in some students, but not in others?

Research has shown that the ability to learn something new depends on an ability to accommodate the new thing to the already known. When the automobile first came on the scene, people called it a "horseless carriage," thus accommodating the new to the old. When a teacher tells a class that electrons go around the nucleus of an atom as the planets go around the sun, that analogy may be helpful for students who already know about the solar system, but not for students who don't. Relevant background knowledge gives students a greater variety of means for capturing the new ideas.

This enabling function of relevant prior knowledge is essential at every stage of learning.

When a child "gets" what is being offered in a classroom, it is like someone getting a joke. A click occurs. People with the requisite background knowledge will get the joke, but those who lack it will be puzzled until somebody explains the background knowledge that was assumed in telling the joke. A classroom of 25 to 30 children cannot move forward as a group until all students have gained the taken-for-granted knowledge necessary for "getting" the next step in learning. If the class must pause too often while its lagging members are given background knowledge they should have gained in earlier grades, the progress of the class is bound to be excruciatingly slow for better-prepared students. If, on the other hand, instead of slowing down the class for laggards, the teacher presses ahead, the less-prepared students are bound to be left further and further behind.

For effective classroom learning to take place, class members need to share enough common reference points to enable all students to learn steadily, albeit at differing rates and in response to varied approaches. Harold Stevenson and James Stigler in their important book, *The Learning Gap,* show that when this requisite commonality of preparation is lacking, as it is in most American classrooms today, the progress of learning will be slow compared with that of educational systems that do achieve commonality of academic preparation within the classroom. It is arguable that this structural difference between American classrooms and those of more effective systems is an important cause of the poor showing of American students in international comparisons.

The learning gap that Stevenson and Stigler describe is a gap in academic performance between American and Asian students. Subsequently, work by Stevenson and his colleagues has shown that this gap grows wider over time, putting American students much further behind their Asian peers by 11th grade than they were in the sixth grade. The funnel shape of this widening international gap has an eerie similarity to the funnel shape of the widening gap *inside* American schools between advantaged and disadvantaged students as they progress through the grades. A plausible explaimplied breadth that means they could move from their most elementary aspects to the beginning of complex mathematics during the elementary and middle grades.

Another pattern identified in Figure 1 is the number of grades in which a topic is covered in the A+ composite—mathematics topics in these countries are generally intended for an average span of three years. Only eight out of the 32 topics are covered for five or more years. In addition, five out of the 32 topics are covered for only one year in grades one through eight. (These five topics reappear in the upper secondary mathematics curricula of A+ countries, but Figure 1 does not include this information.) As you will see, the short duration of topic coverage stands in stark contrast to the U.S.

These data indicate that across the A+ countries there is a generally agreed-upon set of mathematics topics—those related to whole numbers and measurement—that serve as the foundation for mathematics understanding. They constitute the fundamental mathematics knowledge that students are meant to master during grades one to five. Future mathematics learning builds on this

nation for the widening in both cases is that a lack of academic commonality in the American classroom not only slows down the class as a whole thus making us lag behind other countries, but also creates an increasing discrepancy between students who are lucky enough to have gained the needed background knowledge at home and those who have to depend mainly on what they get sporadically in school. The learning of luckier students snowballs upon their initial advantage while that of the less fortunate ones—those dependent for their learning on what the incoherent American school curricula offer—never even begins to gather momentum. The lack of shared knowledge among American students not only holds back their average progress, creating a national excellence gap, but more drastically, holds back disadvantaged students, thus creating a fairness gap as well.

What chiefly makes our schools unfair, then, even for children who remain in the same school year after year, is that some students are learning less than others, not because of their innate lack of academic ability or their lack of willingness to learn, but because of the inherent shortcomings in curricular organization. A systemic failure to teach all children the knowledge they need in order to understand what the next grade has to offer is *the* major source of avoidable injustice in our schools. It is impossible for a teacher to reach all children when some of them lack the necessary building blocks of learning. Under these circumstances, the most important single task of an individual school is to ensure that all children within that school gain the prior knowledge they will need at the next grade level. Since our system currently leaves that supremely important task to the vagaries of individual classrooms, the result is a systemically imposed unfairnesss even for students who remain in the same school. Such inherent unfairness is greatly exacerbated for children who must change schools, sometimes in the middle of the year.

A systemic failure to teach all children the knowledge they need in order to understand what the next grade has to offer is *the* major source of avoidable injustice in our schools.

Consider the plight of Jane, who enters second grade in a new school. Her former first-grade teacher deferred all world history to a later grade, but in her new school, many first-graders have already learned about ancient Egypt. The new teacher's references to the Nile River, the Pyramids, and hieroglyphics simply mystify Jane and fail to convey to her the new information that the allusions were meant to impart. Multiply that incomprehension by many others in Jane's new environment, and then multiply those by further comprehension failures which accrue because of the initial failures of uptake, and we begin to see why Jane is not flourishing academically in her new school. Add to these academic handicaps the emotional devastation of not understanding what other children are understanding, and add to avoidable academic problems the *un*avoidable ones of adjusting to a new group, and it is not hard to understand why newcomers fail to flourish in American schools. Then add to all of these drawbacks the fact that the social group with the greatest percentage of school changers is made up of low-income families who move for economic reasons, and one understands more fully why disadvantaged children suffer disproportionately from the curricular incoherence of the American educational system.

E.D. Hirsch, Jr. is the founder of the Core Knowledge Foundation and a professor emeritus of education and humanities at the University of Virginia. This sidebar is excerpted with permission of Doubleday, a division of Random House, Inc., from The Schools We Need and Why We Don't Have Them, *©1995 by Doubleday, a division of Bantam Doubleday Dell Publishing Group.*

foundation. At the middle and upper grades, new and more sophisticated topics are added—and, significantly, *the foundation topics then disappear from the curriculum.*

A Structure that Reflects the Discipline of Mathematics

To date, most discussions and evaluations of the quality of American standards have revolved around such characteristics as clarity, specificity, and, often, a particular ideology. For example, in mathematics these distinctions have been revealed in what is called the "math wars," a debate over what constitutes basic mathematics for the school curriculum.

