DRAFT

STEM REFORM: WHICH WAY TO GO[1]

William H. Schmidt[2]
Education Policy Center
Michigan State University

In this paper, I address the issue of STEM related reform for non-STEM K-8 schools. The National Research Council (NRC) was charged with identifying highly successful strategies, practices and schools for STEM education. The majority of students learn STEM subjects in schools that do not have a specific STEM focus. These schools are continually looking to improve or reform their STEM education. Various models for reforming and improving STEM education have been developed across the country. This paper provides greater insight into the types of STEM education reform approaches found in non-STEM focused schools. The focus is on mathematics, as it serves as the language for science, engineering and technology and is the area in which the most reform has taken place. It is also impossible in a short paper to cover all four of these areas in depth. Some specific comments related to science will be incorporated at various points in the paper.

WHY REFORM IS NECESSARY

A realistic appraisal of NAEP data gives grounds for alarm, despite incremental gains in student performance. Since 1996, NAEP scale scores have increased by a statistically significant 16 points in fourth grade (from 224 to 240) and 12 points in eighth grade (271 to 283). Similarly, the percentage of students who are proficient in mathematics has increased from 19 to 33 percent in grade 4 and from 20 to 26 percent in grade 8. The progress of twelfth graders, who are tested

[1] A longer, less specific to STEM reform version of this paper is being considered for journal publication.
[2] Nathan Burroughs was instrumental in the preparation of this paper.

less frequently, is much less impressive with scale scores that increased from 150 to 153 since 2005 and the percent proficient rising from 21 to 23 percent (see Figure 1).

In spite of the 14 point increase in the percent of fourth graders that are deemed proficient – not advanced, only proficient, which is, what we would hope all students would be – it should be emphasized that about three quarters of U.S. eighth graders enter high school *not* proficient and therefore *not* prepared to move to this next level where the level of the mathematics is much more demanding. Therefore, it is not surprising that roughly *three quarters* of American high school students graduate with a relatively poor grasp of mathematics at a time when federal law mandates 100 percent proficiency within the next four years (albeit with much weaker state standards). Even the brightest students do not fare well, with only a tiny percentage of students reaching the advanced level on the NAEP: 6 percent in grade 4, 8 percent in grade 8, and a vanishingly small 3 percent in grade 12. These trends are made all the more discouraging by the fact that gains have largely stalled in the last several years. Between 2007 and 2009 scale scores were unchanged in fourth grade mathematics while the percent proficient actually declined. The gains in grade 8 were marginal at best.

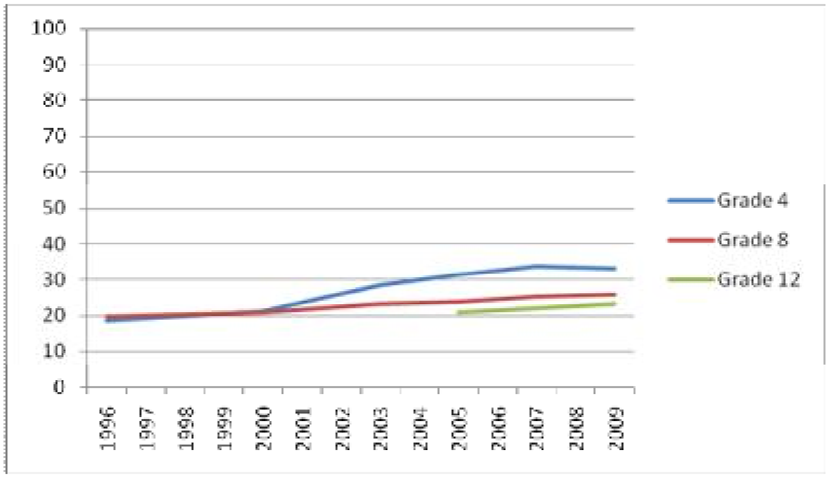

Figure 1. Trends in percent proficient in mathematics on the NAEP, 1996-2009.

The failure of the vast majority of students to achieve mathematics proficiency is more than a matter of not living up to some arbitrary standard of academic excellence. In an increasingly interdependent international economy, American students are competing both with their peers within the United States and millions more in other countries, making comparisons with students in other countries essential. International assessments of academic achievement provide a useful benchmark.

U.S. students mathematics skills develop compared with those in other countries, by falling from rough parity in the early grades until they lag badly at graduation. Whether they complete college or not, many lack the basic quantitative skills required to compete in the international economy. The reality is that the U.S. educational system is fundamentally failing in its task to prepare today's students for the future. Even our best schools and our best students do not perform at high levels by international standards. Looking at our best schools to determine best approaches at reform may not be adequate. We must look beyond our own system. Nearly thirty years after the publication of *A Nation at Risk* and after a decade of accountability reforms, the mediocre performance of American students in mathematics is essentially unchanged in a world that is changing exceptionally rapidly. The results for science paint a similar picture.

In the rest of this paper we will discuss the types of strategies that we believe are needed to guide mathematics education reform as we move forward. We first discuss the related research findings as to how these elements of the system have contributed to the problem detailed in the previous section. Our aim, however, is not to assign blame but to identify potential solutions through these research findings that will help American schools live up to their promise. In the following sections we address five elements that we believe constitute the essential part of the U.S. educational reform agenda in mathematics education: 1) curriculum, including both intended coverage as defined by standards and the mathematics content actually delivered to students by the teachers; 2) teacher knowledge of mathematics content and pedagogical content knowledge by which content is delivered; 3) public expectations of and support for demanding content standards and course requirements; 4) student motivation which influences student engagement in STEM areas; and 5) instructional leadership .

CONTENT DRIVEN STRATEGIES

Standards

One way that most of the developed nations assure the quality of mathematics instruction is by providing uniform national standards that define in detail the mathematics content that is to be covered by teachers at each grade level. This has and continues to serve as the basis for content driven reform. The new Common Core State Standards in Mathematics and the current efforts to develop new science standards only amplify these content driven strategies for reform. Formal content standards provide direction to teachers by defining the content that is to be taught at each grade level, in what sequence, and to what depth. National content standards serve as a national map of content coverage that helps to structure mathematics instruction. As a result curriculum materials (e.g., textbooks, tests), teacher training and professional development are all developed with respect to those standards. In this way the standards serve as the focal point of the educational system and have been employed in various states as the trigger for school reform. Research results from the curriculum analyses of the country standards for almost fifty countries that participated in the 1995 TIMSS indicated that there was substantial variation in both mathematics and science content standards among countries (Schmidt, McKnight, Valverde, Houang, & Wiley, 1997; Schmidt, Raizen, Britton, Bianchi, & Wolfe, 1997).

Most industrialized nations possess content standards that are national in scope (Schmidt, Cogan, Houang, & McKnight, 2009). A few nations follow the U.S. pattern of using regional content standards, with sub-governments possessing a fair degree of latitude in setting standards. Although it is certainly possible *in theory* for regional standards to match (or even exceed) the quality of national standards, the standards of American states generally fail to do so. Based on a comparison of the standards used by the TIMSS highest-performing nations at eighth grade (A+ countries) with those of twenty-one U.S. states, Schmidt, Wang, and McKnight (2005) found that state standards in both mathematics and science were substantially different in three important dimensions: focus, coherence, and rigor.

Curricular focus has to do with the number of content topics included at a given grade level. Coverage of a larger number of topics generally requires that there is less instructional time for any one topic, and hence less focus. At the time of the study and even more recently, U.S. standards tended to have much less focus, with more topics covered in each grade than is the case for A+ countries.

