Using Indicator Data to Drive K–12 STEM Improvements in States and Districts:

Implications and Recommendations for Leaders and Policymakers

February 2016

SRI Education™
A DIVISION OF SRI INTERNATIONAL

Author

Michael Lach, Director of STEM Policy and Strategic Initiatives,
University of Chicago

This material is based upon work supported by the National Science Foundation under Contract No. NSFDACS12C1299. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation

Suggested Citation

Lach, M. (2016). *Using Indicator Data to Drive K–12 STEM Improvements in States and Districts: Implications and Recommendations for Leaders and Policymakers* (SRI Education White Paper). Menlo Park, CA: SRI International.

Contents

Executive Summary

Drawing on his experience as a STEM education leader for Chicago Public Schools and the U.S. Department of Education, Michael Lach explores how district and state education leaders might use data from an indicator system to improve K-12 STEM education performance overall and to reduce inequities in the opportunities available to different student subgroups and communities. Lach argues that new metrics, coupled with appropriate support, can improve education.

Indicators and measures have long served a critical role in the U.S. education system. For district and state education leaders, monitoring data describe current initiatives and support informed decision-making.

In *Monitoring Progress Toward Successful K-12 STEM Education*, the National Research Council (NRC) argues for new and enhanced indicators in science, technology, engineering, and mathematics (STEM) education (see table below). This set of STEM-specific indicators would expand education monitoring beyond the mathematics and English language arts performance concerns of No Child Left Behind to address issues associated with STEM education and its improvement.

Several of the NRC-proposed STEM education indicators can be used by district and school leaders to drive improvement. Specifically,

- **STEM Instructional Materials (Indicator 4).** District strategies that focus on the adoption and implementation of robust instructional materials have been shown to lead to sizable student learning gains. However, many STEM instructional materials are not aligned with rigorous, research-based standards, such as the Next Generation Science Standards, that require more than memorization of vocabulary and rote procedures. This indicator would signal the importance of adoption and alignment of high-quality STEM instructional materials and the importance

Access to Quality STEM Learning	1. Number of and enrollment in different types of STEM schools and programs in each district	Policy and Funding Initiatives	9. Inclusion of science in federal and state accountability systems
	2. Time allocated to teach science in grades K-5		10. Inclusion of science in major federal K-12 education initiatives
	3. Science-related learning opportunities in elementary schools		11. State and district staff dedicated to supporting science instruction
	4. Adoption of instructional materials in grades K-12 that embody rigorous, research-based standards		12. States' use of assessments that measure the core concepts and practices of science and mathematics disciplines
	5. Classroom coverage of content and practices in rigorous, research-based standards		13. State and federal expenditures dedicated to improving the K-12 STEM teaching workforce
Educators' Capacity	6. Teachers' science and mathematics content knowledge for teaching		14. Federal funding for the research that disentangles the effects of school practice from student selection, recognizes the importance of contextual variables, and allows for longitudinal assessments of student outcomes
	7. Teachers' participation in STEM-specific professional development activities		
	8. Instructional leaders' participation in professional development on creating conditions that support STEM learning		

Source: *Monitoring Progress Toward Successful K-12 STM Education: A Nation Advancing?* National Research Council, 2013.

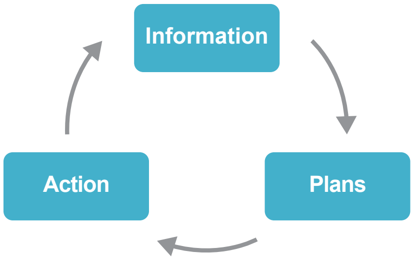
Note: Indicator statements 1, 4, 5, 10, and 14 have been edited slightly to enhance clarity when presented outside the context of the full *Monitoring Progress* report.

of drawing on research to measure the quality of materials considered for adoption.

- **Leadership for Principals (Indicator 8).** School leaders have a significant impact on the learning that occurs in their schools. These leaders must have STEM-specific content knowledge if they are to help shape the school's climate and allocation of resources to promote high-quality STEM learning. Data for this indicator would provide an impetus to provide professional development that helps principals become effective STEM education leaders.
- **Time Allocated to Teach Science (Indicator 2).** It has been a challenge for schools and school systems to make science a focus of improvement efforts when accountability systems address language arts and mathematics almost exclusively. Nationally, little time is devoted to science education in elementary school, despite the facts that it is a national priority and interest in science needs to be developed early. Achievement in science, as in other subject areas, is a function of instructional time. Indicator data would enable districts and schools to know how much time teachers are spending on science and to make informed decisions about whether to adjust the balance of instructional time across subject areas.

These indicators would be a means to improvement, not the ends in themselves. America's challenges with STEM education are numerous, and they will not be solved overnight. However, the STEM education indicators point to a thoughtful path forward: robust indicators that focus on what matters, used by schools and districts that have the trust and mechanisms to learn from and iterate based on indicator data over time.

Implications for Policy and Practice



A set of regularly measured indicators in STEM education could assist district and school leaders in making decisions about education policies, helping redirect efforts to practices that are well supported by the research literature. Common indicators could also facilitate cross-district sharing and learning partnerships.

Districts should adopt STEM curricula that are aligned with rigorous, research-based learning standards. Unfortunately, many curricula do not meet this criterion. Policymakers should invest in research and development of instructional materials that are well aligned with research-based learning standards.

States and districts should support principals in obtaining professional development on STEM education leadership. Content-specific professional development remains a low priority for many leadership development efforts, despite principals' significant role in students' learning. Policymakers should invest in professional development for principals on STEM education leadership.

Schools and districts should devote adequate instructional time to science education. Accurate measures of the time elementary school classrooms spend on science would enable research on what constitutes adequate time and would help school and district administrators make time allocation decisions.

Introduction

Indicators and measures have long served a critical role in the U.S. education system. Tests have scores. Report cards have grades. Our public schools are public entities, so taxpayers expect to understand how our tax dollars are spent. For district and state education leaders, monitoring data that describe current initiatives and support informed decision-making is critical. To evaluate programs and efforts, indicators that describe outcomes and results are similarly essential.

In *Monitoring Progress Toward Successful K-12 STEM Education,* the National Research Council committee argued for a new and enhanced set of indicators (see Exhibit 1), measured regularly at different levels of the educational system, that will enable “Congress and relevant Federal agencies to create and implement a national-level monitoring and reporting system” (Committee on the Evaluation Framework for Successful K-12 STEM Education, 2013). The utility of such a set of indicators is at least as great for state and district leaders who are focused on improving science, technology, engineering, and mathematics (STEM) education within their jurisdictions. A set of indicators not only has the potential to enhance the overall effectiveness of our collective STEM education efforts, but could also support efforts to reduce inequities between student subgroups and communities. As this paper argues, introducing new metrics and then organizing the monitoring system to deliver on them can result in significant improvements for students. But simply choosing new metrics is not likely to make a difference unless they are strategically selected with an eye toward the capacity of the education system to use such measures in ways that matter for overall improvement.

This discussion is based on a review of the research literature, on evaluation data from large-scale science and mathematics improvement efforts, and on the author’s experience as STEM education leader for the Chicago Public Schools and the U. S. Department of Education.

Exhibit 1: Indicators described in *Monitoring Progress: Toward Successful K-12 STEM Education: A Nation Advancing?*

Access to Quality STEM Learning	Educators’ Capacity	Policy and Funding
1. Number of and enrollment in different types of STEM schools and programs in each district 2. Time allocated to teach science in grades K-5 3. Science-related learning opportunities in elementary schools 4. Adoption of instructional materials in grades K-12 that embody rigorous, research-based standards 5. Classroom coverage of content and practices in rigorous, research-based standards	6. Teachers’ science and mathematics content knowledge for teaching 7. Teachers’ participation in STEM-specific professional development activities 8. Instructional leaders’ participation in professional development on creating conditions that support STEM learning	9. Inclusion of science in federal and state accountability systems 10. Inclusion of science in major federal K-12 education initiatives 11. State and district staff dedicated to supporting science instruction 12. States’ use of assessments that measure the core concepts and practices of science and mathematics disciplines 13. State and federal expenditures dedicated to improving the K-12 STEM teaching workforce 14. Federal funding for research that disentangles the effects of school practice from student selection, recognizes the importance of contextual variables, and allows for longitudinal assessments of student outcomes

Source: *Monitoring Progress Toward Successful K-12 STM Education: A Nation Advancing?* Committee on the Evaluation Framework for Successful K-12 STEM Education, National Research Council, 2013.