With our look at the A+ composite, our definition of

The Benefit to Subject-Matter Knowledge

In this article, we discuss America's curriculum gap—the difference between the quality of our curriculum and that of the A+ countries. Others (especially Harold Stevenson and Jim Stigler) have written about a learning gap and a teaching gap. Perhaps one of the biggest gaps—and it's related to the others—is the subject-matter knowledge gap that exists between our mathematics teachers and those in the highest performing countries. If we are serious about making our math curriculum more rigorous, this gap—which reflects the limited subject-matter preparation that many of our teachers receive—will have to be addressed.

In 2001, a survey asked a sample of Michigan teachers if they felt prepared to teach 12 specific mathematics topics such as equations, porportionality concepts, and data representation concepts. How many teachers thought they were prepared to teach all 12? Ten percent of the third-grade teachers, 20 percent of the fourth- and fifth-grade teachers, 45 percent of the sixth-grade teachers, about half of the seventh- and eighth-grade teachers, and only three-fourths of the high-school teachers felt adequately prepared, in a subject matter sense, to teach all 12 topics. Teachers recognize the inadequacy of their training for teaching the more advanced curriculum that we need in order to close the learning gap.

To better understand why this subject-matter gap exists, we must again look abroad to reflect on our own practices. To begin with, in the A+ countries, candidates for middle- and secondary-teaching positions would typically have a strong math background, often including the equivalent of a major in the subject. Even elementary teachers, by virtue of having been educated in these systems, would have quite substantial math backgrounds. This is not trivial and must be addressed as we consider criteria for hiring the next generation of teachers. But I want to focus here on a different aspect of these foreign systems: their equivalent of in-service education, or professional development.

In the high-achieving nations, there is a clearly articulated curriculum specific to each grade, which is usually common for the entire country.

But don't mistake the curriculum itself for the wonder drug. These nations also make carefully planned professional-development investments.

Significantly, these high-achieving nations generally do not attempt generic sorts of professional development, a practice which is fairly common in this country, where, on occasion, you take all the K-12 teachers and put them into one room and call it professional development. Professional development in high-performing countries is generally geared to the grade in which teachers teach. The subject matter content and how to teach it are often the focus. It is about the content that they are teaching their students in the classroom, not about abstract mathematical or other content. In turn, it's not necessary to teach all teachers in a particular field, like mathematics, advanced topics—not all math teachers need to take and know calculus. What fourth-grade teachers need, for example, is an advanced treatment of elementary mathematics. They need to know, for instance, that fractions are part of a rational numbers system. Fractions aren't alien beasts to whole numbers, but they are often presented that way. Deeper knowledge of the structure of the advanced parts of elementary mathematics would enable fourth-grade teachers to carry out the kind of instruction that demonstrates connections between mathematical concepts.

Further, the textbook connection cannot be ignored when thinking about professional development. In the U.S., the correlation between textbook coverage and what teachers teach is .95 (which is comparable to other countries). If we pretend the textbook doesn't exist—and conduct professional development in ways that assume teachers will implement an entirely different approach to the content than the texts take—believe me, the textbook will win. Professional development must be intimately tied to the actual tools teachers use. That's the essence of curriculum-based professional development.

—WILLIAM SCHMIDT

quality moves beyond these issues to what we believe is a deeper, more fundamental characteristic. We feel that one of the most important characteristics defining quality in content standards is what we term coherence.

We define content standards and curricula to be coherent if they are articulated over time as a sequence of topics and performances that are logical and reflect, where appropriate, the sequential or hierarchical nature of the disciplinary content from which the subject matter derives. That is, what and how students are taught should reflect not only the topics that fall within a certain academic discipline, but also the key ideas that determine how knowledge is organized and generated within that discipline.

This implies that "to be coherent," a set of content standards must evolve from particulars (e.g., the meaning and operations of whole numbers, including simple math facts and routine computational procedures associated with whole numbers and fractions) to deeper structures inherent in the discipline. This deeper structure then serves as a means for connecting the particulars (such as an understanding of the rational number system and its properties). The evolution from particulars to deeper structures should occur over the school year within a particular grade level and as the student progresses across grades.

Based on this definition of coherence, the A+ composite is very strong and seems likely to build students' understanding of the big ideas and the particulars of mathematics and to assure that all students are exposed to substantial math content.

In sum, the "upper-triangular" structure of the data in Figure 1 implies that some topics were designed to provide a base for mathematics understanding and, correspondingly,

A Glimpse of an A+ Curriculum…and How It Is Used

Basic Content/Objectives	Detailed Content	Time Ratio	Notes on Teaching
Rate, ratio, and proportion **Objectives:** To develop the ability in the use of rate, ratio, and proportion in problems connected with everyday life.	1.1 Meaning of rate, ratio, and proportion	3	Students are expected to understand clearly the meaning of rate, ratio, and proportion through using everyday examples such as walking rate, reduction rate, and the ratio of the number of boys to that of girls in a class. These examples should lead students to see their relationship.
	1.2 The notion of a two-term ratio a:b or a/b, where b≠0	2	The notion of a two-term ratio a:b is introduced. This can be represented by the fraction a/b, where b≠0. Students should note that a ratio is unaltered if the two numbers (or quantities) of the ratio are both multiplied or divided by the same number. The notion of a two-term ratio may be extended to a three-term ratio or more, e.g. a:b:c=1:2:3.
	1.3 Examples from science and mensuration [i.e., measurement] including similar triangles. Problems on direct and simple inverse proportion. Graphs in two variables	6	Students should be able to deal with rate, ratio, and proportion in examples from science and mensuration, including similar triangles. Practical problems on direct and simple inverse proportion should also be investigated. (N.B. Maps and scale plans are common examples of proportion.) Students may use graphs to see the relationship between two quantities.