Rigor is defined by the level of complexity of the content at each grade level. It is essentially a measure of how demanding the mathematics content is. In A+ countries, students at the eighth grade are expected to study algebra, geometry, physics and chemistry. In the U.S. the vast majority of eighth-graders (variously estimated to be 70 to 80 percent) are expected to continue to study much more elementary topics like fractions, percents, ratio and proportionality and in science: life, earth or physical science. However, the low level of rigor in U.S. standards has begun to improve in the last five to ten years, with many states now requiring algebra. This trend is likely to be consolidated with the development of the Common Core State Standards in Mathematics, which we will discuss further below.

The third characteristic distinguishing U.S. state standards from those of the highest-achieving nations is curricular coherence. We define coherence as:

> … a central defining element of high-quality standards. If one of the major purposes of schooling is to help students develop an understanding of the various subject-matters deemed important by a society, such as mathematics and science, then the definition of 'understanding' is important to examine as a way of viewing each discipline intended for schooling
>
> Bruner (1995:333) suggests that:
>
> > …to understand something well is to sense wherein it is simple, wherein it is an instance of a simpler, general case. …In the main, however, to understand something is to sense the simpler structure that underlies a range of instances, and this is notably true in mathematics.
>
> Bruner's definition implies that the logic of the content in a discipline is important and that, for example, the goal of helping students understand mathematics is facilitated by making visible to them an emerging and progressive sense of its inherent structure. Bruner describes this as:
>
> > …opt[ing] for depth and continuity in our teaching rather than coverage … to give … [the student] the experience of going from a primitive and weak grasp of some subject to a stage in which he has a more refined and powerful grasp of it (p. 334)
>
> We define content standards, in the aggregate, to be coherent if they are articulated over time as a sequence of topics and performances consistent with the logical and, if appropriate, hierarchical nature of the disciplinary content from which the subject-matter derives. This is not to suggest the existence of a single coherent sequence, only that such a sequence reflect the inherent structure of the discipline. This implies that, for a set of content standards 'to be coherent', they must evolve from particulars (e.g. simple mathematics facts and routine computational procedures associated with whole numbers and fractions) to deeper structures. It is these deeper structures by which the particulars are connected (such as an understanding of the rational number system and its properties). This evolution should occur both over time within a particular grade level and as the student progresses across grades. (Schmidt et al., 2005).

This characteristic is the most important of the three.

The lack of focused, coherent, and rigorous content standards has real consequences. What follows focuses on mathematics. A cross-national comparative set of analyses by Schmidt and Houang (2007) examined the effects of coherence and focus on mathematics achievement.

Using 1995 TIMSS data, the researchers found that curricular coherence and focus had a statistically significant relationship with student outcomes. The greater the degree of alignment of a country's standards with the A+ model of coherence, the greater the predicted achievement in seventh and eighth grade ($p<.02$). The relationship was weaker in earlier grades; marginally so at fourth grade ($p < .08$) but not at third grade. Similarly, greater focus was also related to achievement, with statistical significance in third, fourth, seventh, and eighth grades. The explanatory power of focus and coherence was fairly robust, accounting for between 22 and 26 percent of the variance in student achievement across countries. Effect sizes ranged from one fourth to three fourths of a standard deviation at eighth grade with the predicted effects generally greater at the earlier grades.

Aside from the dimensions of rigor, coherence, and focus, cross-national analysis of the actual content that students are expected to learn indicates a strong relationship between intended curriculum and student achievement (Schmidt et al., 2001). As stipulated by standards, curricular expectations are a statistically significant predictor of mathematics achievement in the eighth grade. The effects of curriculum are more than simply cumulative; their importance are underscored by the fact that intended curriculum at eighth grade has a particular influence on the cohort *gains* in student achievement between the seventh and eighth grade. Although Schmidt et al. (2001) did not find a relationship between intended content and achievement within the United States, this is not because curriculum has no effect on student learning, but rather that the U.S. content standards did not vary enough between states to register any statistical effect. Instead it was the content of textbooks that served as a sort of de facto curriculum which was related to academic performance.

It is important to note that the research described above matches content coverage to student achievement by topic areas. The analysis specifically related the content coverage in specific content topics (as defined by standards) to gains in student achievement on subtests of those same topics. The analysis clearly demonstrates a link between intended curricular coverage and achievement gains in the same topic area. This result reinforces the proposed relationship between what standards indicate students should learn (i.e., intended Opportunity to Learn) and student achievement – a relationship that is now viewed by many as an established fact (Suter, 2000; Floden, 2002; Schmidt & Maier, 2009).

Implemented Content Coverage

Content standards wouldn't mean a great deal if they had no relationship to what was actually taught in the classroom, but the cross-country analyses cited in the previous section also provided evidence suggesting that content standards are generally implemented by schools and teachers. As measured by either teacher coverage of a specific topic or by the amount of instructional time allocated to that topic, cross-country analyses have shown that the content countries emphasized in their standards is the content to which students were exposed (Schmidt et al. 2001).

This process of implementing standards is not perfectly clear cut, however. Teachers remain a critical mediator in the transmission of content. Content standards do not just have a direct effect on student achievement but operate through other factors (such as textbooks) and by providing the guidelines by which teachers instruct their students. Content standards therefore work indirectly through teacher coverage of the content. A country's implemented curriculum coverage can be measured in two different ways: the percentage of teachers who covered a given topic or the average amount of time devoted to that topic. The effect of content coverage on both achievement and achievement gains was discernable both across countries and across classrooms within the United States (Schmidt et al., 2001). While both methods of assessing curriculum coverage suggested a positive relationship with gains in student achievement, the effects of instructional time were somewhat muted ($p<.126$) (Schmidt et al., 2001).

There is a great deal of variation in the relationship between curriculum coverage and achievement based on which topic is being considered. For some topics coverage doesn't seem to matter – achievement gains are much more related to content standards or textbook coverage. For other topics it matters a great deal. For example, student learning about the geometric topic of congruence and similarity are very strongly related to teacher coverage defined as either coverage or instructional time. Japan's eighth grade students exhibited this relationship especially well (Schmidt, McKnight, Cogan, Jakwerth, & Houang 1999). They didn't have the highest score in the 1995 TIMSS, but they did have the highest score on congruence and similarity. It would be an extraordinary coincidence if this success had nothing to do with the fact that they also spent more instructional time on that topic than any other country.

The evidence suggests that the more curriculum content coverage of a topic area, the larger the achievement gains in that topic area, whether content coverage is measured as emphasis in content standards, teacher coverage of a topic, or the amount of time allocated to the coverage of the topic. The curricular priority of a country as determined by content standards or by teacher content coverage was related to the profile of achievement gains across topics for that country. The specific nature of these relationships was not the same for all countries. When coupled with the results of other researchers (Gamoran, Porter, Smithson, & White, 1997;

Rowan, Correnti, & Miller, 2002), these studies make a strong case for the link between achievement and Opportunity to Learn (OTL).

Content standards therefore play a multi-faceted role in the learning process, both directly by shaping the content of the curriculum and indirectly by influencing the structure of that curriculum and the behavior of teachers. Weak state standards could therefore bear some responsibility for the mediocrity of mathematics learning in the United States. A comparison of state and international standards by Schmidt et al. (1999) demonstrated that state standards generally lacked the focus, rigor, and coherence of high-achieving countries. Teacher content coverage in the U.S., especially in the eighth grade, also fell far below that of the best-performing nations. Analyses both of the standards and what was actually taught indicated that eighth grade content coverage in the U.S. fell behind most of Europe and Asia. The typical eighth-grader in the U.S. was one to two years behind her peers in the OECD. While lower secondary (middle school) students in most nations were exposed to algebra and geometry, roughly three fifths of U.S. students[1] were focused on simple arithmetic topics like fractions, decimals, percentages, ratios, and proportions – topics that students in other nations had already moved well beyond.