Note: Indicator statements 1, 4, 5, 10, and 14 have been edited slightly to enhance clarity when presented outside the context of the full *Monitoring Progress* report.

Context of Indicators in Public Education: Potential for Improving STEM Education

By providing clear goals, defining a common purpose, and offering a foundation of consistent facts, indicators can help inform and support improvement in education as they do in other fields. The No Child Left Behind (NCLB) legislation ("No Child Left Behind Act of 2001," 2001) required that states develop annual adequate yearly progress targets for schools and districts for all students and for key subgroups of students based on state test results, student test participation rates, and one other academic indicator (such as graduation rate). As a result, every state publishes district and school report cards "to provide a snapshot of how well that district and school is educating its students" (Mikulecky & Christie, 2014). While NCLB was replaced with the Every Student Succeeds Act (ESSA) in December 2015, district and school report cards are likely to continue due to similar rules in ESSA. Especially in large urban school districts, many other indicators are in use to describe the core operational functions of those systems (for an example of management-focused indicators, see Casserly & Carlson, 2014).

A review of these reports at the state and district levels (Education Commission Of The States, 2013) shows a consensus about the metrics that are most important to report: student achievement, student growth or academic improvement, a measure of achievement gap closure, high school graduation rate, and postsecondary readiness (Mikulecky & Christie, 2014). Moreover, these efforts are not limited to the state or district. Over the past decade, efforts to improve "data-driven decision making" from the classroom level on up have taken root (Gill, Borden, & Hallgren, 2014), and the indicators used at the state and district levels generally address the priorities laid out by NCLB. From a STEM education perspective, they include the following considerations:

- **Mathematics.** Given that mathematics continues to be a priority in ESSA, all states use mathematics performance in their reporting arrangements.
- **Science.** ESSA requires that science be assessed in one grade at each level of schooling but does not require science scores to be used in calculating adequate yearly progress for purposes of accountability. Although there have been efforts to add science to statewide accountability provisions (Ayers, Owen, Partee, & Chang, 2012), few state or district indicator systems use science as a core indicator.
- **Additional STEM-specific metrics.** Some states have developed indicators that relate to particular STEM education issues or contexts. For example, and with some debate, Illinois plans to include the percentage of students passing Algebra I by the end of eighth grade in 2015 (Illinois State Board Of Education, 2015) as part of its school report cards.

It is generally accepted that what is measured is what matters for practitioners. Entire programs and strategies have been developed to help schools, districts, and states better identify important indicators and organize their efforts to achieve results on those priorities (for instance, Barber, Moffit, & Kihn, 2010). However, there are some cautions to consider as well.

Limitations of the Current System for Improvement

The current system of STEM education data collection and use has limitations, particularly if the goal is improvement at all levels. Many of the lessons to be learned are driven by the policies established in the context of NCLB, so by unpacking the impacts of that law we can begin to understand the sorts of adjustments that might be needed for a system of STEM indicators to have positive effects. Even though ESSA has replaced NCLB, much remains to be learned from the implementation of the 2001 law.

Choosing the Right Metrics Is Not Trivial

NCLB charged schools and systems to improve their overall test performance in English language arts and mathematics, operationalized as the percentage of students at each grade level and within defined demographic subgroups meeting a state-defined proficiency level as measured on a state test. There is evidence, however, that in many places the law's accountability provisions led schools to focus on the "bubble kids"—those close to the state's cutoff score for proficiency.

In a study of testing and accountability policy in Chicago between 1996 and 2002, Neal and Schanzenbach (Neal & Schanzenbach, 2007) demonstrate how district-driven accountability policies in 1996 and NCLB-driven policies in 2002 led to a focus on students in the middle of the achievement distribution, but the subsequent improvements did not address those students who were significantly below or above the proficiency cutoff score. They conclude that "the choice of the proficiency standard in such accountability systems determines the amount of time that teachers devote to students of different ability levels," undermining the overall goals of the NCLB legislation (Neal & Schanzenbach, 2007, page 2).

The best metrics relate closely to an important aspect of system performance and are relatively impervious to gaming. Choosing the right metrics matters.

Educational Systems Are Complex

A parallel criticism of NCLB-era accountability builds on the complexities of schools and school systems. Schools and school systems consist of many different pieces that interlock and intersect in complicated ways (Cuban, 2013). Particularly in the case of high-need students, these interactions may result in irrational combinations of policies and practices (Payne, 2008). Further, although there is some evidence that efforts to shine a light on performance gaps can cause systems to change in response (Lauen & Gaddis, 2012), as the gaps get closer to the classroom and the student and focus more on instruction and less on access, identifying suitable strategies to close (not just identify) the gaps becomes more challenging. Setting targets (e.g., increase mathematics performance by 10% next year) is an initial step. But determining precisely how to reach those goals—what changes in curriculum, instruction, and leadership will result in that gain—is much more difficult. As Stigler and Hiebert (1997) commented, "a focus on standards and accountability that ignores the processes of teaching and learning in classrooms will not provide the direction that teachers need in their quest to improve" (Stigler & Hiebert, 1997). If teachers, principals, and system leaders do not understand the nuances of how the system works or how student learning can be promoted, the improvements stemming from their change efforts are likely to be limited (also see R. Elmore, 2004).

For instance, in the mid-2000s, the Chicago Public Schools launched an initiative to better support students who were behind in mathematics by enrolling entering high school students who were below the 50th percentile in mathematics into an extended “double-dose” algebra course. The idea behind this essentially was that students who were behind in mathematics needed additional instructional time to catch up and that attempts to remediate them by slowing them down (programming them into pre-algebra courses, for instance) would not be effective. After evaluating this effort, researchers found that as the program designers intended, test scores did go up both for below-average students and for above-average students. Program planners were surprised to find, however, that above-average students had higher algebra failure rates after implementation of the new policy (Nomi & Allensworth, 2012). To understand the data, one had to understand that to implement the policy, many schools instituted student sorting, with above-average students going into a set of single-period algebra classes and below-average students going into a set of double-period classes. Thus, the outcomes reflected the results of an increase in the homogeneity of algebra classes as well as an increase in instructional time for lower achieving students. As Nomi and Allensworth (2014) described in a follow-up study,

> This happened partly because teachers demanded more from students in classes with higher-achieving students, making it more difficult to pass. More critically, the double-dose algebra policy caused some high-skilled students to become the lowest-skilled students in their class—particularly if their math skills were just above the national average. Students with skills at the bottom of their class were much more likely to fail. This might be due to teachers’ grading practices, or reduced effort among students who feel frustrated from falling behind. (Nomi & Allensworth, 2014, page 6)

While long-term results for the double-dose strategy were positive (Cortes, Goodman, & Nomi, 2013), the implementation research revealed complexities within the system that surprised program planners and researchers alike.

Furthermore, identifying a goal is not the same as determining how to accomplish it. The idea of the “instructional core” (City, Elmore, Fiarman, & Teitel, 2009) has great traction in the school improvement literature. But connecting school or district interventions to the instructional core is challenging—for a typical school, a myriad of interventions could possibly be put in place to improve mathematics achievement. The nature of those interventions will determine the impact they have on student learning, and frequently well-meaning but uninformed efforts fail to show results because the ability to implement the selected intervention does not exist at the school (e.g., E. Allensworth, Correa, & Ponisciak, 2008). As Ladd (2007) commented in her review of educational accountability systems, we have “too much reliance on test-based accountability, and too little attention to promoting effective process and practice within schools” (Ladd, 2007).

Improvement Within Systems

Part of the process of systems improvement is increasing the capacity of the individuals and teams that make up the system (R. F. Elmore, 2002). Establishing data systems is a common intervention across many school system improvement strategies (Mourshed, Chijioke, & Barber, 2010). A new set of K-12 STEM education indicators has the potential to significantly increase the rate of improvement as part of such data systems by providing data related to practices within the control of districts and schools.

Several examples demonstrate how introducing new metrics and then organizing the system to deliver on them can result in significant improvements for students. Below are three examples of leadership decisions or policies that involved introducing new metrics and resulted in systemic changes that improved student outcomes. These examples share several important qualities that contributed to their success: The metrics were both easy to understand and (relatively) easy to measure, and leaders understood system complexities with sufficient depth to design strategies that could move the needle not only on the metric in question, but also on the underlying long-term goals for students' success. Equally important, the education system had sufficient capacity (even if diffuse) to implement the new strategies.