Source: Hong Kong eighth-grade curriculum, excerpted from the Syllabus for Mathematics: Forms I-V, *the curriculum that was in effect until spring of 2001 (and during the TIMSS).*

Unlike a typical set of state standards in this country, the Hong Kong curriculum contains much more than just the content that teachers ought to cover, yet the information it provides is not overwhelming. The time ratio provided allows teachers and others taking guidance from the curriculum to easily see which topics should be emphasized, though it does not put teachers on a strict schedule. The notes on teaching offer valuable tips and examples and explain how topics and subtopics relate. As another example, consider these notes on teaching seventh-graders the use of letters to represent numbers:

> The use of letters to represent numbers arises quite naturally when formulae such as A = b + l and P = 2(b + l), where A stands for area, b for breadth, l for length, and P for perimeter of a rectangle, are considered. Teachers can point out that even in daily life, letters may be used to represent numbers, e.g. in a secret code.

Keep in mind that this curriculum is the beginning, not the end, of support for instruction. It serves as the basis for a raft of well-aligned classroom materials, including:

- classroom assessments for teachers to use at their discretion;
- highly focused textbooks that flesh out the curriculum with closely-aligned explanations and problem sets;
- preservice education that prepares teachers to teach the curriculum; and
- multiple opportunities for content-based professional development.

were covered in the early grades. Increasingly over the grades, the curricula of the top achieving countries becomes more sophisticated and rigorous in terms of the mathematics topics covered. As a result, it reflects a logic that we would argue is inherent in the nature of mathematics itself. As we will see, the U.S. state and district standards do not reflect a comparable logical structure.

The A+ composite is stunningly coherent, and it's a pole star that can guide our curriculum and standards-writing efforts. But the huge educational impact of the curriculum in A+ countries lies in several additional related facts: In every A+ country, there is a single national curriculum.[5] It does not sit on a shelf unread and unused, nor is it an exceedingly long document that teachers pick through on their own, selecting which topics to emphasize and de-emphasize. The national curriculum as a whole is meant to be the enacted curriculum; related training, tools, and assessments are provided that make such enactment possible (and likely). The curriculum's coherence is translated into textbooks, workbooks, diagnostic tests for teacher use, and other classroom materials that enable teachers to bring the curriculum into the classroom in a relatively consistent, effective way. In turn, the curriculum serves as an important basis for the nation's preservice teacher education and for ongoing professional development, which again adds to the generally consistent, high quality of teaching across classrooms and schools.

Underlying all of this and making it all possible, is the fact that the curriculum is common—that is, the same coherent set of topics is intended to be taught in the same grade to virtually every child in the country—at least from grades one through eight (the focus of our study). Regardless of which school you attend or to which teacher you are assigned, the system is designed so that you will be exposed to the same material in the same grade.

PHOTOGRAPH COURTESY OF LELAND COGAN

Mathematics textbooks in the U.S. cover more topics than texts in other countries, and, as a result, are substantially larger. The photo above compares five eighth-grade texts commonly used in the U.S. (right) to the eighth-grade texts from five of the A+ countries, which often use two slim books per year (left).

This common, coherent curriculum makes possible a cascade of benefits for students' education. The possible net effects of these benefits are: 1) to positively influence overall student achievement (as reported in the opening section of this article); 2) to greatly reduce the differential achievement effects that are produced (in the U.S.) by standards and curricula of different quality; and, as a result, 3) to substantially weaken the relationship between student achievement and socioeconomic status (a link which is quite strong in the U.S.).

III. Repetition and Incoherence in the U.S.

As we know, unlike the A+ countries, the U.S. does not have a single, national curriculum. To determine the intended math curriculum, we looked primarily at the math standards that have been established at the state level. We also reviewed district-level standards.

State Standards

In Figure 2 we show a composite of the math standards in the 21 states that volunteered for our study. Since Figure 1 includes topics that were intended by at least two-thirds of the A+ countries, a similar two-thirds majority was applied to create the state composite shown in Figure 2 (on page 5).[6] The resulting pattern for the composite of U.S. states is very different from that of the A+ countries. The state standards do not reflect the three-tier structure described previously. The majority of the 32 mathematics topics that A+ countries teach at some point in grades one through eight are likely to be taught to American students repeatedly throughout elementary and middle school.[7] *In fact, the average duration of a topic in state standards is almost six years. This is twice as long as for the A+ countries.*

This long duration means that U.S. states include many more topics at each grade than do A+ countries. That, in turn, means each topic is addressed in less depth. In general, the state standards increase the duration of a typical topic by introducing it at an earlier grade. For instance, even more demanding topics such as geometric transformations, measurement error, three-dimensional geometry, and functions are introduced as early as first grade. In the A+ composite, these same topics are first covered in middle school.

If coherence means that the internal structure of the academic discipline is reflected within and across grades, then clearly these results for U.S. states suggest a lack of coherence, even if the claim is that these topics are only presented initially in an elementary or introductory fashion. The U.S. standards, with their early introduction and frequent repetition of topics, appear to be just an arbitrary collection of topics. Here are several specific examples of this incoherence:

■ **Prerequisite knowledge doesn't come first.** For example, properties of whole number operations (such as the distributive property) are intended to be covered in first grade, the same time that children are beginning to study basic whole-number operations. This topic is first typically introduced at grade four (and not earlier than grade three) in the top-achieving countries.

■ **Topics endure endlessly.** The A+ composite did not intend for any topic to be covered at all eight grades, yet 10 topics were intended for such enduring coverage in the state composite.

■ **Consensus about when to teach topics is lacking.** The state composite has blank rows for three fundamental topics—rounding and significant figures, the properties of common and decimal fractions, and slope. This odd finding reflects the lack of consensus among states as to the appropriate grade level for these topics. The state standards all cover rounding and significant figures, as well as common and decimal fractions, but these topics cannot be part of the state composite because at least two-thirds of the states do not agree on the proper grade placement for these topics. The absence of slope from the state composite reflects both a lack of agreement and a lack of rigor—most states do not

FIGURE 3

District Composite: Mathematics topics intended at each grade by at least two-thirds of 50 districts in one state.