Has anything changed in the last ten years? Perhaps. The reforms of the last decade have probably expanded the proportion of middle school students who study algebra. However, there is reason to worry that there are still vast inequalities in the rigor of content coverage. Because of the phenomenon of tracking, many students are placed on learning paths where they are guaranteed to learn less than other students, with little opportunity to close the gap. Despite long-standing criticism, tracking still remains a fairly common practice in the U.S. In other countries it remains relatively rare in earlier grades (1-8).

In TIMSS, only a quarter of U.S. students attended a school without some form of tracking and although somewhat dated, there is scant evidence to suggest this has changed (Schmidt, 2009). That same study found that middle school students in the same school were being placed in as many as six different types of courses with labels such as regular mathematics, pre-Algebra, and Algebra. This has the effect of increasing variability across U.S. classrooms in terms of content coverage thereby contributing to the large variation in achievement across classrooms and the relationship of OTL to achievement found in TIMSS at the classroom level within the U.S. In fact the proportion of the total variation across U.S. classrooms in tracked schools at eighth grade in the rigor of the content coverage related to tracking is estimated as 40 percent (Schmidt, 2009).

Schools without tracking have approximately half the variation in content coverage as schools that have tracking. Tracking does provide some benefit to students who are placed in the most rigorous courses, with eighth grade algebra scores that are higher than algebra scores at non-tracked schools; but students placed in regular mathematics courses at non-tracked schools outperform those at tracked schools.

The discussion of tracking is relevant to the overall problem of mathematics achievement in two ways. First, it reveals that there is a large variation in content coverage in eighth grade classrooms in the U.S., a variation built into the middle school structure. The variation in content coverage means that some students in the U.S. are exposed to content that is comparable to that of eighth graders in other countries, but that overall U.S. performance is reduced because many other students are taking the international equivalent of a sixth or seventh grade mathematics course.

Second, it provides part of the explanation as to why the majority of U.S. students receive less rigorous content coverage than those in other nations. The differentiation in content coverage exists because the educational system was designed to provide variable coverage. It is a deliberate policy – a policy that can be changed.

There are a number of substantive steps that educational leaders can take to improve student learning. To date much of the debate about how to improve schools has focused on strictly structural questions like charter schools and pay for performance that do not have a direct impact on what students learn (although they may have an indirect effect). Setting aside the utility of these proposals, addressing content standards and content coverage provides a very straightforward form of intervention, one that holds considerable promise. Changing the substantive content schools convey to students by improving the rigor, focus, and coherence of intended coverage (standards) and implemented coverage can have a significant impact on mathematics achievement.

The beginnings of this process may already be underway. In the summer and fall of 2010, forty-four states and the District of Columbia adopted a new set of mathematics standards called the Common Core. The Common Core moves the U.S. much closer to a system of national standards and represents a fairly dramatic break with past practices. If properly implemented, it could substantially reduce cross-state variation in intended content coverage and expose all students in those states to internationally benchmarked standards. Since the Common Core embodies standards for all students (even in middle grades), it has the potential to affect the entire distribution of students, particularly those currently tracked into less demanding mathematics classes, thereby likely improving overall mathematics achievement on international assessments.

A good example of this potential can be found by examining the curriculum of two states that participated in the 2007 TIMSS as test "countries" and hence can be directly compared to other countries – Massachusetts and Minnesota (other such state comparisons are achieved through psychometric equating not actually participating in the study). These two states did fairly well on the most recent TIMSS, outperforming the U.S. average and posting very competitive scores internationally (SciMathMN, 2008; Olson, Martin, & Mulllis, 2008). These two states have also placed near the top on NAEP, Massachusetts ranking number one in both fourth and eighth grade and Minnesota ranking third and second, respectively. These states have also received moderately favorable ratings for their mathematics standards (Carmichael, Martino, Porter-Magee, & Wilson, 2010).

The case of Minnesota is particularly informative. The state participated in the TIMSS in both 1995 and 2007, but in contrast to its excellent 2007 results it ranked near the U.S. average in 1995, placing it in the lower middle of the distribution of participants. At that time, Minnesota did not have state standards in place but was in the process of adopting them. Specifically, TIMSS 1995 teacher data reporting on classroom coverage demonstrated a lack of focus and coherence. In 1995 Minnesota teachers in the fourth and eighth grade were covering topics at grade levels that were not covered in A+ countries, i.e., Minnesota was an example of the "laundry list" approach to curriculum (Schmidt, McKnight, & Raizen, 1996).

By 2007 Minnesota had thoroughly revamped its content standards, reducing its teacher coverage of topics that were inconsistent with the international model of coherence from about half to less than five percent. Although there is no conclusive causal evidence that Minnesota's gains between 1995 and 2007 were primarily due to changes in its standards, the data do support the hypothesis that there is a relationship between standards and achievement – that content coverage led by coherent, focused, and rigorous standards properly implemented by teachers can improve student outcomes in mathematics. Most importantly, this improvement can happen in an American state. As a species of national standards, the Common Core has the potential of applying the lessons of Minnesota's accomplishment nationwide. But if the Fordham Institute's recent analysis of state standards is correct, most states still have quite a long way to go if they are to meet the requirements of the Common Core (Carmichael et al., 2010).

Improving overall student performance by means of a strengthened curriculum is a critical task, but equally critical is ensuring greater equality in learning opportunities. We have already discussed the variation in content standards between states. A concern for educational equality requires that we also examine variation in content coverage within states, even within districts, and schools. Data from Promoting Rigorous Outcomes in Mathematics and Science Education (PROM/SE) shows students receive wildly different content coverage than their schoolmates at the same grade level, even within the same school, both in terms of the amount of time allocated to content coverage and the sequence of such coverage (see Schmidt & McKnight, in press). The resulting inequalities are not always randomly distributed. A student's opportunity to learn mathematics is greatly influenced by where he lives, his race, his family's income, and by the particular teachers he happens to have.

Such inequalities in content coverage affect the entire system not just minority or low-income students, making a mockery of the idea of equal opportunity for all. Two students of similar backgrounds in the same district and school are not assured equal content coverage. No matter what other policies are put into place, it is implausible to think that students will receive an adequate education in mathematics if they do not have the chance to learn the same important mathematics. Further, these inequalities tend to compound over time. Rather than a system in which struggling students receive concentrated attention so that they have a chance to catch up, they are effectively locked into a lower tier, a strata with lesser educational and hence fewer lifetime opportunities. These are differentiations which sometimes may have nothing to do with a student's native ability or work ethic. They are in effect a kind of educational caste system.

TEACHER RELATED STRATEGIES

Content knowledge does not just happen; it is not simply absorbed by students who sit in its presence. Instead, it is presented by teachers who must use their mathematics and pedagogical knowledge to convey that content to the students in an intelligible fashion and in such a way that it creates meaningful learning experiences for students. Teachers represent a second major factor affecting student performance: the ability of a teacher to create instructional experiences in which student learning takes place. This ability has been described as either teacher quality or (in recent literature) teacher effectiveness.