Example 1: Freshman On-Track

Increasing the graduation rate in the Chicago Public Schools has historically been a vexing challenge—the 5-year cohort graduation rate of the Chicago Public Schools hovered under 53% for the eight-year span between 1999 and 2006 (Chicago Public Schools, 2015). An analysis by researchers at the University of Chicago showed that ninth-grade semester core course grades were a strong predictor of the likelihood of students graduating 3 or 4 years later, leading to the development of a new indicator called "freshman on-track" based on ninth-grade credit accumulation and the number of failing semester grades. Research demonstrated that the freshman on-track rating is highly predictive of eventual graduation (E. Allensworth & Easton, 2005).

Systems were established to provide additional supports to students, teachers, and schools around the ninth grade (E. M. Allensworth & Easton, 2007); these included monitoring reports, additional tutoring for students who showed a risk of becoming off track, and supports for counselors and school administrators to redesign school systems to better enable a focus on reducing off-track behavior (Network For College Success, 2015). Given the local attention, capacity was quickly developed to implement and make good on the information the freshman on-track metric provided.

Since 2007, the high school graduation rate has improved (Ahmed-Ullah & Byrne, 2014). Based on this work, efforts have been under way to explore similar on-track measures at the middle school level (E. M. Allensworth, Gwynne, Moore, & Torre, 2014) and in other jurisdictions (Burke, 2015; Kemple, Segeritz, & Stephenson, 2013).

Example 2: FAFSA Completion

Another example of system changes in response to a new indicator is the Chicago Public Schools' rate of completion of the Free Application for Federal Student Aid (FAFSA) form. Completing FAFSA maximizes a student's prospects for receiving federal, state, and institutional financial aid for college, something that is extremely important for low-income students who are likely to be the first members of their family to seek a postsecondary degree.

After research emerged that showed failure to complete FAFSA (and hence to receive needed aid) was a significant barrier to college going among the city's low-income students (Roderick et al., 2008), Chicago Public Schools responded by measuring school FAFSA completion rates and providing a series of supports for students and schools to increase them. The district's actions included a weekly monitoring system and incentives to schools for improved FAFSA completion rates. These efforts resulted in considerable improvements in this metric in subsequent years (Chicago Public Schools, 2011) and have led to national attention to the issue of improving this indicator for all schools (Aldeman, 2015).

Example 3: Eighth-Grade Algebra

In the mid-2000s, the Chicago Public Schools partnered with several local universities on the Algebra Initiative, with the goal that "every [Chicago] middle school…could offer algebra to qualified students in the eighth grade" (Jabon et al., 2010). The initial work was to develop a series of university-based courses that prospective algebra teachers could take to ensure they understood the mathematics content, and the district counted both the number of teachers who successfully completed the coursework and the number of schools with completing teachers. As in the previous examples, there was already considerable capacity within Chicago and the district to initiate such an effort and to expect it might succeed (Deiger et al., 2009).

As Jabon et al. (2010) described, the district soon found that better prepared teachers did not translate immediately into increases in middle school mathematics outcomes at scale. Better prepared teachers appeared to be a necessary but not sufficient condition for raising the rates of algebra completion in grade 8. District administrators subsequently added and revised several Algebra Initiative components, including an approval process for schools to complete before offering eighth-grade algebra, a qualifying exam for teachers, and a comprehensive eighth-grade algebra exam for students (Fendt, Harris, Lent, & Wenzel, 2008).

Over time, the district's results on these metrics have increased dramatically, from a baseline close to zero in 2007. In 2014, 525 teachers had completed the university-based coursework and passed the teacher qualifying exam, 207 schools offered an approved middle grades algebra course, and 5,720 middle grade students took and 55% of them passed the citywide algebra exit exam, compared with 1,114 students taking and just 29% passing the exam in 2007 (Fulton, 2014). Moreover, students who performed well in eighth-grade algebra have been found to have a greater likelihood of higher grades and test scores in future mathematics work in Chicago (Schmidt, 2009).

Improvement Across Schools and Districts

Schools and districts must increase learning within their systems. But they face numerous challenges—challenges that often are in proportion to their resources and institutional capacity for improvement. The commonality of the problems districts face suggests that they would be amenable to joint solutions, so districts also can and should learn from one another. The "not invented here" syndrome is unfortunately all too common among schools, districts, and states, although given the inherent complexities with partnerships and the rapid turnover of district superintendents (Grissom & Andersen, 2012), this is not surprising. If not a prerequisite then certainly a strong support for cross-district learning is use of common metrics across contexts to compare and adapt strategies, interventions, and implementations. Several metrics and processes currently in existence do just this.

- The National Assessment of Educational Progress (NAEP) provides a common baseline for comparison of state-level progress; the Trial Urban District Assessment (http://www.nationsreportcard.gov/tuda.aspx), based on district-level NAEP data, provides a similar function for large urban districts.
- Several efforts are under way to ratchet up the learning and sharing between school districts and nearby charter management organizations, such as Schools That Can. A key component of these efforts is the articulation of common goals across organizations and the determination of common metrics.
- Although practitioners often cringe when external organizations assign "grades" to aspects of their work (e.g., *Education Week's* regular Quality Counts feature), these efforts can certainly spur comparisons and reflection, if not collaboration.
- Several mathematics-specific efforts to foster learning between districts, such as the Suburban Cook County Mathematics Initiative in Illinois and the Math-In-Common online community for principals based in California, focus on common metrics as a lever to increase mathematics learning across systems.

The Case for STEM-Specific Indicators

The curriculum-narrowing effects of No Child Left Behind are well documented (Jennings & Rentner, 2006; Pederson, 2007) and much lamented. There have been some efforts to include science as part of state accountability frameworks (Judson, 2012), but they have been limited and certainly have not included much focus on the engineering or technology components of STEM education. A set of STEM-specific indicators, such as those envisioned by *Monitoring Progress*, would expand monitoring beyond the mathematics and English language arts concerns of NCLB to address the particular issues associated with STEM education improvement efforts.

Because science, technology, engineering, and mathematics are different content areas, they demand content-specific strategies for school and system improvement. A content-agnostic "standards implementation" strategy or an English language arts strategy repurposed for science most likely will not be successful. Below is a sampling of reasons why the STEM disciplines should be treated distinctly. Similar reasons support the separation of science from mathematics, engineering, and technology within STEM (C-STEMEC, 2013).

- **STEM disciplines are highly sequential.** A review of various learning progressions in science and mathematics (such as at http://ime.math.arizona.edu/progressions/ and National Research Council, 2007) clearly shows how the content in one grade builds on content presumably learned in previous years. Because the content to be learned at a particular point in time depends on the knowledge of prior content, remediation strategies for students behind in mathematics need to be adjusted (Daro, 2007; Stoelinga & Lynn, 2013). Although sequences appear to be more flexible in science than in mathematics, by the time students get to high school science, prior science learning as well as a set of mathematics skills are needed to succeed. For multischool districts with significant within-district student mobility, the issue of when particular content is taught becomes particularly pressing.
- **Schools organize around science and mathematics differently than they do around literacy.** In general, schools have much tighter knowledge networks and more frequent adult interactions associated with English language arts instruction than with mathematics or science instruction (Spillane, 2015). Schools perceive the source of support differently for English language arts and mathematics, with support for mathematics improvement more likely than that for literacy improvement to originate outside the school (Spillane, 2005; Stein & Spillane, 2005). Given the sequential nature of STEM disciplines, the need for vertical alignment within a school becomes especially important, with all the associated organizational and school leadership issues entailed.
- **School leaders do not perceive improving STEM outcomes as all that important.** There has been a widespread push to improve STEM education outcomes for decades (for instance, Committee on Prospering in the Global Economy of the 21st Century, 2006; The National Commission on Mathematics and Science Teaching for the 21st Century, 2000). Despite these efforts, surveys show that most principals and superintendents think that improving science and mathematics education is "not a serious problem in their local schools" (Johnson, Arumi, & Ott, 2006).