Note that the structure of the district composite is very similar to that of the state composite—and likewise, lacks a visible structure.

TOPIC **GRADE:**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Whole Number Meaning	▦	▦	▦	▦	▦	□	□	□
Whole Number Operations	▦	▦	▦	▦	▦	▦	□	□
Measurement Units	■	■	▦	▦	▦	▦	▦	▦
Common Fractions		▦	▦	▦	▦	▦	▦	□
Equations & Formulas	□	▦	□	▦	▦	▦	▦	▦
Data Representation & Analysis	▦	▦	▦	▦	▦	▦	▦	▦
2-D Geometry: Basics		□	▦	□	□	▦	▦	▦
Polygons & Circles	▦	▦	▦	▦	▦	▦	▦	▦
Perimeter, Area & Volume			▦	▦	▦	▦	▦	▦
Rounding & Significant Figures				□			□	
Estimating Computations		□	□	▦	□	□	□	□
Properties of Whole Number Operations		□	□					
Estimating Quantity & Size								
Decimal Fractions			□	□	▦	▦	▦	▦
Relationship of Common & Decimal Fractions							□	
Properties of Common & Decimal Fractions								
Percentages					□	▦	▦	▦
Proportionality Concepts						□	□	□
Proportionality Problems						□	□	▦
2-D Coordinate Geometry						□	□	□
Geometry: Transformations		□	□	□	□		□	□
Negative Numbers, Integers & Their Properties							□	□
Number Theory					□		□	□
Exponents, Roots & Radicals							□	□
Exponents & Orders of Magnitude							▦	□
Measurement Estimation & Errors			□			□	□	□
Constructions w/ Straightedge & Compass								
3-D Geometry	▦	▦	▦	▦	▦	▦	▦	▦
Congruence & Similarity							□	□
Rational Numbers & Their Properties							□	▦
Patterns, Relations & Functions	▦	▦	▦	▦	▦	▦	▦	▦
Slope & Trigonometry								
Number of topics covered by at least 67% of the districts	8	13	16	15	16	18	27	25
Number of additional topics intended by districts to complete a typical curriculum at each grade level	9	6	4	7	8	9	3	4

□ – intended by 67% of the districts ▦ – intended by 83% of the districts ■ – intended by 100% of the districts intend for slope to be covered until high school.

The longer topic coverage combined with the absence of the three-tier structure suggest that state standards are developed from a laundry-list approach to mathematics that lacks any sense of the logic of mathematics as a discipline. For many of the individual states it seems that almost all topics are intended to be taught to all students at all grades.

District Standards

Arguably, teachers pay more attention to district standards than to state standards. Are they substantially different? It doesn't appear so. We have done dozens of analyses of district standards from across the U.S. In this article, we present a composite of district-level standards from one selected state.[8] Looking at this composite (Figure 3, page 11), it is clear that the districts' standards tend to include slightly fewer topics than are specified in state standards. But, like the states, the districts still specify many more topics per grade than do the A+ countries. Furthermore, the district data, like the state data, indicate a great deal of repetition of the topics across grades. Five of the 10 topics intended for coverage in all eight grades in the state composite are similarly intended for such coverage in the district composite; an additional three of the topics are intended for coverage in seven of the eight grades. Overall, then, we can see that the districts' standards are nearly as incoherent as the states' standards.

One can assume that given the broad scope of these standards, teachers are forced to cut back from what's intended

The Benefit to Professional Development

Most studies of professional development don't even consider the effect on student achievement; and most studies of educational reform that include a teacher-training component do not isolate the impact of the training. But the few studies that do examine the link between professional development and student achievement suggest this: Professional development is most effective 1) when it is focused on the content teachers must teach and how to teach it, or 2) when it is provided in concert with a curriculum and helps teachers to understand and apply that curriculum. Such professional development can raise achievement substantially.

Some evidence for this comes directly from TIMSS. Unlike the rest of the United States, eighth-graders in Minnesota attained scores that were second only to Singapore's eighth-graders in science. Intrigued, the National Education Goals Panel commissioned a case study of the state's approach to science in the seventh and eighth grades. The study found that through an "incremental but cumulative" process, a consensus was built in Minnesota about what constituted good science content and instruction in the middle grades.

By the time TIMSS was administered in 1995, the vast majority of Minnesota seventh-graders took life science and eighth-graders took earth science. There had been a large number of professional-development activities geared to these courses, and "science teachers in the middle grades were more likely to use the same or similar texts and common instructional practices." Not only was the curriculum common, it was also coherent. Unlike the typical science curriculum in the U.S. (in which large numbers of topics are introduced each year, with few covered in depth), in Minnesota "there were far fewer topics introduced and more time devoted to developing them in depth." The National Educational Goals Panel concluded that, "This research suggests the necessity of aligning teacher training, professional development, and other teacher support mechanisms with the overall reform process." (To read the Panel's full report, please visit **www.negp.gov/promprac/promprac00/promprac00.pdf**.)

Further evidence for curriculum-based professional development was reviewed by Grover Whitehurst, assistant secretary for research and improvement, U.S. Department of Education, for the White House Conference on Preparing Tomorrow's Teachers. He stated that out of seven teacher characteristics that could increase achievement (things like certification, workshop attendance, and experience), participation in professional development that is focused on academic content and curriculum was second only to a teacher's cognitive ability. In contrast, participation in typical professional-development workshops was the least effective of the seven characteristics. Summarizing the relevant research on in-service training, Whitehurst said, "when professional development is focused on academic content and curriculum that is aligned with standards-based reform, teaching practice and student achievement are likely to improve."