Teacher quality has become a major focus of attention, both by researchers and policymakers. Recent work has argued that teacher quality is a powerful determinant of student outcomes, especially in mathematics, as well as a significant limiting factor for U.S. schools (Clotfelter, Ladd, & Vigdor, 2007; Rivkin, Hanushek, & Kain, 2005; Wayne & Youngs, 2003). These concerns have sparked federal efforts such as No Child Left Behind (NCLB), and local efforts in districts like Baltimore, Denver, and Washington, D.C. to improve teacher quality. Federal legislation now requires that all states define "qualified" teachers and to have such teachers constitute a large percentage of the teaching force. A glaring problem with this approach is that, as with content standards, there is both tremendous variation among states in what a "qualified" teacher is as well as expectations that are far too low. The type of mathematics knowledge a teacher should possess has been an especially vibrant area of debate (Begle, 1979; Begle & Geeslin, 1972; Darling-Hammond, 1999; Darling-Hammond & Youngs, 2002; Good & Grouws, 1987; Zumwalt & Craig, 2005).

In this paper we chose not to address the more general as well as the more difficult concept of what makes a teacher "effective." This broader notion is a much more complex issue as it is often coupled with notions of student achievement or achievement gains. This then involves complex statistical and psychometric issues about which there is much discussion and disagreement (National Research Council, 2010). We limit our focus more narrowly to the concept of a teacher's professional competence and specifically on the professional knowledge of teachers that is used to successfully solve core, job-related problems i.e. the teaching of mathematics. Professional competence as defined by Weinert (1999) also includes motivational and volitional predispositions which we do not address here. There is wide body of literature dealing with professional competence generally and teacher professional competence in particular (Weinert 1999, 2001; Bromme, 1992; Blömeke, 2005).

Professional knowledge for mathematics teachers is defined as mathematics content knowledge and pedagogical knowledge. Pedagogical knowledge includes both general issues of classroom and instructional organization (including such issues as motivation, classroom management and lesson planning) and issues that are tailored to the teaching of mathematics such as the psychology of learning mathematics, methods of teaching algebra and more practical issues such as probing student understanding of mathematics. These areas constitute the professional knowledge teachers need to have in order to undertake the typical activities of instruction and assessment.

As a practical matter it is quite difficult to directly test teachers' level of mathematics knowledge due to political sensitivities. However, there are still a variety of means to approach the problem. One method is to ask teachers how well prepared they feel they are to teach specific mathematics topics. For example, a 2006 study based on the PROM/SE project surveyed over 5000 teachers from 60 school districts about 24 mathematics topics, asking:

> "how well prepared academically do you feel you are – that is you feel you have the necessary disciplinary coursework and understanding – to teach each of the following at the grade level you are currently teaching (PROM/SE, 2006)…"

The list of topics varied based on the grade level of the teacher in question. Teachers were grouped into three broad categories: elementary (1-5), middle (6-8) and high school (9-12). The results of the survey are summarized below.

Although PROM/SE was restricted to only two states (Michigan and Ohio), the demographic and achievement characteristics reflect those of the overall U.S. student population. As such, if the results of the study are generalizable to the general population of elementary teachers in the 42 states that have subscribed to the Common Core as well as to the nation as a whole, it suggests that implementation of the new standards will be difficult without some form of professional development.

The majority of ***elementary teachers*** reported that they felt very well prepared to teach most of the topics that were a part of the elementary school curriculum that they taught. This result is cold comfort, however, since in grades 1-5 there were only two topics (whole number meaning/place value and operations and their properties) in which more than 75 percent expressed confidence in their preparation to teach. For teachers only at grades four and five, there was one additional topic – common fractions – that met the 75 percent threshold.

Equally alarming, elementary school teachers demonstrated a lack of confidence in topics that would be covered in later grades. This is a disturbing result, given the relationship of advanced to more basic mathematics. Teachers should have sufficient grounding in certain mathematics taught at grade levels above their own grade since it serves as the necessary background for what they are presently teaching, but the survey results suggested that for many teachers this is simply not the case. A case in point: less than a quarter of elementary teachers stated that they felt well prepared to teach proportionality concepts. However, proportionality is related to the understanding of fractions – an elementary school topic.

These elementary school teachers clearly recognized the limits of their knowledge. Although these were self-evaluations in which teachers stated their own sense of competence, in all likelihood these results were if anything *inflated*; it is quite probable that many teachers had an exaggerated degree of confidence in their preparation and competence. For basic topics such as whole number computations and fractions there are likely gaps in teacher knowledge. Even if the results are taken at face value, elementary teachers do not express strong confidence in their ability to teach the topics that are included in the Common Core and which are fundamental to enhanced student achievement in mathematics.

Middle school teachers (grades 6-8) had similar results. Faced with a list of more advanced mathematics topics, the number of topics for which at least three quarters of teachers claimed to feel very well prepared to teach was *zero*. Algebra plays an important part in the Common Core standards for middle grades, yet only half of PROM/SE middle school teachers stated that they were very well prepared to teach linear equations and inequalities. An even smaller proportion felt well prepared to teach algebraic related concepts like proportionality (41%), slope (38%), and functions (39%).

We should expect that all middle school teachers will feel capable of teaching the full range of K-8 topics, and probably many high school topics as well. This quite modest goal is regrettably very far out of reach, which raises serious questions about the ability of states to live up to the Common Core standards to which they have subscribed. Again, as with elementary school students, the similarity between the socioeconomic and ethnic profile of the PROM/SE students as those students from other states suggests that the problem is likely national in scope.

Teacher Preparation

The results of the PROM/SE survey of teacher confidence are instructive. Elementary and middle school teachers appear quite similar, with both sets of teachers demonstrating lower levels of confidence to teach specific mathematics topics than high school teachers. One possible explanation is the very different training received by K-8 as opposed to high school teachers. There is considerable variation in the proportion of elementary school teachers that have majors or minors in mathematics, ranging from almost none in the first grade to 65 percent in grade 8. Teachers are telling us that they have not been adequately prepared to teach mathematics and we should listen to them.

A 2010 study of teacher preparation (Teacher Education and Development Study in Mathematics – TEDS-M) was conducted in 16 countries to evaluate the amount of mathematics training primary and lower secondary school teachers received (Center for Research in Mathematics Education, 2010; Schmidt, Cogan, & Houang, in press). In the U.S., 81 public and private universities and colleges were randomly selected to participate in an examination of teachers who were about to enter the teaching workforce. Nearly 3300 potential future teachers participated in the U.S.-based study and 23,000 future teachers around the world did so.

Participants in the study were tested on their mathematics content knowledge and pedagogical content knowledge.

Country	Mn	(se)
Taiwan	623	(4.2)
Singapore	590	(3.1)
Switzerland	543	(1.9)
Russian Federation	535	(9.9)
Thailand	528	(2.3)
United States-Private	527	(3.6)
Norway	519	(2.6)
United States-Public	518	(4.1)
Germany	510	(2.7)
Poland	490	(2.2)
Malaysia	488	(1.8)
Spain	481	(2.6)
Botswana	441	(5.9)
Philippines	440	(7.6)
Chile	413	(2.1)
Georgia	345	(3.9)

Significantly above US-Public

Not significantly different from US-Public

Significantly below US-Public

TEDS-M Countries' Overall Performance with Respect to Mathematics Content Knowledge at the Primary Level. (From Center for Research in Mathematics and Science Education 2010)

Country	Mn	(se)
Singapore	593	(3.4)
Taiwan	592	(2.3)
Norway	545	(2.4)
United States-Private	545	(3.1)
United States-Public	544	(2.5)
Switzerland	537	(1.6)
Russian Federation	512	(8.1)
Thailand	506	(2.3)
Malaysia	503	(3.1)
Germany	502	(4.0)
Spain	492	(2.2)
Poland	478	(1.8)
Philippines	457	(9.7)
Botswana	448	(8.8)
Chile	425	(3.7)
Georgia	345	(4.9)

Significantly above US-Public

Not significantly different from US-Public

Significantly below US-Public

TEDS-M Countries' Overall Performance with Respect to Pedagogical Content Knowledge at the Primary Level. (From Center for Research in Mathematics and Science Education 2010)

Figure 2. TEDS-M countries' overall performance with respect to mathematics content knowledge and pedagogical content knowledge at the primary level.