- **Instructional materials play a unique role in mathematics and science.** In mathematics and science more than in other academic subjects, instructional materials tend to determine the nature and sequence of the content delivered to students as well as the way that content is taught (Banilower et al., 2013). Efforts to enhance the quality of instructional materials have the potential to drive improvements (Ball & Cohen, 1996; Banilower, Boyd, Pasley, & Weiss, 2006; Briars, 2002; Sconiers, Isaacs, Higgins, McBride, & Kelso, 2003).
- **The general public—including many educators—does not understand the practice of science.** On issues like climate change, evolution, or vaccination, a large segment of the U.S. population has views outside the scientific mainstream, resulting in some cases in antagonism toward or a fear of science (Mooney & Kirshenbaum, 2010). This fear and antagonism extend to teachers (Beilock, Gunderson, Ramirez, & Levine, 2010) and can make science instruction very challenging (Beilock & Willingham, 2014).

Perhaps the most complex application of content-specific education quality metrics to systems and districts lies in discussions about autonomy. Evidence indicates that providing teachers with the autonomy to design their own science and mathematics lessons is less advantageous than the well-supported enactment of well-designed instructional materials (Banilower et al., 2006) Penuel, Gallagher, & Moorthy, 2011). Some data show that efforts to increase autonomy for principals, particularly on issues of budgeting and curriculum choice, have a positive impact on reading performance but no effect on mathematics (Steinberg, 2014), which may be because "while school leaders understood the expertise for leading change in literacy instruction to be home-grown, the expertise for leading change in mathematics was outside the schoolhouse" (Stein & Spillane, 2005). A similar circumstance no doubt exists for science, technology, and engineering. If knowledge for leadership in the STEM disciplines is distinct and is underdeveloped in most educators, policies that promote autonomy for decisions associated with STEM issues will likely be unsuccessful.

Making Use of STEM Indicators

A set of national K-12 STEM education indicators, as envisioned by the Monitoring Progress report, has the potential to help districts and states improve their practice and get better results for three reasons.

First, having national data on a shared set of STEM indicators could open the "black box" of system functioning. By black box, I mean essentially the same phenomenon Black and Wiliam (1998) referred to in discussing the black box of the classroom but at the system level: a component defined by a series of inputs (such as students, resources, communities) and outputs (better educated students) without delineation of the myriad of mediating processes that occur in between (e.g., teaching and learning). Many scholars have attempted to describe educational systems via their components and connecting relationships, developing state-level models (Smith & O'Day, 1990), district-level models (Stein & Coburn, 2007), and descriptions of the mechanisms of management (Resnick, Besterfield-Sacre, Mehalik, Sherer, & Halverson, 2007). A K-12 STEM Education Indicator System, as described in *Monitoring Progress,* could focus attention on critically important but commonly neglected aspects of systemic improvement that occur within the district black box, shining a spotlight on them and enabling research that tests their impacts on desired outcomes.

Second, a shared set of STEM education indicators could also provide politically acceptable justification for difficult decisions about education resources and policies, helping redirect efforts to practices that are well supported by the research literature and that require supporting policies and resources to enact.

Finally, a set of national STEM indicators would help encourage more research on answering questions about improvement at scale, which is particularly important for system leaders, and less research that attempts to answer questions of interest mostly to researchers.

Let's examine three of the *Monitoring Progress* indicators to show how they might help system leaders drive additional improvements in STEM teaching, learning, and leading. In each of these cases, I describe how such an indicator could open the black box of system functioning, provide politically acceptable justification for difficult decisions, and set an agenda for future research.

Case 1: STEM instructional materials (Indicator 4)

An indicator for instructional materials has the potential to at least partially open the black box of system improvement in some powerful ways. *Monitoring Progress* defines indicator 4 as the "adoption of instructional materials in grades K-12 that embody the Common Core State Standards for Mathematics and the Framework for K-12 Science Education" (Committee on the Evaluation Framework for Successful K-12 STEM Education, 2013). Note that this indicator attempts to account for both the adoption of particular instructional materials and the degree to which those materials are consistent with the essential elements of new college- and career-ready standards.

Instructional materials—whether traditional textbooks or online collections of lessons and units—are part of the fabric of U.S. schooling (Goodlad, 2004). Yager and Stodghill commented back in 1979 that

> Science in the school program can be characterized by one word—textbooks. The science curriculum exists as the facts and concepts that are

> traditionally packaged in textbooks. The textbook not only determines the content, but the order, and the examples, and the application of that content. (Yager & Stodghill, 1979)

Since textbook content is so closely aligned with classroom coverage, Indicator 4 is closely tied to Indicator 5, defined by *Monitoring Progress* as "classroom coverage of content and practices in rigorous, research-based standards" (Committee on the Evaluation Framework for Successful K-12 STEM Education, 2013). Efforts to improve the quality of mathematics and science curriculum date back to the post-Sputnik era (DeBoer, 1991) and were the subject of major investments by the National Science Foundation in the 1990s (Tushnet et al., 2000). There is considerable evidence that research-based instructional materials, when well implemented, contribute significantly to student improvement gains in mathematics (Post et al., 2008; Sconiers et al., 2003; Wendt & Rice, 2013) and science (Harris, Penuel, DeBarger, D'Angelo, & Gallagher, 2014; Penuel, Gallagher, & Moorthy, 2011). There is also evidence that in general most teachers are not particularly good at designing their own lessons and produce better results for students when they tailor or enact well-designed instructional materials in their classrooms (Banilower et al., 2006). District strategies for the adoption and implementation of robust instructional materials have been shown to lead to sizable student learning gains, especially in large urban school districts (Briars, 2002; Wenzel, 2010).

The strategy underlying research-based curriculum development efforts, in a nutshell, has been to develop high-quality instructional materials that are educative for teachers (E. A. Davis & Krajcik, 2005) as well as for students, thereby enhancing student learning both directly (as they interact with the materials) and indirectly (by building teachers' capacity to interact with their students on the content to be learned). It is difficult for a single teacher or a small group of local teachers to attain the quality and consistency of materials development possible with federal and private funding or when a networked group of teachers collaborate to collectively produce, try out, and refine a set of materials over time.

A similar challenge arises with respect to the professional development needed to accompany new STEM instructional materials geared to today's more challenging science standards (Wilson, 2013). The problem of designing effective professional development is compounded many times over when the school district uses many different sets of science instructional materials in different classrooms. Without common instructional materials, it is inefficient to invest significant resources in professional development on science learning activities, content, and instructional models that are used in only a few classrooms. There is power in both better materials and in using a consistent set of materials throughout the district.

Challenges of Common Adoption as a Path to Coherence and Quality

However, both materials adoption and alignment issues present challenges for schools and districts. Some local education agencies are reluctant to make districtwide adoption decisions, instead leaving the choice to individual teachers or schools. This can result in considerable curricular variance within the district. Whitehurst (2009) bemoaned the lack of attention paid to curriculum-based reforms, commenting that "the effect sizes for curriculum are larger, more certain, and less expensive" (p.9) than those for more popular and politically palatable reform strategies, such as reconstituting the teacher workforce and opening charter schools. At the school and district levels, there is resistance as well based on a desire to let teachers create their own materials in order to increase their sense of professional autonomy (for instance, Ball & Cohen, 1996; Kozol, 2005) and also from a desire for digital materials that are connected or integrated on site (M. R. Davis, 2013).

A teacher-created curriculum strategy should not be adopted without considering equity. Many teachers lack the content and instructional design expertise to create high-quality learning activities and curricula, and teachers who work with low-income students are generally less qualified in this creation than those who do not (for instance, Clotfelter, Ladd, Vigdor, & Wheeler, 2007). Teachers who work with large proportions of low-income students also are often working in more challenging environments. Relative to a unified and coherent approach to STEM instructional materials, a teacher-created curriculum strategy runs the risk of exacerbating both the incoherence and the quality and opportunity gaps within a school or district, making systemic instructional improvement for the high-need students who need it most even more difficult.

Moreover, the STEM education community has not helped administrators much in this regard. The “math wars” highlighted divisions within the math community, as mathematicians and mathematics educators debated the best ways to educate children with “extreme and intemperate dialogue” (Schoenfeld, 2004). To a nonspecialist, like most principals or superintendents who generally have the ultimate procurement authority within their systems, the most reasonable reaction is to keep the controversy at arm’s length: If even the “experts” cannot agree on what is appropriate science or mathematics curriculum and instruction, an administrator’s safest course is to avoid spending money on any new purchase that will invite criticism.

What Is Alignment, Anyway?