To illustrate his point, Whitehurst described a study of Pittsburgh schools that implemented a standards-based mathematics curriculum. The resulting differences in student achievement between the strong and weak implementers of the curriculum were dramatic. In the strong implementation schools, 74 percent of African-American students and 71 percent of white students met the established performance standard on the New Standards Mathematics Reference Exam. But in the weak implementation schools, only 30 percent of African-American students and 48 percent of white students met the standard. After pointing out that strong implementation eliminated racial differences in the outcome measure, Whitehurst explained that the impressive results were in fact due to the implementation, not differences in the teachers: "There is no reason to believe that any...individual differences in teachers..., such as cognitive ability or education, differed among the weak...versus the strong implementation schools. Yet the teachers in the strong implementation schools were dramatically more effective than teachers in in state and district standards. It's not likely that many can distill a coherent curriculum from the incoherence that's offered them. Further, teachers are likely to prune back the state/local standards in different, idiosyncratic ways. This is what leads to the well-known American phenomenon—and special bane of transient students—in which what's actually taught in a given grade varies wildly from class to class, even in the same school, district, or state.

It goes without saying that under these circumstances, a serious investment in curriculum-based professional development is not feasible; nor is it really feasible to align preservice education or texts to a non-existent curriculum. Any statewide assessment must choose between asking vague or low-level questions—or risk asking specific questions about particular content that teachers haven't taught.

Overall, Figures 2 and 3, representing composites of state and district standards, suggest that in America we tend to treat mathematics as an arbitrary collection of topics. There is no visible sense-making or structure. The math—for both students and teachers—looks and feels like a bunch of disconnected topics rather than a continuing development of the main concepts of mathematics that fit together in a structured, disciplinary way.

To complete this picture of the intended American math curriculum, we must take note of the especially huge curricular variation that becomes visible in the eighth grade, when most schools offer a variety of math courses, each with different content and rigor. In our study of eighth-grade math courses offered in American

the weak implementation schools. Thus [the]…effect of curriculum implementation swamped effects of individual differences in background among teachers." Diane Briars, head of Mathematics and Science Education for the Pittsburgh Public Schools and an author of the study, told the Committee on Education and the Workforce in the U.S. House of Representatives that in strong implementation schools "teachers were given time to meet and work together to improve their instruction."

Noting that most studies of professional development do not address student achievement, Whitehurst recommended just a handful of high-quality studies. Excerpts from two of them are provided below. (To read Dr.Whitehurst's full presentation, visit **www.ed.gov/inits/preparingteachersconference/whitehurst.html**.)

—*EDITORS*

The Case of California

By David Cohen and Heather Hill

Most reformers, including many governors, President George W. Bush, and many business officials concerned with schools, have argued that schools need to be shaped up with stronger academic standards, stiffer state tests, and accountability for students' scores. Our decade of detailed study on California's effort to improve mathematics teaching and learning shows that standards, assessments, and accountability are more likely to succeed if they are accompanied by extended opportunities for professional learning that are grounded in teachers' practice. But our study also strongly suggests that not all opportunities for teachers to learn are created equal.

The 1985 Mathematics Framework for California Public Schools was one of the first major state reforms. The goal was to provide much more academically demanding work for students. The initiative offered more detailed guidance for teaching and learning—in assessments, curricular frameworks, student curricula, and professional education—than has been commonly provided by most state governments during most of our history.

Having failed to persuade textbook publishers to produce much less conventional textbooks, in 1989 the reformers began encouraging curriculum developers to create "replacement units" on specific topics like fractions. To aid teachers further, these units were accompanied by "replacement unit workshops"—two-and-a-half-day sessions in which teachers would do the mathematics themselves, talk with each other about the content, and observe examples of student work on the materials.

These kinds of opportunities to learn seemed not only to increase teaching practices associated with the new math framework but to decrease use of conventional methods; teachers did not simply add new practices to a conventional core, but also changed that core teaching approach. This is quite significant when compared with the "Christmas tree" approach most teachers bring to their learning from professional development, in which they festoon an otherwise stable and conventional practice with attractive, new—and often inconsistent—additions.

In contrast, when teachers used their professional-development time to attend special-topics workshops, there was nearly zero association with teachers' ideas and practices (whether conventional or innovative). We suspect that this occurred because special-topics workshops were not chiefly about the mathematical content, though they were consonant with the state math frameworks in some respects.

(Continued on page 14)

David Cohen is a John Dewey Professor of Education and an Annenberg Professor of Public Policy at the University of Michigan where Heather Hill is an assistant research scientist with the university's School of Education. Cohen and Hill's material was excerpted with permission from Learning Policy: When State Education Reform Works, *©2001 by Yale University Press, New Haven, CT, 1-800-YUP-READ;* **www.yale.edu/gap.**

schools, we learned that eighth-graders tend to be enrolled in any of about six different types of mathematics courses, ranging from remedial math focused on arithmetic, to pre-algebra, algebra, and even geometry.[9] Not surprisingly, student achievement at the end of eighth grade roughly corresponded to the courses students had taken. In short, a student's achievement corresponded substantially to his or her opportunity to be exposed to more or less rigorous material.

It is probably no surprise to report another finding: that a student's opportunity to study in a higher-level math course was related to his or her geographic location. We determined that while 80 percent of eighth-graders had access to a "regular" math course, only 66.5 percent of eighth-graders attend schools that even offered an algebra course. That is, a full third of eighth-graders don't even have such a course as an option. In rural and urban settings, 60 percent of students attended schools that offered algebra and other more challenging classes. In suburban and mid-sized cities, 80 percent of students attended schools with such classes.

As with the farming ideas available from states and towns in Country B, it's not a great loss that the various state and district standards are so difficult to implement consistently, as they are of questionable quality. Like the farmers in Country B, American teachers often don't have the tools (textbooks or classroom materials) or training to make use of any wisdom they might be able to cull from the standards anyway. But without the benefit of the distilled national wisdom about mathematics education or the tools and training to go with it, American teachers are at a great disadvantage. Some get a hold of excellent curricula; some have a knack—coupled with a lot of blood and

The Benefit to Professional Development

(Continued from page 13)

Such workshops might have encouraged cooperative learning or new techniques for students who have not traditionally performed well in math rather than any change in core beliefs and practices concerning mathematics and teaching mathematics.