Like the TIMSS results, future elementary teachers in the U.S. ranked somewhere in the middle of the international distribution in overall mathematics knowledge, with a performance that was comparable to that of Germany, Russia, and Norway but well behind the high-achieving countries of Taiwan, Singapore, and Switzerland (See Figure 2). Future elementary teachers performed somewhat better on questions about pedagogical content knowledge, although still trailing teachers from the top nations. The performance of U.S. future teachers could best be described as mediocre - certainly not the basis for a major improvement in mathematics performance by American students and what is needed to implement the new Common Core State Standards.

The story for future middle school teachers is even more troublesome. As found in the PROM/SE study of current middle school teachers in two states, there were no topics in which 75 percent of teachers expressed full confidence in their own preparation. TEDS-M also found serious weaknesses in the preparation of middle school mathematics teachers. Particularly alarming is that we would expect recent future teachers to do somewhat better, as they are still taking mathematics coursework. Additionally they are more recently trained than those already in practice and given No Child Left Behind (NCLB) and other recent reforms one might expect that the standards for mathematics teacher preparation would have become more demanding.

U.S. future middle school teachers, both public and private, found themselves on mathematics content knowledge, in the middle of the international distribution dividing the TEDS-M countries into two distinct groups, those countries whose middle school students do better than the U.S. on international tests and those who don’t – the only exception being Malaysia (See Figure 3).

Country	Mn	(se)
Taiwan	667	(3.9)
Russian Federation	594	(12.8)
Singapore	570	(2.8)
Poland	540	(3.1)
Switzerland	531	(3.7)
Germany	519	(3.6)
United States-Private	512	(16.3)
United States-Public	505	(9.7)
Malaysia	493	(2.4)
Thailand	479	(1.6)
Oman	472	(2.4)
Norway	444	(2.3)
Philippines	442	(4.6)
Botswana	441	(5.3)
Georgia	424	(8.9)
Chile	354	(2.5)

Significantly above US-Public
Not significantly different from US-Public
Significantly below US-Public

TEDS-M Countries' Overall Performance with Respect to Mathematics Content Knowledge at the Lower Secondary Level. (From Center for Research in Mathematics and Science Education 2010)

Country	Mn	(se)
Taiwan	649	(5.2)
Russian Federation	566	(10.1)
Singapore	553	(4.7)
Switzerland	549	(5.9)
Germany	540	(5.1)
Poland	524	(4.2)
United States-Private	505	(13.0)
United States-Public	502	(8.7)
Thailand	476	(2.5)
Oman	474	(3.8)
Malaysia	472	(3.3)
Norway	463	(3.4)
Philippines	450	(4.7)
Georgia	443	(9.6)
Botswana	425	(8.2)
Chile	394	(3.8)

Significantly above US-Public
Not significantly different from US-Public
Significantly below US-Public

TEDS-M Countries' Overall Performance with Respect to Mathematics Content Knowledge at the Lower Secondary Level. (From Center for Research in Mathematics and Science Education 2010)

Figure 3. TEDS-M countries' overall performance with respect to mathematics content knowledge and pedagogical content knowledge at the middle school level.

Unlike in elementary school there were major differences in the coursework taken by U.S. future teachers and those in better-performing countries. In Taiwan and Russia, nearly half (48.5%) of all coursework concentrated on formal mathematics, leaving 30 percent for mathematics pedagogy and 20 percent for general pedagogy. Even within the general category of pedagogy, sixty percent was specifically related to mathematics. The difference with the U.S. is stark. Only 38 percent of the courses in the U.S. were mathematics content classes. Like Taiwan and Russia 30 percent was allocated to mathematics pedagogy, the difference was that the greater focus was placed on general pedagogy (30%). Given the usually unchallenging nature of the U.S. middle school mathematics curriculum it should not be a great surprise that future middle school teachers did not spend a substantially greater amount of time taking mathematics content than did future elementary teachers (38% vs. 33%).

Teacher Knowledge and the Common Core

The challenge of providing knowledgeable teachers in every classroom is made all the greater by the adoption of the Common Core. The Common Core is rigorous, focused and demanding, especially at the middle school level. This is all to the good, but the lack of attention to implementation could prove disastrous. The mismatch between teacher preparation and the expectations mandated by the Common Core has been described as a “perfect storm” (Center for Research on Mathematics and Science Education, 2010), and for good reason. Forty-four states will now require middle school teachers to teach topics that ***they are not adequately prepared to teach.*** Their inability to do so may well have less to do with their motivation or how hard they work, than with the fact that they do not have an adequate preparation to teach these subjects. To put it simply, unless a concerted effort is made to improve teachers’ mathematics content and pedagogical knowledge – both for current and incoming teachers – the Common Core will not take hold. This necessitates better teacher preparation and serious content-driven professional development (PD) – the latter designed around the Common Core topics delivered over the full school year. The literature is clear as to the type of PD needed (Ball, Hill & Bass, 2005; Wu, 2008) and the means by which it is best delivered (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007).

PUBLIC AND PARENTAL SUPPORT

Education is performed within a cultural context, and educational systems and practices are inevitably shaped by that context. The degree to which educational systems and practices are determined by national culture have been the subject of heated debate, especially about the extent to which instructional practices can be borrowed from one country and successfully applied in another (LeTendre, Baker, Akiba, Goesling, & Wiseman, 2001). What is relevant in the present situation is the relevance of educational *expectations*, which can have an effect both at the individual level (Davis-Kean, 2005; Goyette & Xie, 1999), as well as in the aggregate. Public attitudes can have a profound effect. For example, if there is a widespread belief that some students "can't do mathematics" or that mathematics is not important, educational standards may be watered down. If it is believed that algebra is too difficult for all children to be expected to learn it in the eighth grade, then policy may be shaped in light of this belief. The results would be different content standards for certain groups of students – one for those who "can" learn algebra and another (weaker) curriculum for those who "can't." This attitude probably contributed to the policy of tracking students. In education, as with other matters of public import, popular support is critical.

Public opinion surveys suggest that the public does indeed support stronger mathematics standards. Over 90 percent of parents participating in the PROM/SE study saw a need for mathematics courses in every year of school (unpublished PROM/SE data). Strong support for continuous mathematics education was registered by surveys of Michigan parents, growing from 70 percent in 1994 to 80 percent in 2009. When asked exactly which courses should be required, Algebra I had the most support (92%). According to 2009 Education Next survey, 72 percent supported common standards (Howell, Peterson, & West, 2009). Finally, mathematics has been rated as "extremely important" and "more in need of improvement" by parents, more than any other subject (Harrison Group, 2009). These results suggest broad public support for a more rigorous mathematics curriculum.