If adoption of curriculum materials presents challenges, so does alignment. “Aligned to standards” is a pretty empty phrase—a quick online search at Amazon.com will reveal thousands of products that claim to connect to the Common Core State Standards. While there are tools and procedures that have been developed to determine the degree to which a set of instructional materials is aligned with a set of standards (for instance, DiRanna, Bertrand, & Janulaw, 2000; Porter, 2002; Roseman, Stern, & Koppal, 2010), alignment remains a deceptively difficult and expensive process to conduct. As standards reach higher and involve greater conceptual understanding for students, the challenge increases: Recent attempts have shown that many commercially available materials are not particularly well aligned with the Common Core State Standards (Polikoff, 2014), and efforts to create a teacher-driven Consumer Reports-style resource to gauge the quality of instructional materials remain fraught with challenges (Heitin, 2015).

Achieving alignment is nuanced and labor intensive, and alignment judgments depend on the definitions and processes used by the judges. Attempting to generate school-, district-, or state-level indices of the quality of alignment when curriculum enactment occurs amid the varying contexts at the classroom level is even more problematical.

A national indicator here could provide justification (i.e., much-needed political “cover”) for districts asserting a larger role in orchestrating and supporting districtwide processes for selecting and implementing instructional materials in the manner that most of the research literature recommends.

Such an indicator would signal to local education leaders the unique and important role that instructional materials play for STEM improvement and would direct attention to both adoption of and alignment with standards. In this way, the indicator could drive change. Availability of a "degree of alignment" metric for prominent STEM curriculum products would push the education system away from the popular everyone-creates-his-or-her-own model and encourage districts to make uniform, coherent selections of instructional materials and then support robust, context-sensitive enactment of those materials. Assuming that some relatively transparent process is established to measure the degree of alignment, an alignment indicator would encourage the procurement of higher quality materials. For districts that wanted to go a different route—letting individual teachers create their own materials, for instance—the burden of proof for showing that their strategy will lead to good results for students would be clear. Of course, concerns about scripting lessons and de-professionalizing the practice of teaching would need to be considered in the course of implementation, alongside issues about what professional teaching comprises.

At the state level, a materials-alignment indicator would also have merit. A variety of state models exist for textbook procurement, and while some states have "adoption" or "open territory" models (Silber & Chein, 2014), in most cases schools and districts make final instructional material decisions. When the state of Louisiana launched an ambitious effort to support districts as they made instructional materials decisions,[1] the result was that half the districts chose the same set of (highly rated) instructional materials, a remarkable improvement in large-scale instructional coherence (Cowert, 2015).

Curriculum alignment data would also stimulate research. Certainly, schools and districts want to know which sorts of curricula and instructional materials are most effective and in what contexts. Most likely, a myriad of factors drive overall effectiveness, and so understanding the relative effects of professional development, school leadership, or teacher qualifications would help districts make smarter decisions about issues they regularly confront. Moreover, questions about the implementation of these tools should trigger important research questions: Which students and schools get access to better materials? When does that access occur? How do district and state policies influence the ways teachers and students implement instructional materials?

Case 2: Leadership For Principals (Indicator 8)

The *Monitoring Progress* report defines an indicator related to building principal capacity through "participation in professional development on creating conditions that support STEM learning" (Committee on the Evaluation Framework for Successful K-12 STEM Education, 2013). School leaders have a significant impact on the learning in their schools (Leithwood, Louis, Anderson, & Wahlstrom, 2004), and indeed it is hard to find a great school absent a great leader. Well-regarded products have provided STEM content to school and district leaders (Davidson, Nelson, & Weinberg, 1998), as well as research describing the need for content-specific knowledge for school and system leaders (Halverson, Feinstein, & Neshoulam, 2011; Stein & Nelson, 2003), although content-specific principal professional development remains a low priority for most leadership development efforts.

---

[1] For details, see http://www.louisianabelieves.com/academics/ONLINE-INSTRUCTIONAL-MATERIALS-REVIEWS/curricular-resources-annotated-reviews .

Much of the need for content-specific knowledge at the leadership level builds on the subject-specific issues associated with STEM education improvement more broadly. As Spillane (2005) has commented, mathematics and science improvement expertise tends to initiate from outside the schoolhouse, whereas in literacy the expertise typically resides within the school. As such, this indicator is related to indicator 7, "teachers' participation in STEM-specific professional development activities" (Committee on the Evaluation Framework for Successful K-12 STEM Education, 2013, p.25). But while sizable improvements in student learning can be generated from interventions external to the school, sustaining those improvements nearly always depends on the principal and school leaders understanding the content-specific decisions they need to make to sustain the progress.

This phenomenon can be illustrated by the experience of the Chicago Math and Science Initiative in K-8 mathematics. This initiative demonstrated significant year-over-year improvement in elementary mathematics achievement by using robust instructional materials, focused professional development, and some in-school coaching, but the rate of improvement decreased with successive years (Deiger et al., 2009; Lach, 2006). One interpretation of these outcomes is that teachers could learn the mechanics of the new instructional materials via externally driven supports, and since the design of those curriculum packages was somewhat robust, student learning improved. However, deeper implementation—where teachers were thoughtfully enacting and adjusting based on their students and their school context—required teaming routines and local analysis activities that depend on high-capacity school leadership. Creating such conditions at scale was not possible because of various system characteristics, so curriculum implementation and the rate of improvement in successive years leveled off after the initial success. Similar challenges arose in related efforts to improve student learning at the high school level, with evaluators commenting that "lacking knowledge of the new curriculum and teaching strategies, [principals] were unable to give full support to their teachers" (Humphrey & Shields, 2009).

Principals have complicated jobs with many demands on their time (Horng, Klasik, & Loeb, 2009). Many have called for principals to have more "control over the learning process" (Bottoms & Fry, 2009), yet given both the need for STEM-specific improvement strategies and the controversies and ambivalences around STEM education discussed above, it is unlikely that principals and school leaders will learn how to support STEM learning in their schools on their own. A national indicator that addressed principals' opportunity to learn about effective STEM teaching and learning practices from a leadership perspective would draw attention to the critical need in most districts and schools and would give district leaders and state policymakers an additional impetus to focus on this aspect as they work to increase the capacity of their principals and assistant principals.

Such an indicator also has the potential to spur research and development on the best ways to design and implement such professional development. Professional development for school leaders could be connected to efforts to create improved organizational structures and healthy climates for STEM instruction, with school leaders working with one another and with researchers to investigate alternative school configurations, teacher teaming structures, and tools for classroom observations and providing feedback to teachers.

Case 3: Time Allocated to Teach Science (Indicator 2)

Schools and school systems have been challenged to make science a focus of improvement efforts for some time (Dewey, 1910), particularly in the NCLB era. A thorough study of science education in California showed that “little time is devoted to science learning” particularly at the K-5 level (Rene Dorph, Shields, Tiffany-Morales, & McCaffrey, 2011), and this finding has been replicated in broader national surveys (Trygstad, Smith, Banilower, & Nelson, 2013). This decline in the time spent teaching science in elementary school has occurred despite the evidence that science achievement is, to a significant extent, a function of the time spent teaching science (Blank, 2013).

For a school or district leader, increasing the time allocated to science at the elementary level is challenging. The disciplines of English language arts and mathematics took precedence in NCLB accountability systems. Despite efforts to show how learning science can help improve literacy (Greenleaf et al., 2011; Pearson, Moje, & Greenleaf, 2010; Snow, 2010), the prevalent literacy-focused culture within elementary schools, coupled with the lack of scientific expertise at this level, makes change difficult. Moreover, evidence from the Bay Area (Rena Dorph et al., 2007) and California (Rene Dorph et al., 2011) shows that teachers perceive that they spend less time teaching science than administrators think they do, implying that efforts to increase the time spent on science from the top do not always produce that result.

A national indicator that reports the amount of time allocated to teaching science K-5 would help districts and schools readjust the balance of instructional time across subjects. When coupled with the science performance reporting rules (part of NCLB and presumably continuing in ESSA), such an indicator would make it easier to apply pressure to increase the time spent on science, and it could become a rallying cry for current elementary science education experts, giving them the reach and platform to advocate for their own capacities and ability to help move the student learning forward. Of course, pressure from the top has the potential to drive increased reporting of time spent teaching science with little change in actual instruction. For this reason, it may be better not to attach high stakes to teachers’ reports of the amount of time they spend on science per se.

An instructional time indicator would also help direct attention to important research questions. For instance, how much science learning time is needed for students to achieve competence in the disciplinary core ideas, cross-cutting themes, and science and engineering practices in the new science standards by the end of grade 5? Holding the amount of science instructional time constant, what sorts of teacher and school factors contribute most to the effectiveness of that time?