Our central finding is that California's effort to improve teaching and learning did meet with some success, but only in this circumstance: When California teachers had significant opportunities to learn how to improve students' learning, their practices changed appreciably and students' learning improved. The things that made a difference to changes in their practice were those things that were integral to instruction: curricular materials for teachers and students to use in class, assessments that enabled students to demonstrate their mathematical performance—and teachers to consider it—and instruction for teachers that was grounded in these curriculum materials and assessments.

The difficulty with countless efforts to change teachers' practices through professional development has been that they bore no relation to central features of the curriculum that students would study, and consequently have had no observable effect on students' learning. Many efforts to "drive" instruction by using "high-stakes" tests failed either to link the tests to the student curriculum or to offer teachers substantial opportunities to learn. These and other interventions assume that working on only one of the many elements that shape instruction will affect all the others. The evidence presented here, however, suggests that instructional improvement works best when 1) it focuses on specific academic content, 2) there is a curriculum for improving teaching that overlaps with curriculum and assessment for students, and 3) teachers have substantial opportunities to learn about the academic content, how students are likely to make sense of it, and how it can be taught.

Content Matters Most

By Mary Kennedy

The one-shot workshop is a much maligned event in education. Researchers and policy analysts have generated a number of proposals for how inservice education programs should be organized instead. Surprisingly, these reform proposals generally deal with the *structure* of the professional development, but rarely specify the *content* that inservice teacher education programs should provide. Specifically what the content should be—generic teaching techniques versus research findings on how students learn specific content, for instance—is rarely discussed.

Although the literature on inservice programs is voluminous, that volume subsides quickly when you limit yourself, as I did, to studies that include evidence of student learning and concentrate on either mathematics or science. The studies I found are organized into four groups according to the content they provide teachers. While the study addressed both mathematics and science, only the mathematics findings are presented here:

■ The two studies in group 1 prescribe a set of teaching behaviors that are expected to apply generically to all school subjects. These behaviors might include things like cooperative grouping, and the methods are expected to be equally effective across school subjects.

■ The seven studies in group 2 prescribe a set of teaching behaviors that seem generic, but are proffered as applying to mathematics. Though presented in the context of a particular subject, the behaviors themselves have a generic quality to them in that they are expected to be generally applicable in that subject.

■ The two studies in group 3 provide teachers with some theory about student learning and then move to a recomsweat—for figuring out how to teach even the most challenging students fairly well. The most effective and most affluent school districts can attract a disproportionate share of the most well-prepared teachers; plus, many of these districts provide reasonable materials and training to their faculty.

Yet most teachers, especially those working in the poorest school districts and poorest schools, cannot turn to their districts or states for much help. For most teachers, it's an ongoing, consuming challenge to dream up a basic curriculum and the daily lesson plans to execute it. Not many teachers have the additional time or resources to go beyond that to devise special, unique ways of reaching the kids in the class (or, in secondary school, in a number of classes) who aren't catching on for a wide variety of different reasons.

This lack of curriculum, materials, and training produces the same results for American students as Country B's policy produced for its crops. Curriculum really matters. Schools are supposed to provide opportunities for students to acquire the knowledge that society deems important, and structuring those learning opportunities is essential if the material is to be covered in a meaningful way. The particular topics that are presented at each grade level, the sequence in which those topics are presented, and the depth into which the teacher goes are all critical decisions surrounding the curriculum that have major implications for what children learn.

IV. The U.S. Result: Lower Achievement and Less Equity

Based on our findings of curriculum differences between A+ countries and the U.S., we can say that our students and teachers are severely hampered—both by the inadequacy of the curriculum in this country and by the loss of the benefits

mended set of teaching strategies and a recommended curriculum that is justified by that knowledge of student learning.

■ The one study in group 4 focuses on the particular mathematical content that students will learn and on the particular kinds of difficulties they are likely to have in learning this content. Teachers were not provided with a set of invariant teaching strategies, but the researchers engaged teachers in discussions about different ways of teaching different types of math problems to children.

The table below shows the average size per group of program effects on student achievement outcomes in mathematics. Groups 3 and 4 clearly had greater impacts on student achievement than did groups 1 and 2.

Average Standardized Effect Sizes in Mathematics

Group	Basic Skills	Reasoning & Problem Solving
1	–.14	.10
2	.17	.05
3	.13	.50
4	.52	.40

This pattern of outcomes suggests that the content of inservice programs does indeed make a difference and that programs that focus on subject-matter knowledge and on student learning of particular subject matter are likely to have larger positive effects on student learning than are programs that focus mainly on teaching behaviors.

These more successful programs provided knowledge that tended *not* to be purely about the subject matter—that is, they were not courses in mathematics—but instead were about *how students learn* that subject-matter. The programs in groups 3 and 4 were very specific in their focus. They did not address generic learning, but instead addressed the learning of particular mathematical ideas.

I suspect this type of program content benefits teachers in two ways. First, in order to understand how students understand particular content, teachers also have to understand the content itself so that subject-matter understanding is likely to be a by-product of any program that focuses on how students understand subject matter. Second, by focusing on how students learn subject matter, inservice programs help teachers learn both what students should be learning and how to recognize signs of learning and signs of confusion. So teachers leave these programs with very specific ideas about what the subject matter they will teach consists of, what students should be learning about that subject matter, and how to tell whether students are learning or not. This content makes the greatest difference in student learning.

Mary Kennedy is a professor in the College of Education at Michigan State University. Her material was excerpted with permission from "Form and Substance in Inservice Teacher Education," which is available online at ***www.msu.edu/~mkennedy/publications/docs/NISE/NISE.pdf****.*

that can flow from making a quality curriculum common.