In general there is considerable anxiety about the state of education in the United States, and a degree of support for more rigorous mathematics standards. Although parents recognize the importance of mathematics, they, too, are trapped in the vicious loop of weak mathematics skills begetting weak mathematics skills. Parents have expressed frustration at their lack of ability to assist their children in mathematics – so much so that they feel more comfortable talking to their children about illegal drugs than about mathematics (Intel, 2009). An alarming 34 percent of parents would accept a child stating that they are "not good at mathematics" (Harrison Group, 2010), and between a quarter and a fifth of parents think the ability to learn mathematics is hereditary (Harrison Group, 2010; Johnson, Rochkind, & Ott, 2010). Given such conflicting opinions, schools must continuously work to gain the support of parents and the local community as to the importance of STEM courses and the fact that this is important for all children.

STUDENT ENGAGEMENT

Inclusion of important mathematics content in the curriculum does not assure student learning of the content even if it is taught by knowledgeable teachers. Students, themselves, must engage in the schooling process for learning to take place. The basic notion is that by increasing a student's desire to learn a subject (or simply by mitigating psychological barriers), students will more readily learn academic material. As well as school-based interventions, a host of outreach programs in both the public and private sector seek to increase student engagement in STEM, such as Intel's Inspiring Young Innovators Program and 176 programs funded through the NSF ITEST grants. A limited body of research exists suggesting that increasing student engagement in STEM subjects can lead to improved student learning (Hulleman & Harackiewicz, 2009; Laukenmann, Bleicher, Fuß, Gläser-Zikuda, Mayring, & von Rhöneck, 2003, Patrick & Yoon, 2004; Singh, Granville, Dika, 2002). One analysis suggests that as much as 38% of the variation in student achievement could be explained by student engagement indicators. One of the more promising features of student engagement as a means to improving student achievement is that it appears quite malleable (Fredericks, Blumenfeld, & Paris, 2004). A variety of interventions have been explored that appear to increase student interest in science, including: relating science to students' daily lives (Hulleman & Harackiewicz, 2009), supportive teachers (Klem & Connell, 2009), employing group activities to stimulate social engagement (Olitsky, 2007), the use of hands-on tasks (Blumenfield & Meece, 1988; Silk, Higashi, & Schunn, 2011), incorporating novelty and decision-making into classroom activities (Palmer, 2009), and perhaps involvement by instructional leaders (Quinn, 2002; but see Kruger, Witziers, & Sleegers, 2007).

How to interpret student engagement remains under dispute, however. In a review of the literature on engagement in general, Fredericks et al. (2004) identified three differing conceptions of student engagement: behavioral (participation/effort), emotion (affect/attitudes), and cognitive engagement (psychological involvement). According to Fredericks et al., each of these types of engagement is fraught with definitional and measurement issues, and evidence supporting the impact of engagement often lack an adequate longitudinal element and are vulnerable to charges of spuriousness.

The complexities of studying student engagement, and of designing effective interventions that will stimulate student learning, afflict science and mathematics as much as other disciplines. There is considerable evidence that the effects of instructional practices can vary depending on the type of student motivation (Lee & Brophy 1996, Lee & Anderson 1993, Patrick & Yoon, 2004; Singh et al., 2002). Students who are oriented towards performance goals as opposed to those who have an intrinsic interest in learning science will react less well to some common methods of stimulating engagement and student learning. The challenge therefore is not just in alleviating the negative attitudes toward science exhibited by some students, but in developing approaches to stimulating interest that are flexible enough to meet the needs of many different kinds of students.

INSTRUCTIONAL LEADERSHIP

Instructional leadership has been a major focus of school reform efforts. The meaning of "leadership" has been somewhat murky, however. Leadership has been conceptualized in a variety of different ways, largely depending on the unit of analysis (superintendents, principals, or a more broad-based collaborative approach) and distinguishing direct from indirect effects on student achievement. What is most striking about the research on instructional leadership is the ambiguity of the results: the ability of leadership of whatever type to shape student outcomes appears to be very contingent on circumstances and of equivocal impact.

In the case of district-level leaders, claims that superintendents can influence student achievement by focusing on goals and aligning incentives to meet those goals (Murphy & Hallinger, 1986; Waters & Marzano, 2006) has been questioned by those who argue that superintendents are generally disengaged from curriculum issues, focusing instead on basic administrative functions (Floden et al., 1988; Bredeson & Kose, 2007).

The bulk of research involving instructional leadership has focused on the role of principals. Meta-analyses by Nettles and Harrington (2007), Waters, Marzano and McNulty (2003) and Witziers, Bosker and Krüger (2003) suggest that principals can have a direct effect on improving student outcomes by shaping the institutional mission of the school and promoting an orderly environment conducive to learning. The magnitude of these effects tends to be quite modest however. There is somewhat more evidence for an indirect influence on student achievement, with principal effects mediated by teacher behavior, for example with the proper organization of teacher professional development activities (Robinson, Lloyd, & Rowe, 2008; Graczewski, Knudson, & Holtzman, 2009; Hallinger & Heck, 1998).

A number of scholars have moved away from a top-down model of instructional leadership towards a more collaborative model based on bottom-up organizational transformation in which incentives are aligned with more rigorous instruction (Hallinger, 2003). Leithwood and Mascall (2008) suggest that leadership is a collective rather than individual effort. Hallinger (2005, 2011) argues that strategies relying on leadership (which he calls “leadership for learning”) should recognize the institutional environment will constrain leadership activities. According to Elmore (2000), principals should act as “buffers” for teachers on non-instructional issues, permitting the latter to focus on their core tasks. As a consequence responsibility for school reform should be broadly distributed, with principals acting as “value leaders” and facilitators rather than generals-in-chief.

Despite the attention paid to instructional leadership in general, there has been little research on its specific effects on achievement in STEM areas, Spillane, Diamond, Walker, Halverson, and Jita (2001) and Johnson (2009) being two noteworthy exceptions in science and Heck and Hallinger (2009) in mathematics. Consistent with the collaborative, indirect approach to leadership, these studies found that the proper use of resources, the development of human and social capital, and the management of teacher attitudes can improve student learning in STEM. Even should we accept the more optimistic interpretations of the evidence for the generic impact of instructional leadership, these lessons should be applied with caution in the case of science, as discussions of instructional leadership can go astray when they fail to observe the distinctions among different subjects and grade levels (Spillane 2005). Leadership strategies that may be appropriate in the case of reading could work less well in mathematics or science.

FINAL THOUGHTS

I would argue that of the five main reform efforts discussed in the paper, the two most important are: the curriculum and the teachers who teach it. The others, including external support for a demanding curriculum from parents and the public; methods for fostering student motivation; and instructional leadership on the part of school administrators, especially principals, are also important but more in a supporting role to teachers and the curriculum. I believe that a focused, rigorous and coherent K-8 curriculum backed up by teachers with deep content knowledge and the pedagogical knowledge to make that content accessible to students, plays a powerful and significant role in improving students' ability to learn mathematics and science.

These five factors together play a critical role because their interdependence. The curriculum defines the core of the learning experience (the content, skills, and reasoning to be learned by the students) that is delivered by a teacher who must know that content and be able to present it to students in an accessible fashion. Furthermore, the curriculum and teacher must arouse student interest and motivation but must also be supported externally by parents and the public who endorse the importance of rigorous mathematics courses and internally by principals and superintendents who act not only as administrators but as instructional leaders themselves. The failure to generate sufficient support and the lack of a consensus about goals can lead to debilitating conflicts, such as the "math wars" that in some communities persist to this day.