Putting STEM Indicators into Practice

A common set of STEM education indicators, such as those described in *Measuring Progress*, has the potential to focus improvement efforts, elevating education leaders' awareness of these issues (putting them "on the dashboard") and thus the likelihood of their addressing them. Such indicators would be a means to improvement, not the ends in themselves. The evidence base varies for the different indicators—we know more about what it takes to measure and change some of them than others—and the *Monitoring Progress* report calls for more education research spending as well. But for state and district leaders, the greatest advantages of such an indicator system most likely would be the attention it would bring to certain vexing problems of practice, both to highlight the complexities of the system and to provide political cover for difficult decisions, as well as to drive additional research at the local, state, and national levels. As research focuses more on the problems that schools and districts care about and less on the problems needed to generate publications and favorable tenure reviews, those in the K–12 education system will probably learn more about what it would take to deliver the sorts of STEM learning at scale that *Measuring Progress* envisions.

Common metrics are an important initial step in improving STEM education nationwide, but the STEM indicators are certainly only a piece of the puzzle. To achieve sustained, long-term improvement—indeed, the only real kind of improvement—schools and districts need much more. In a research-practitioner partnership between Fresno and Long Beach, California, districts, researchers identified additional touchstones—including mechanisms and routines to use data, strong relationships and trust between participants, and a culture of evidence-based practice—that helped drive and sustain practice (California Collaborative on District Reform, 2012). Similarly, efforts to improve mathematics and science instruction in the Chicago Public Schools were grounded in broad community support (particularly from local universities), frequent revision and refinement of strategies and efforts, and attention to data and evaluation (Deiger et al., 2009). Similarly, a recent review of the Sanger, California, school district's improvement efforts highlighted the focus on a long-term view coupled with a robust culture and healthy use of evidence to make dramatic improvements (David & Talbert, 2013). Educational systems are incredibly complex, and while STEM indicators can help leaders treat parts of the system less as a black box and more as a critical set of components to be designed and managed, the complexity remains. Evidence-based justification is useful for getting agencies and organizations to do something but is insufficient to secure quality.

America's challenges with STEM education are numerous, and they will not be solved overnight. These examples of STEM education indicators point to a thoughtful path forward: robust indicators that focus on what matters, coupled with schools and districts that have the internal trust and management mechanisms to enable ongoing learning and iteration based on indicator data over time. In a context of multiple participants collaborating to drive improvement, the indicators described in *Monitoring Progress* can be an important step forward.

References

Ahmed-Ullah, N. S., & Byrne, J. (2014, August 27). Chicago Public Schools reports graduation rate up. *Chicago Tribune.* Retrieved from http://www.chicagotribune.com/news/ct-cps-graduation-rate-met-0827-20140827-story.html

Aldeman, C. (2015). On college signing day, a look at FAFSA completion rates. Retrieved from http://aheadoftheheard.org/on-college-signing-day-a-look-at-fafsa-completion-rates/

Allensworth, E., Correa, M., & Ponisciak, S. (2008). *ACT Preparation: Too much, too late.* Retrieved from Chicago, IL: Consortium on Chicago School Research.

Allensworth, E., & Easton, J. Q. (2005). *The on-track indicator as a predictor of high school graduation.* Retrieved from Chicago, IL: http://ccsr.uchicago.edu/content/publications.php?pub_id=10

Allensworth, E. M., & Easton, J. Q. (2007). *What matters for staying on-track and graduating in Chicago Public High Schools.* Retrieved from Chicago, IL: http://ccsr.uchicago.edu/sites/default/files/publications/07 What Matters Final.pdf

Allensworth, E. M., Gwynne, J. A., Moore, P., & Torre, M. d. l. (2014). *Looking forward to high school and college: Middle grade indicators of readiness in Chicago Public Schools.* Retrieved from Chicago, IL: https://ccsr.uchicago.edu/sites/default/files/publications/Middle%20Grades%20Report.pdf

Ayers, J., Owen, I., Partee, G., & Chang, T. (2012). *No Child Left Behind waivers: Promising ideas from second round applications.* Retrieved from Washington, DC: https://www.americanprogress.org/wp-content/uploads/issues/2012/07/pdf/nochildwaivers.pdf

Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher, 25*(9), 6-8, 14.

Banilower, E. R., Boyd, S. E., Pasley, J. D., & Weiss, I. R. (2006). *Lessons from a decade of mathematics and science reform: A capstone report for the Local Systemic Change through Teacher Enhancement Initiative.* Retrieved from Chapel Hill, NC: http://www.pdmathsci.net/findings/report/32

Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 National Survey of Science and Mathematics Education.* Retrieved from Chapel Hill, NC: http://www.horizon-research.com/2012nssme/wp-content/uploads/2013/02/2012-NSSME-Full-Report-updated-11-13-13.pdf

Barber, M., Moffit, A., & Kihn, P. (2010). *Deliverology 101: A field guide for educational leaders.* Thousand Oaks, CA: Corwin.

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences.* Retrieved from http://www.pnas.org/cgi/doi/10.1073/pnas.0910967107

Beilock, S. L., & Willingham, D. T. (2014). Math anxiety: Can teachers help students reduce it? *American Educator 38*(2) (Summer 2014), 28-32, 43.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-144, 146-148.

Blank, R. K. (2013). Science instructional time is declining in elementary schools: What are the implications for student achievement and closing the gap? *Science Education, 97*(6), 830–847. doi:10.1002/sce.21078

Bottoms, G., & Fry, B. (2009). *The district leadership* *challenge: Empowering principals to improve teaching and learning.* Retrieved from http://publications.sreb.org/2009/09V11_District_Leadership_Challenge_color.pdf

Briars, D. (2002, March). *The Pittsburgh story: Successes and challenges in implementing a standards-based mathematics program.* Paper presented at the Success For Everyone Conference, University of Illinois at Chicago.

Burke, A. (2015). *Early identification of high school graduation outcomes in Oregon Leadership Network schools* (REL 2015–079). Retrieved from Washington, DC: http://ies.ed.gov/ncee/edlabs/regions/northwest/pdf/REL_2015079.pdf

C-STEMEC. (2013). *Putting it all together: Supporting K-12 STEM education in Illinois*. Retrieved from http://c-stemec.org/wp-content/uploads/2013/10/putting_it_all_together.pdf

California Collaborative on District Reform. (2012). Special Series on the Fresno–Long Beach learning partnership: Overview. Retrieved from http://cacollaborative.org/sites/default/files/CA_Collaborative_Fresno_LB_Overview_1.pdf

Casserly, M., & Carlson, R. (2014). *Managing for results in America's great city schools 2014.* Retrieved from http://files.eric.ed.gov/fulltext/ED496477.pdf

Chicago Public Schools. (2011). CPS students post record FAFSA completion rate [Press release]. Retrieved from http://cps.edu/News/Press_releases/Pages/02_28_2011_PR1.aspx

Chicago Public Schools. (2015). Citywide cohort dropout and graduation rates. Retrieved from http://www.cps.edu/Performance/Documents/Datafiles/website_cohort_5Yr_city_wide_1999through2014.xls

City, E. A., Elmore, R. F., Fiarman, S. E., & Teitel, L. (2009). *Instructional rounds in education: A network approach to improving teaching and learning.* Cambridge, MA: Harvard Education Press.

Clotfelter, C., Ladd, H. F., Vigdor, J., & Wheeler, J. (2007). *High-poverty schools and the distribution of teachers and principals.* Retrieved from Washington, DC: http://www.caldercenter.org/sites/default/files/1001057_High_Poverty.pdf

Committee on Prospering in the Global Economy of the 21st Century. (2006). *Rising above the gathering storm: Energizing and employing America for a brighter economic future.* Washington, DC: National Academies Press.

Committee on the Evaluation Framework for Successful K-12 STEM Education. (2013). *Monitoring progress toward successful K-12 STEM education: A nation advancing?* Retrieved from Washington, DC: http://www.nap.edu/catalog/13509/monitoring-progress-toward-successful-k-12-stem-education-a-nation?utm_expid=4418042-5.krRTDpXJQISoXLpdo-1Ynw.0

Cortes, K., Goodman, J., & Nomi, T. (2013). *Intensive math instruction and educational attainment: Long-run impacts of double-dose algebra.* Retrieved from Cambridge, MA: https://research.hks.harvard.edu/publications/getFile.aspx?Id=928

Cowert, J. (2015, September 11). [IMR Discussion with Michael Lach and Jill Cowert].