We saw at the beginning of this article that the average achievement in the U.S. is low in comparison to many other countries. Moreover, the gap in students' achievement between our most- and least-advantaged schools is much greater than the comparable gap in most TIMSS countries. In fact, a recent study conducted by researchers at Boston College demonstrated that in the U.S. about 40 percent of the variation among schools in students' test scores is explained by socioeconomic factors. In comparison, across all of the TIMSS countries, socioeconomic factors explain less than 20 percent of this type of variation.[10]

We believe that America's poor average achievement, as well as our strong link between achievement and SES, can be traced in part to our lack of a common, coherent curriculum. The A+ countries have a common curriculum for virtually all students through the eighth grade. In those countries, all schools have roughly comparable access to the full array of materials, professional development, and assessments that can help teachers lead students to high achievement.

Further, students' opportunities to learn are enhanced by the benefits that accompany a common curriculum: teachers can work together with a shared language and shared goals; new teachers can receive clear guidance on what to teach; professional development may be anchored in the curriculum that teachers teach; textbooks may be more focused and go into greater depth with a smaller set of topics; and transient students (and teachers) may more easily adapt to new schools. All of this contributes to greater consistency and quality across schools.

We intend to conduct additional studies to further test the veracity of these arguments. But we would argue strongly that the weight of the evidence—and the high stakes, which include reducing the achievement gap and raising average achievement—should dissuade us from waiting around for more evidence before acting.

As we said at the outset, the practices of other nations can rarely be imported whole-cloth. Institutions and cultures differ too much. But we can learn from other nations and find ways to adapt to our own use those practices that seem particularly effective. In all likelihood, we won't adopt—certainly not in the near term—a national curriculum like the A+ countries have—after all, most of the A+ countries are small (though the largest is almost half our size).

But similar benefits could flow from adaptive arrangements that provide a common, coherent, rigorous curriculum to large groups of our students, such as adopting curriculum at the state level, or facilitating groups of states in adopting a common curriculum.

One way or another, we should be moving on a variety of fronts to bring about a more common, coherent curriculum and to let the benefits of that flow to our schools, our teachers, and especially our students—who deserve no less than the quality of education experienced by children in the A+ countries. □

How Would Your District Standards Compare?

Working with the TIMSS researchers, the North Central Regional Educational Laboratory created a Web site that allows districts to create maps of their mathematics and science standards. Just indicate which of 44 math and 79 science topics are supposed to be taught at each grade, and the site will develop the map. Then, you'll have the option of comparing your district standards to those of top-achieving countries. Visit **http://currmap.ncrel.org** to develop your map.

Endnotes

1 Schmidt, W.H., McKnight, C.C., Houang, R.T., Wang, H., Wiley, D.E., Cogan, L.S., and Wolfe, R.G. (2001). *Why Schools Matter: A Cross-National Comparison of Curriculum and Learning.* San Francisco, CA: Jossey-Bass.

2 In each of these countries there is a document outlining the content that is to be taught to virtually all children in the school system. Some students may receive additional advanced problems for specific topics. In Hong Kong, for example, textbooks may indicate Level 2 problems that teachers are encouraged to assign to their more advanced students. But the composite presented on page 14 (Figure 1) is based on the material that all students are exposed to.

3 Schmidt, W.H., Wang, H.A., and McKnight, C.C. (no date). Curriculum Coherence: An Examination of U.S. Mathematics and Science Content Standards from an International Perspective. Paper being prepared for publication.

4 To make sure that our analysis of the A+ composite did in fact apply to a complete curriculum, we developed a second composite that included all of the additional topics from the A+ countries. This complete composite confirmed that the basic three-tier structure that is discussed in the section on the A+ composite is retained even after the additional topcics are added.

5 Belgium actually has two national curricula, one for each of its two national language groups. For all practical purposes, though, a given group of teachers and students are only governed by one, so it functions like a single national curriculum.

6 A methodological note: The majority of states had grade-specific content standards. But several states specify a cluster of grades in which a topic could be taught, then leave it up to local districts to determine in which grades the topic is actually taught. For the few states that used a cluster approach, our method assumes that the topic is intended in each of the cluster grades. This seems reasonable since some data indicate that districts and textbook publishers tend to use the clusters in this fashion.

7 This holds true for *each* of the states studied—not just for the composite. When we did individual displays of each state's standards, we found that most were even more repetitive than the state composite. In addition, none of the state's standards were even remotely as coherent as the A+ composite.

8 This state volunteered for the district analysis, however the results presented here are consonant with the results from our other district studies.

9 Cogan, L.S., Schmidt, W.H., and Wiley, D.E. (2001). Who Takes What Math in Which Track? Using TIMSS to Characterize U.S. Students' Eighth-Grade Mathematics Learning Opportunities. *Educational Evaluation and Policy Analysis, 23*(4) 323-341.

[10] Martin, M.O., Mullis, I., Gregory, K.D., Hoyle, C., and Shen, C. (2000). *Effective schools in science and mathematics, IEA's Third International Mathematics and Science Study.* Chestnut Hill, MA: International Study Center, Lynch School of Education, Boston College.

Appendix: Methodology

Development of the A+ Composite

To identify the top achieving (A+) countries in mathematics, we rank ordered countries from highest to lowest using their eighth-grade score. We then compared each country's score with every other country's score to determine which ones were statistically significantly different. The following countries, which statistically outperformed at least 35 other countries, became the A+ countries: Singapore, Korea, Japan, Hong Kong, Belgium (Flemish-speaking), and the Czech Republic.*

To analyze the A+ countries' intended content, a procedure called General Topic Trace Mapping (GTTM) was used. Education officials were given extensive lists of topics in mathematics and asked to use their national curriculum to indicate for each grade level whether or not a topic was supposed to be covered. The result was a map reflecting the grade level coverage of each topic for each country. Although none of the countries' maps were identical, the A+ countries' maps all bore strong similarities.