For mathematics, we, as a nation, have reason to hope that with the rigorous, focused and coherent Common Core State Standards we are at the beginning of a process that will promote excellence, but we will only succeed if we recognize that teachers, principals, superintendents and the public are an essential part of this endeavor. Teachers need to have the content knowledge and pedagogical skills to teach according to those standards attained both through their preparation and through continuing professional development both linked to the Common Core content, and they must be supported by their administrators and the community. Schools operating in such a system with strong teachers are the ones in which all children will have the opportunity to succeed in mathematics. It takes both the individual school and the system in which it finds itself to succeed.

References

Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator, 29*(3), 14-22.

Begle, E. G. (1979). *Critical variables in mathematics education.* Washington, DC: Mathematical Association of American and National Council of Teachers of Mathematics.

Begle, E. G., & Geeslin, W. (1972). *Teacher effectiveness in mathematics instruction*. (National Longitudinal Study of Mathematical Abilities Report No. 28). Washington, DC: Mathematical Association of America and National Council of Teachers of Mathematics.

Blömeke, S. (2005). *Lehrerausbildung—lehrerhandeln—schülerleistungen: Perspektiven nationaler und internationaler empirischer bildungsforschung [Teacher education—teacher performance—student achievement:Perspectives of national and international empirical research in education]*. Public Lecture on December 10, 2003, at Humboldt University. Berlin, Germany: Humboldt-Universität.

Blumenfeld, P. C., & Meece, J. L. (1998). Task factors, teacher behavior, and students' involvement and use of learning strategies in science. *The Elementary School Journal, 88*(3), 235–250.

Bredeson, P. V., & Kose, B. W. (2007). Responding to the education reform agenda: A study of school superintendents' instructional leadership. *Education Policy Analysis Archives, 15*(5). Retrieved April 10, 2011 from http://epaa.asu.edu/epaa/v15n5/.

Bromme, R. (1992*). Der Lehrer als Experte. Zur Psychologie des professionellen Lehrerwissens [The teacher as expert. On the psychology of professional knowledge]*. Bern: Huber.

Carmichael, S. B., Martino, G., Porter-Magee, K., & Wilson, W. S. (2010). *The State of State Standards-and the Common Core-in 2010.* Washington, DC: Thomas B. Fordham Institute.

Center for Research in Mathematics and Science Education (2010). *Breaking the cycle: An international comparison of U.S. mathematics teacher preparation.* East Lansing, MI: Michigan State University.

Clotfelter, C. T. & Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*(6), 673–682.

Darling-Hammond, L. (1999). The case for university-based teacher education. In R. Roth (Ed.), *The role of the university in the preparation of teachers* (pp. 13–30). Philadelphia, PA: Falmer Press.

Darling-Hammond, L., & Youngs, P. (2002). Defining “highly qualified teachers”: What does “scientifically-based research” actually tell us? *Educational Researcher, 31*(9), 13–25.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology, 19*(2), 294–304.

Elmore, R. F. (2000). *Building a new structure of school leadership.* The Albert Shanker Institute.

Floden, R. E. (2002). The measurement of opportunity to learn. In A.C. Porter & A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 231-266). Washington, DC: National Academy Press.

Floden, R. E., Porter, A. C., Alford, L. E., Freeman, D. J., Irwin, S., Schmidt, W. H., & Schwille, J. R. (1988). Instructional leadership at the district level: A closer look at autonomy and control. *Educational Administration Quarterly, 24,* 96-124. Retrieved on April 15, 2011 from http://eaq.sagepub.com/content/24/2/96

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of t he concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109.

Gamoran, A., Porter, A. C., Smithson, J., & White, P. A. (1997). Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth. *Educational Evaluation and Policy Analysis, 19*(4), 325–338.

Good, T. L., & Grouws, D. A. (1987). Increasing teachers' understanding of mathematical ideas through inservice training. *Phi Delta Kappan,68*(10), 778–783.

Goyette, K. & Xie, Y. (1999). Educational expectations of Asian American youths: Determinants and ethnic differences. *Sociology of Education, 72*(1), 22–36.

Graczewski, C., Knudson, J., & Holtzman, D. J. (2009). Instructional leadership in practice: What does it look like, and what influence does it have? *Journal of Education for Students Placed at Risk, 14*(1), 72–96.

Hallinger, P. (2003). Leading educational change: Reflections on the practice of instructional and transformational leadership. *Cambridge Journal of Education, 33*(3), 329–352.

Hallinger, P. (2005). Instructional leadership and the school principal: A passing fancy that refuses to fade away. *Leadership and Policy in Schools, 4*(3), 221–239.

Hallinger, P. (2011). Leadership for learning: lessons from 40 years of empirical research. *Journal of Educational Administration, 49*(2), 125–142.

Hallinger, P., & Heck, R. H. (1998). Exploring the principal's contribution to school effectiveness: 1980-1995. *School Effectiveness and School Improvement, 9*(2), 157–191.

Harrison Group, Inc. (2009). [Findings among parents: Attitudes towards school, math, and science]. Unpublished raw data.

Harrison Group, Inc. (2010). [Findings among parents: Attitudes towards school, math, and science]. Unpublished raw data.

Heck, R. H., & Hallinger, P. (2009). Assessing the contribution of distributed leadership to school improvement and growth in math achievement. *American Educational Research Journal, 46,*(3), 659–689.

Howell, W., Peterson, P. E., & West, M. (2009). 2009 Education next – PEPG survey of public opinion. *Education Next, 9,*(4).

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes [Electronic version]. *Science, 326,* 1410–1412.

Intel. (2009, October 21). *Intel news release: Parents more comfortable talking drugs than science.* Retrieved December 23, 2010, from http://www.intel.com/pressroom/archive/releases/2009/20091021edu.htm

Johnson, C. C. (2009). An examination of effective practice: Moving toward elimination of achievement gaps in science. *Journal of Science Teacher Education, 20,* 287–306.

Johnson, J., Rochkind, J., & Ott, A. (2010). Are we beginning to see the light? Public and parents buy into the need to ramp up math and science education but most still think their local schools are doing fine. *Public Agenda.* Retrieved December 23, 2010 from http://www.publicagenda.org/pages/math-and-science-ed-2010

Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health, 74,* 262–273. doi: 10.1111/j.1746-1561.2004.tb08283.x

Kruger, M. L., Witziers, B., & Sleegers, P. (2007). The impact of school leadership on school level factors: Validation of a causal model. *School Effectiveness and School Improvement, 18*(1), 1–20.

Laukenmann, M., Bleicher, M., Fuß, S., Gläser-Zikuda, M., Mayring, P., & von Rhöneck, C. (2003). An investigation of the influence of emotional factors on learning in physics instruction. *International Journal of Science Education, 25*(4), 489–507.

Lee, O., & Anderson, C.W. (1993). Task engagement and conceptual change in middle school science classrooms. *American Educational Research Journal, 30*(3), 585 -610.

Lee, O., & Brophy, J. (1996), Motivational patterns observed in sixth-grade science classrooms. *Journal of Research in Science Teaching, 33,* 303–318. doi: 10.1002/(SICI)1098-2736(199603)33:3<303::AID-TEA4>3.0.CO;2-X

Leithwood, K., & Mascall, B. (2008). Collective leadership effects on student achievement. *Educational Administration Quarterly, 44,* 529–561.

LeTendre, G. K., Baker, D. P., Akiba, M., Goesling, B., & Wiseman, A. (2001). Teachers' work: Institutional isomorphism and cultural variation in the U.S., Germany, and Japan. *Educational Researcher, 30,* 3.