Cuban, L. (2013). *Inside the black box of classroom practice: Change without reform in American education.* Cambridge, MA: Harvard Education Press.

Daro, P. (2007). Systematic catch-up for Middle School students. *Welcome to the toolkit: resources for leaders in mathematics education.* Retrieved from http://toolkitforchange.org/toolkit/view.php?obj=1128&link=429

David, J. L., & Talbert, J. E. (2013). *Turning around a high-poverty district: Lessons from Sanger.* Retrieved from San Francisco, CA: http://web.stanford.edu/group/suse-crc/cgi-bin/drupal/sites/default/files/Sanger%20Turnaround%2010-14-12.pdf

Davidson, E., Nelson, B. S., & Weinberg, A. (1998). *Lenses on learning.* New York, NY: Educational Development Center, Inc.

Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher, 34*(3), 3-14.

Davis, M. R. (2013, February 6). "Big three" publishers rethink K-12 strategies. *Education Week.* Retrieved from http://www.edweek.org/dd/articles/2013/02/06/02textbooks.h06.html

DeBoer, G. (1991). *A history of ideas in science education: Implications for practice*. New York, NY: Teachers College Press.

Deiger, M., Fendt, C., Feranchak, B., Hurtig, J., Shefner, R., Slavsky, D., & Wenzel, S. A. (2009). *Math and science education systemic reform in Chicago, 2002–2008.* Retrieved from Chicago, IL: http://www.luc.edu/media/lucedu/cse/programs/pdfs/White_Paper_MASTER_012609.pdf

Dewey, J. (1910). Science as subject-matter and as method. *Science, 31*(787), 121-127.

DiRanna, K., Bertrand, C., & Janulaw, A. (2000). *Selecting and purchasing science instructional materials* (K. DiRanna, C. Bertrand, & A. Janulaw Eds.). Sacramento, CA: California Science Teachers Association.

Dorph, R., Goldstein, D., Lee, S., Lepori, K., Schneider, S., & Venkatesan, S. (2007). The status of science education in the Bay Area: Research study. *Unpublished manuscript, University of California, Berkeley.*

Dorph, R., Shields, P., Tiffany-Morales, J., & McCaffrey, T. (2011). *High hopes, few opportunities: The status of elementary science education in California*. Retrieved from Research study e-report. Berkeley: Lawrence Hall of Science, University of California.

Education Commission Of The States. (2013). State school accountability "report cards." Retrieved from http://ecs.force.com/mbdata/MBquest3RT?Rep=ar10

Elmore, R. (2004). *School reform from the inside out: Policy, practice, and performance.* Cambridge, MA: Harvard Education Press.

Elmore, R. F. (2002). *Bridging the gap between standards and achievement: The imperative for professional development in education*. Retrieved from http://www.gtlcenter.org/sites/default/files/docs/pa/3_PDPartnershipsandStandards/TheImperativeforPD.pdf

Fendt, C., Harris, R., Lent, J., & Wenzel, S. (2008). *CMSI high school Algebra I for middle grade students: Logistics, students, instruction, and teachers and their preparation, 2007-2008.*

Fulton, J. (2014, October 24). *The Chicago Algebra Initiative: Expanding access & supporting teacher development.* Paper presented at the Council of the Great City Schools 58th Annual Fall Conference, Long Beach, CA.

Gill, B., Borden, B. C., & Hallgren, K. (2014). *A conceptual framework for data-driven decision making.* Retrieved from https://www.disabilitypolicyresearch.org/~/media/publications/pdfs/education/framework_data-driven_decision_making.pdf

Goodlad, J. (2004). *A place called school: Twentieth anniversary edition* (2nd ed.). New York, NY: McGraw Hill.

Greenleaf, C. L., Litman, C., Hanson, T. L., Rosen, R., Boscardin, C. K., Herman, J., . . . Jones, B. (2011). Integrating literacy and science in biology: Teaching and learning impacts of reading apprenticeship professional development. *American Educational Research Journal, 48*(3), 647–717. doi:10.3102/0002831210384839

Grissom, J. A., & Andersen, S. (2012). Why superintendents turn over. *American Educational Research Journal, 49*(6), 1146–1180. doi:10.3102/0002831212462622

Halverson, R., Feinstein, N. R., & Neshoulam, D. (2011). School leadership for science education. In G. E. DeBoer (Ed.), *The Role of Public Policy In K-12 Science Education* (pp. 397-431). Charlotte, NC: Information Age Publishing.

Harris, C. J., Penuel, W. R., DeBarger, A. H., D'Angelo, C., & Gallagher, L. P. (2014). *Curriculum materials make a difference for next generation science learning: Results from Year 1 of a randomized control trial.* Retrieved from https://www.sri.com/sites/default/files/publications/pbis-efficacy-study-y1-outcomes-report-2014_0.pdf

Heitin, L. (2015, March 17). Review of math programs comes under fire. *Education Week.* Retrieved from http://www.edweek.org/ew/articles/2015/03/18/review-of-math-programs-comes-under-fire.html

Horng, E. L., Klasik, D., & Loeb, S. (2009). *Principal time-use and school effectiveness*. Retrieved from Palo Alto, CA: http://web.stanford.edu/~sloeb/papers/Principal Time-Use (revised).pdf

Humphrey, D. C., & Shields, P. M. (2009). *High school reform in Chicago Public Schools: An overview*. Retrieved from https://ccsr.uchicago.edu/sites/default/files/publications/Overview.pdf

Illinois State Board Of Education. (2015). Illinois Report Card. Retrieved from http://illinoisreportcard.com/

Jabon, D., Narasimhan, L., Boller, J., Sally, P., Baldwin, J., & Slaughter, R. (2010). The Chicago Algebra Initiative. *Doceamus: Notices of the American Mathematical Society* (August 2010), 865-867. Retrieved from http://www.ams.org/notices/201007/rtx100700865p.pdf

Jennings, J., & Rentner, D. S. (2006). Ten big effects of the No Child Left Behind Act on public schools. *Phi Delta Kappan,* 110-113. doi:10.1177/003172170608800206

Johnson, J., Arumi, A. M., & Ott, A. (2006). *Reality check 2006: The insiders: How principals and superintendents see public education today.* Retrieved from http://www.publicagenda.org/research/pdfs/rc0604.pdf

Judson, E. (2012). When science counts as much as reading and mathematics: An examination of differing state accountability policies. *Education Policy Analysis Archives, 20*(26).

Kemple, J. J., Segeritz, M. D., & Stephenson, N. (2013). Building on-track indicators for high school graduation and college readiness: Evidence from New York City. *Journal of Education for Students Placed at Risk, 18*(1), 7-28. doi:10.1080/10824669.2013.747945

Kozol, J. (2005). *The shame of the nation.* New York, NY: Random House.

Lach, M. (2006). *K-8 mathematics performance update.* Chicago, IL: Chicago Public Schools.

Ladd, H. F. (2007). *Holding schools accountable revisited.* Paper presented at the Association for Public Policy Analysis and Management. http://www.appam.org/assets/1/7/Holding_Schools_Accountable_Revisited.pdf

Lauen, D. L., & Gaddis, S. M. (2012). Shining a light or fumbling in the dark? The effects of NCLB's subgroup-specific accountability on student achievement. *Educational Evaluation and Policy Analysis, 34*(2), 185–208. doi:10.3102/0162373711429989

Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *Review of research: How leadership influences student learning.* Retrieved from http://www.wallacefoundation.org/knowledge-center/school-leadership/key-research/documents/how-leadership-influences-student-learning.pdf

Mikulecky, M., & Christie, K. (2014). *Rating states, grading schools: What parents and experts say states should consider to make school accountability systems meaningful.* Retrieved from Denver, CO: http://www.ecs.org/docs/rating-states,grading-schools.pdf

Mooney, C., & Kirshenbaum, S. (2010). *Unscientific America: How scientific illiteracy threatens our future.* New York, NY: Basic Books.