The A+ countries' topic maps were synthesized to develop a composite of the topics intended by at least two-thirds of the A+ countries (see Figure 1, page 14). The synthesis was done in three steps. First, we determined the A+ countries' average number of intended topics at each grade level. Second, we ordered the topics at each grade level based on the percentage of the A+ countries that included a particular topic in their curriculum. For example, since all of the countries included the topic "whole number meaning" in the first grade, that topic was placed at the top of the list for first grade. Third, we used the information from steps one and two to develop the A+ composite. At each grade, the composite was to include no more than the average number of intended topics. The composite was also to include only topics that were intended by at least two-thirds of the A+ countries. Therefore, the topics intended by the greatest percentage of countries were selected for the composite first, and only as many were chosen as were indicated by the mean number of intended topics at each grade level. Therefore, the topics in the A+ composite constitute the "core curriculum." In addition to these core topics, each country taught additional topics. The number of additional topics beyond the core that are intended at each grade level can be seen in the number found in the last row in Figure 1 (see page 4).

* Valverde, G.A. & Schmidt, W.H. (2000). Greater expectations: learning from other nations in the quest for 'world-class standards' in U.S. school mathematics and science. *Journal of Curriculum Studies, 32*(5), 651-687.

Development of the U.S. Content Standards

The data on U.S. content standards in mathematics were collected from two sources: a sample of 21 states' standards and a sample of 50 districts' standards. These data indicated topics intended for instruction at each grade level through eighth grade.

Because the U.S. has so many sets of standards, using the General Topic Trace Mapping procedure would have been very difficult. Instead of using education officials' judgments about intended content, coders (graduate students with degrees in mathematics, engineering, and the various sciences) compared the actual standards documents referenced above to the same extensive list of mathematics topics that was used for the GTTM. More complex standards were identified with more than one topic as appropriate. Once the standards were coded by topic, state and district composites were developed in the same manner as the A+ composite. □

Getting There In America's Decentralized Education System

For over a decade, there's been a consensus among American leaders and the public that our schools can and should be improved based on the vision outlined in these pages: clear standards for what students should know and be able to do; a coherent curriculum that maps a route to the standards; professional development tied to the curriculum; excellent texts and materials; quality assessments; and a fair accountability system that encourages students to put forward their best effort and assures that schools get the intervention they need.

With America's traditional wariness of federal involvement in curriculum matters, however, there has also been a consensus that this vision should be achieved at the state level. But the ambition of this vision has exceeded the resource capacity of most states. Perhaps not surprisingly, most states have only gotten as far as developing student achievement standards (that are often vague) and generally inadequate assessments.* Without a curriculum and without the training materials to teach the curriculum, many teachers (and parents and students) feel that the assessments are simply a "gotcha" exercise—not an instructionally useful and valid tool. On these rough shoals, America's longest running education reform movement could founder.

If standards-based reform is to succeed in lifting student achievement, we need new ideas and structures. If the development costs for quality curriculum, training, and assessments are too great for a single state, let a number of states come together and jointly develop them. If states find it politically impossible to gain agreement on the details of a specific curriculum, perhaps we can turn to independent organizations like the International Baccalaureate described in this issue (see page 28). States could certify the curricula and assessments of these groups as being consistent with the state's vaguer standards; and schools or districts could be encouraged to adopt them and make use of their training opportunities and materials. In Virginia, for example, students who do well on an IB exam are exempt from the corresponding state exam. Likewise in Florida, students have an incentive to take the IB courses (and schools, therefore, have an incentive to offer them) because IB diploma holders receive full scholarships to state colleges.

One very promising initiative, the Mathematics Achievement Partnership (MAP), is being launched by Achieve, an organization representing the nation's governors and business leaders.

We highlight MAP as a project that's well-along and generally well conceived. We look forward to other initiatives that find ways to navigate a path from America's traditional embrace of local control of curriculum to a higher-quality, aligned educational system that students abroad enjoy and benefit from—and students here so far don't.

—EDITOR

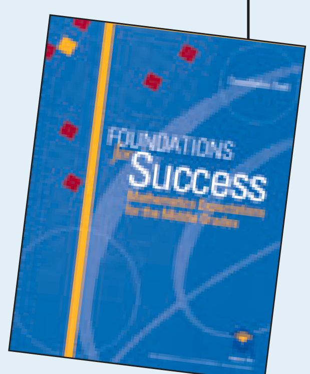
---

* For a full report see *Making Standards Matter 2001*, published by the AFT, available online at **www.aft.org/edissues/standards/MSM2001** or prepaid ($10 each; $8 for orders of five or more) from the AFT Order Department, 555 New Jersey Ave. N.W., Washington, DC 20001. Please reference item No. 39-0262.

MAP: A Promising Initiative

Achieve's Mathematics Achievement Partnership has brought together a consortium of states to jointly develop key components of standards-based reform, all focused on middle-school math and culminating with an end-of-eighth-grade assessment. Its coordinated components will include:

■ **Focused and rigorous expectations** for what students should know and be able to do at the end of eighth grade: Called *Foundations for Success*, a consultation draft of these world-class expectations is currently available at **www.achieve.org/dstore.nsf/Lookup/Foundations/$file/Foundations.pdf**. Unlike most expectations documents, *Foundations* includes sample problems that illustrate the depth of conceptual understanding that students should attain. Achieve hopes to publish a final version of these expectations in late 2003.



■ **A grade-by-grade sequence:** Also expected in 2003, this sequence will suggest what material students need to learn in sixth, seventh, and eighth grades in order to meet the *Foundations* expectations at the end of the eighth grade.

■ **Content-based professional development:** The professional development component, which enables teachers to increase their knowledge of mathematics and their skill in teaching it, is now being piloted in several districts.

■ **Diagnostic and cumulative assessments:** MAP will include diagnostic, classroom-based tests aligned to the sequence that will help teachers ensure that all students progress toward meeting the expectations. At the end of eighth grade, there will be an internationally benchmarked assessment that is aligned with the MAP expectations.

As noted in these pages, a curriculum with grade-by-grade specifics, including teaching ideas, is an indispensable element for designing effective professional development, classroom materials, and assessments—and for assuring that all these pieces are aligned with each other. We hope that as MAP's grade-by-grade sequence takes shape, it will include the specifics that will make such alignment possible and give teachers the guidance they need and deserve.

To learn more about MAP, visit ***www.achieve.org/achieve.nsf/MAP?OpenForm***.