Murphy, J., & Hallinger, P. (1986). The superintendent as instructional leader: Findings from effective school districts. *The Journal of Educational Administration, XXIV*(2), 213–236.

Nettles, S. M., & Herrington, C. (2007). Revisiting the importance of the direct effects of school leadership on student achievement: The implications for school improvement policy. *Peabody Journal of Education, 82*(4), 724–736.

National Research Council. (2010). *Preparing teachers: Building evidence for sound policy.* Committee on the Study of Teacher Preparation Programs in the United States, Division of Behavioral and Social Sciences and Education (p. 232). Washington, DC: National Academy Press.

Olitsky, S. (2007). Promoting student engagement in science: Interaction rituals and the pursuit of a community of practice. *Journal of Research in Science Teaching, 44*(1), 33–56.

Olson, J. F., Martin, M. O., & Mullis, I. V. S. (Eds.). (2008). *TIMSS 2007 technical report.* Chestnut Hill, MA: Boston College, TIMSS & PIRLS International Study Center.

Palmer, D. H. (2009). Student interest generated during an inquiry skills lesson. *Journal of Research in Science Teaching, 3,* 147-169.

Patrick, H. & Yoon, C. (2004). Early adolescents' motivation during science investigation. *The Journal of Educational Research, 97*(6), 319–328.

PROM/SE. (2006). *Knowing mathematics: What we can learn from teachers* (Research Report Vol. 2). East Lansing, MI: Michigan State University.

Quinn, D. M. (2002). The impact of principal leadership behaviours on instructional practice ands tudent engagement. *Journal of Educational Administration, 40* (5), 447-467.

Rivkin, S., Hanushek, E., & Kain, J. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458.

Robinson, V. M. J., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly, 44*, 635–674.

Rowan, B., Correnti, R., & Miller, R. J. (2002). *What large-scale survey research tells us about teacher effects on student achievement: Insights from the prospects study of elementary schools* (Tech. Rep.). Philadelphia, PA: Consortium for Policy Research in Education.

Schmidt, W. H. (2009). *Exploring the relationship between content coverage and achievement: Unpacking the meaning of tracking in eighth grade mathematics.* East Lansing, MI: Michigan State University, The Education Policy Center.

Schmidt, W. H., Cogan, L. S., & Houang, R. T. (in press). The role of opportunity to learn in teacher preparation: An international context. *Journal of Teacher Education.*

Schmidt, W. H., Cogan, L. S., Houang, R. T., & McKnight, C. (2009). *Equality of educational opportunity: A myth or reality in U.S. schooling.* East Lansing, MI: Michigan State University, The Education Policy Center.

Schmidt, W. H., & Houang, R. T. (2007). Lack of focus in the mathematics curriculum. In T. Loveless (Ed.), *Lessons learned: What international assessments tell us about math achievement.* Washington, DC: Brookings Institution Press.

Schmidt, W. H., & Maier, A. (2009). Opportunity to learn. In G. Sykes, B. L. Schneider, & D. N. Plank (Eds.), *Handbook on Education Policy Research* (pp. 541–549). New York: Routledge.

Schmidt, W. H., & McKnight, C. C. (in press). *Inequality for all: Why America's schools are failing our children.* New York: Teachers College Press.

Schmidt, W. H., McKnight, C., Cogan, L.S., Jakwerth, P.M., & Houang, R.T. (1999). *Facing the consequences: Using TIMSS for a closer look at U.S. mathematics and science education*. Dordrecht/Boston/London: Kluwer.

Schmidt, W. H., McKnight, C. C., Houang, R. T., Wang, H., Wiley, D., Cogan, L. S., et al. (2001). *Why schools matter: A cross-national comparison of curriculum and learning*. San Francisco, CA: Jossey-Bass.

Schmidt, W. H., McKnight, C. C., & Raizen, S. A. (1996). *Splintered vision: An investigation of U.S. science and mathematics education: Executive summary*. East Lansing, MI: Michigan State University, U.S. National Research Center for the Third International Mathematics and Science Study.

Schmidt, W. H., McKnight, C. C., Valverde, G. A., Houang, R. T., & Wiley, D. E. (1997). *Many visions, many aims, volume I: A cross-national investigation of curricular intentions in school mathematics*. Dordrecht/Boston/London: Kluwer.

Schmidt, W. H., Raizen, S. A., Britton, E. D., Bianchi, L. J., & Wolfe, R. G. (1997). *Many visions, many aims, volume II: A cross-national investigation of curricular intentions in school science*. Dordrecht/Boston/London: Kluwer.

Schmidt, W. H., Wang, H. A., & McKnight, C. C. (2005). Curriculum coherence: An examination of US mathematics and science content standards from an international perspective. *Journal of Curriculum Studies, 37*(5), 525–529.

SciMathMN (2008). Minnesota TIMSS report December 2008: A preliminary summary of results. Retrieved December 23, 2010 from http://www.scimathmn.org/timss/timss2007/2007timss_report.pdf

Silk, E. M., Higashi, R., & Schunn, C. D. (2011). Resources for robot competition success: Assessing math use in grade-school-level engineering design. *American Society for Engineering Education.*

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *The Journal of Educational Research, 95*(6), 323–332.

Spillane, J. P. (2005). Primary school leadership practice: How the subject matters. *School Leadership & Management, 25*(4), 383–397.

Spillane, J. P., Diamond, J. B., Walker, L. J., Halverson, R., & Jita, L. (2001). Urban school leadership for elementary science instruction: Identifying and activating resources in an undervalued school subject. *Journal of Research in Science Teaching, 38*(8), 918–940.

Suter, L. E. (2000). Is student achievement immutable? Evidence from international studies on schooling and student achievement. *Review of Educational Research, 70*, 529–545.

Waters, J. T., & Marzano, R. J. (2006). School district leadership that works: The effect of superintendent leadership on student achievement (working paper). Retrieved on April 15, 2011 from www.mcrel.org.

Waters, T., Marzano, R. J., & McNulty, B. (2003). Balanced leadership: What 30 years of research tells us about the effect of leadership on student achievement (working paper). Retrieved on April 15, 2011 from www.mcrel.org.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89–122.

Weinert, F.E. (2001). Concept of competence: A conceptual clarification. In: D.S. Rychen & L.H. Salganik (eds.), Defining and Selecting Key Competencies (pp. 45–66). Göttingen: Hogrefe.

Weinert, F. E. (1999). *Konzepte der kompetenz. Gutachten zum OECD-projekt "Definition and selection of competencies: Theoretical and conceptual foundations (DeSeCo)" [Concepts of competence. Expertise on the OECD project "definition and selection of competencies: Theoretical and conceptual foundations (DeSeCo)"]*. Neuchatel, Switzerland: Bundesamt für Statistik.

Witziers, B., Bosker, R. J., & Krüger, M. L. (2003). Educational leadership and student achievement: The elusive search for an association. *Educational Administration Quarterly, 39*, 398–425.

Wu, H. -H. (December 19, 2008). The Mathematics K-12 Teachers Need to Know. (pp. 72): University of California, Berkeley.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from http://ies.ed.gov/ncee/edlabs

Zumwalt, K., & Craig, E. (2005). Teachers' characteristics: Research on the indicators of quality. In M. Cochran-Smith & K. Zeichner (Eds.*) Studying teacher education: The report of the AERA panel on research and teacher education.* Mahwah, NJ: Lawerence Erlbaum Associates.

End Notes

[1] The exact percentage is difficult to calculate because of varying definitions of what constitutes Algebra I. In some cases course are listed as Algebra but in fact spend much of their time on arithmetic.