Mourshed, M., Chijioke, C., & Barber, M. (2010). *How the world's most improved school systems keep getting better.* Retrieved from http://www.mckinsey.com/client_service/social_sector/latest_thinking/worlds_most_improved_schools

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8.* Committee on Science Learning, Kindergarten Through Eighth Grade. Richard A. Duschl, Heidi A. Schweingruber, and Andrew W. Shouse, Eds., Board on Science Education, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press. Retrieved from http://www.nap.edu/read/11625/chapter/10#213

Neal, D., & Schanzenbach, D. W. (2007). Left behind by design: proficiency counts and test-based accountability. *National Bureau of Economic Research Working Paper* (13293). Retrieved from http://www.nber.org/papers/w13293.pdf

Network for College Success. (2015). On-track to graduation and college readiness. Retrieved from https://ncs.uchicago.edu/page/track-graduation-and-college-readiness

No Child Left Behind Act of 2001, H. R. 1, United States Congress (2001).

Nomi, T., & Allensworth, E. (2012). Sorting and supporting: why double-dose algebra led to better test scores but more course failures. *American Educational Research Journal, 50*(4), 756-788. doi:10.3102/0002831212469997

Nomi, T., & Allensworth, E. M. (2014). *Skill-based sorting in the era of college prep for all: costs and benefits.* Retrieved from Chicago, IL: https://ccsr.uchicago.edu/sites/default/files/publications/Sorting%20Brief.pdf

Payne, C. M. (2008). *So much reform, so little change: The persistence of failure in urban schools.* Cambridge, MA: Harvard Education Press.

Pearson, P. D., Moje, E., & Greenleaf, C. (2010). Literacy and science: each in the service of the other. *Science, 328.*

Pederson, P. V. (2007). What is measured is treasured: The impact of No Child Left Behind on nonassessed subjects. The Clearing House: *A Journal of Educational Strategies, Issues and Ideas, 80*(6), 287-291. doi:10.3200/TCHS.80.6.287-291

Penuel, W. R., Gallagher, L. P., & Moorthy, S. (2011). Preparing teachers to design sequences of instruction in earth systems science: A comparison of three professional development programs. *American Educational Research Journal, 48*(4), 996–1025. doi:10.3102/0002831211410864

Polikoff, M. S. (2014). *How well aligned are textbooks to the common core standards in mathematics?* Paper presented at the Annual Conference of the Association for Education Finance and Policy, San Antonio, TX.

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher, 31*(7), 3-14.

Post, T. R., Harwell, M. R., Davis, J. D., Maeda, Y., Cutler, A., Andersen, E. Norman, K. W. (2008). Standards-based mathematics curricula and middle-grades students' performance on standardized achievement tests. *Journal For Research In Mathematics Education, 39*(2), 184–212.

Resnick, L., Besterfield-Sacre, M., Mehalik, M., Sherer, J. Z., & Halverson, E. (2007). A framework for effective management of school system performance. *Yearbook of the National Society for the Study of Education, 106*(1), 155–185. doi:10.1111/j.1744-7984.2007.00101.x

Roderick, M., Nagaoka, J., Coca, V., Moeller, E., Roddie, K., Gilliam, J., & Patton, D. (2008). *From high school to the future: potholes on the road to college.* Retrieved from Chicago, IL: https://ccsr.uchicago.edu/publications/high-school-future-potholes-road-college

Roseman, J. E., Stern, L., & Koppal, M. (2010). A method for analyzing the coherence of high school biology textbooks. *Journal of Research in Science Education, 47*(1), 47-70.

Schmidt, R. A. (2009). *High school algebra in the middle grades: examining ninth grade outcomes in the 2007-2008 school year.* CPS internal evaluation report (May 2009). Retrieved from http://research.cps.k12.il.us/cps/accountweb/Research

Schoenfeld, A. H. (2004). The math wars. *Educational Policy, 18*(1), 253-286. doi:10.1177/0895904803260042

Sconiers, S., Isaacs, A., Higgins, T., McBride, J., & Kelso, K. R. (2003). *The ARC Center Tri-State Student Achievement Study.* Retrieved from Lexington, MA: http://www.whiteplainspublicschools.org/cms/lib5/NY01000029/Centricity/Domain/35/ARC_Center_Tri-State_Achievement_Study.pdf

Silber, J. M., & Chein, H. S. (2014). *Education And Training.*

Smith, M. S., & O'Day, J. (1990). Systemic school reform. Politics of Education Association Yearbook. (233-267).

Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science, 328(*5977), 450-452.

Spillane, J. P. (2005). Primary school leadership practice: How the subject matters. *School Leadership and Management, 25*(4), 383-397. doi:10.1080/13634230500197231

Spillane, J. P. (2015). Broadening the educational capability conversation: leveraging the social dimension. Retrieved from http://www.shankerinstitute.org/blog/broadening-educational-capability-conversation-leveraging-social-dimension

Stein, M. K., & Coburn, C. (2007). *Architectures for learning: a comparative analysis of two urban school districts.* Retrieved from Seattle, WA: http://www.jstor.org/stable/pdfplus/10.1086/589315.pdf?acceptTC=true&jpdConfirm=true

Stein, M. K., & Nelson, B. S. (2003). Leadership content knowledge. *Educational Evaluation and Policy Analysis, 25*(4), 423–448. doi:10.3102/01623737025004423

Stein, M. K., & Spillane, J. (2005). What can researchers on educational leadership learn from research on teaching? Building a bridge. In W. A. Firestone & C. Riehl (Eds.), *A new agenda for research in educational leadership: Critical issues in educational leadership* (pp. 28-45). New York, NY: Teachers College Press.

Steinberg, M. P. (2014). Does greater autonomy improve school performance? Evidence from a regression discontinuity analysis in chicago. *Education Finance And Policy, 9(*1), 1-35. doi:10.1162/EDFP_a_00118

Stigler, J., & Hiebert, J. (1997). Understanding and improving classroom mathematics instruction: an overview of the TIMSS Video Study. *Phi Delta Kappan,* 14-21.

Stoelinga, T., & Lynn, J. (2013). *Algebra And The Unprepared Learner.* Retrieved from Chicago, IL: http://c-stemec.org/wp-content/uploads/2013/08/Algebra-and-Underprepared-Learner.pdf

The National Commission on Mathematics and Science Teaching for the 21st Century. (2000). *Before it's too late: A report to the nation from the National Commission on Mathematics and Science Teaching for the 21st Century.* Education Publications Center.

Trygstad, P. J., Smith, P. S., Banilower, E. R., & Nelson, M. M. (2013). *The status of elementary science education: Are we ready for the Next Generation Science Standards?* Retrieved from http://www.horizon-research.com/horizonresearchwp/wp-content/uploads/2013/12/The-Status-of-Elementary-Science-Education_paper.pdf

Tushnet, N. C., Millsap, M. A., Abdullah-Welsh, N., Brigham, N., Cooley, E., Elliott, J. Rosenblum, S. (2000). *Final report on the evaluation of the National Science Foundation's instructional materials development program.* Retrieved from Washington, DC: http://www.nsf.gov/pubs/2000/nsf0071/nsf0071.pdf

Wendt, S., & Rice, J. (2013). *Evaluation of ST Math in the Los Angeles Unified School District.* Retrieved from San Francisco, CA: http://www.wested.org/wp-content/files_mf/1371188648resource1314.pdf

Wenzel, S. (2010). Systemic reform of math and science education in chicago: an interactive web-based resource. Retrieved from http://www.luc.edu/scaleup/index.php

Whitehurst, G. R. (2009). Don't forget curriculum. *Brown Center letters on education* (October 2009). Retrieved from http://www.brookings.edu/research/papers/2009/10/14-curriculum-whitehurst

Wilson, S. (2013). Professional development for science teachers. *Science, 340*(310). doi:10.1126/science.1230725

Yager, R. E., & Stodghill, R. (1979). School science in an age of science. *Educational Leadership, 36*(6), 439-445.

SRI Education™

SRI Education, a division of SRI International, is tackling the most complex issues in education to identify trends, understand outcomes, and guide policy and practice. We work with federal and state agencies, school districts, foundations, nonprofit organizations, and businesses to provide research-based solutions to challenges posed by rapid social, technological and economic change. SRI International is a nonprofit research institute whose innovations have created new industries, extraordinary marketplace value, and lasting benefits to society.

Silicon Valley
(SRI International headquarters)
333 Ravenswood Avenue
Menlo Park, CA 94025
+1.650.859.2000
education@sri.com

Washington, D.C.
1100 Wilson Boulevard, Suite 2800
Arlington, VA 22209
+1.703.524.2053

www.sri.com/education

Stay Connected