STEM Indicators

Measuring the *Monitoring Progress* K-12 STEM Education Indicators: A Road Map

Final Road Map

August 28, 2015

SRI Education™
A DIVISION OF SRI INTERNATIONAL

Authors

Barbara Means, Jessica Mislevy, Thomas Smith, Vanessa Peters, Sarah Nixon Gerard

SRI Education

This material is based upon work supported by the National Science Foundation under Grant NSFDACS13T1274. 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.

SRI Education™

Contents

Introduction

On behalf of the National Science Board (NSB), the National Science Foundation's (NSF) National Center for Science and Engineering Statistics (NCSES) produces *Science and Engineering Indicators* (SEI), a biennial compendium of quantitative data on U.S. and international science and engineering education and productivity. Each SEI report is accompanied by companion pieces in which the NSB focuses on trends that raise important policy concerns.

In 2011, the National Research Council (NRC) released the report *Successful K-12 STEM Education*, which made recommendations for improvement in the areas of students' access to STEM learning, educators' capacity, and funding and policy initiatives. In 2013, a second NRC committee released the report *Monitoring Progress Toward Successful K-12 Education: A Nation Advancing?* calling for a national indicator system that could be used by both policymakers and practitioners to improve STEM education. The report lists 14 Indicators that, if measured regularly, could catalyze improvement. This call has gained policymakers' attention at the highest levels, with Congress directing NSF to begin implementing a progress monitoring system for the indicators. However, some of the Indicators require further conceptual work and preliminary research before a data collection plan can be solidified.

SRI Education, a division of SRI International, is supporting NSF by developing a road map for measuring and reporting on the 14 Indicators. SRI is compiling data on each indicator where available and recommending an agenda to support the collection of enhanced data in the future. The purpose of these activities is to enable the NSB Committee on Science and Engineering Indicators to consider and approve companion reports for upcoming SEI reports addressing the Indicators.

SRI Data Collection and Consulting Activities

To inform the road map, SRI conducted two primary types of data collection and consulting activities, each of which is described in additional detail in Appendix A:

1. convened an advisory group of STEM education experts to guide the indicator system data gathering, collaboration, and reporting efforts and
2. interviewed stakeholders from NSF and other key organizations about available data and desired data uses for the 14 Indicators.

Consistent with the Statement of Work, SRI's project did not include a comprehensive review of the literature on these topics. Instead, we focused on the literature the experts we spoke with identified as most relevant to measurement of the 14 indicators. Thus, this report does not cite all the literature that may support one measurement choice versus another and varies in depth and breadth across the 14 indicators based on the experts' resource recommendations.

Advisory Group Meetings

SRI biannually convenes an advisory group of 12 national experts in STEM education policy and practice, education databases, data analysis, indicator systems, and education staffing and finance to help guide the work.

Specifically,

- SRI hosted an advisory group meeting in October 2013, which included discussions on the project history and scope, the set of 14 indicators recommended in the Monitoring Progress report and their operational definitions and potential data sources, prospective stakeholders to interview concerning the development of the indicator system and interview protocols, and plans for ongoing collaboration and future meetings.
- SRI held a virtual advisory group meeting in April 2014 to review progress to date, solicit feedback on the results on the stakeholder interviews, and receive guidance on next steps, including prioritizing potential activities.
- SRI hosted a second in-person advisory group meeting in October 2014 to collect feedback on the current, near-term, and long-term activities for measuring the indicators outlined in the draft road map and to devise strategies for broadening the impact of NSF's STEM Indicators efforts, including potential topics for concept papers.

Stakeholder Interviews

The research team met with stakeholders from key organizations that collect, report, or use K-12 STEM education data. Specifically,

- SRI organized meetings with key statistical entities—the National Assessment Governing Board, National Center for Education Statistics, and Horizon Research, Inc.—to discuss the feasibility and costs of expanding or modifying data collections to obtain more relevant data for the indicators.
- SRI interviewed experts with experience measuring one or more of the indicators to help create operational definitions for terms that need further definition (e.g., STEM schools), identify what sources of information for each indicator exist or could be adjusted to obtain the necessary information, and determine what additional research might be needed to ensure a robust indicator. One to five researchers were interviewed per indicator.
- SRI interviewed two congressional staffers and three additionally policymakers to ascertain the level and locus of interest in the indicators, determine how to make the indicator information most useful to policymakers, and collect recommendations for engaging policy groups.

Dear Colleague Letter to EAGER Grantees

In addition to SRI's data collection and consulting efforts, NSF issued a Dear Colleague Letter (DCL) to fund a series of EAGER (Early-concept Grants for Exploratory Research) awards to advance knowledge about how to measure the indicators (see Appendix A for a list of EAGER grants NSF awarded in response to the DCL). SRI is hosting convenings of an expanded consultancy group that includes the DCL awardees, additional NSF-funded researchers with indicators-related projects, and other identified stakeholders for the purposes of consensus-building and making connections among the indicators-related work being conducted by SRI and others. Researchers discuss their plans, share progress updates, and collect feedback and advice from colleagues doing related work. The first convening was in October 2014.

To promote coherence across projects, SRI will review findings from the DCL awards on an ongoing basis to inform our recommendations for collecting enhanced data on the indicators in the near and longer term. For example, SRI may refine the proposed operational definition or reprioritize activities for an indicator based on the recommendations stemming from a DCL award.

Priority Recurring Data Collections

Extracting information from existing data collections is one way to help control the costs of developing the indicator system. Data for many of the 14 indicators are—or could be—available through surveys administered by that National Center for Educational Statistics (NCES) or other nationally representative recurring data collections. The table below provides an overview of the recurring data collections we identified as most relevant to the STEM indicators work, including the acronyms for each data collection referenced in this report. Additional details regarding these data collections are provided in Appendix B.

Data Source	Agency	Frequency	Sample Overview
Baccalaureate and Beyond Survey, 2008 Cohort	NCES	Longitudinal	National sample of recent recipients of bachelor's degrees
Common Core of Data (CCD)	NCES	Annual	National census of K-12 schools
Civil Rights Data Collection (CRDC)	NCES	Every 2 years	National census of K-12 schools
Early Childhood Longitudinal Program (ECLS) Kindergarten Class of 2010-11 (ECLS-K:2011)	NCES	Longitudinal	National sample of kindergarten class of 2011
High School Longitudinal Study of 2009 (HSLS)	NCES	Longitudinal	National sample of 9th-grade class of 2009
National Assessment of Educational Progress (NAEP) Math Contextual Variables	NCES	Every 2 years	National sample of 4th-, 8th-, and 12th-graders
NAEP Science Contextual Variables	NCES	Every 4 years	National sample of 4th-, 8th-, and 12th-graders
NAEP Technology and Engineering Literacy (TEL) Contextual Variables	NCES	Every 4 years	National sample of 8th-graders
National Survey of Science and Mathematics Education (NSSME)	Horizon Research	Varies	National sample of K-12 math and science teachers
National Teacher and Principal Survey (NTPS; formerly SASS)	NCES	Every 2 years	National sample of K-12 schools
Programme for International Student Assessment (PISA)	OECD	Every 3 years	International sample of 15-year-olds
Schools and Staffing Survey (SASS)	NCES	Every 4 years	National sample of K-12 schools
Survey of the Use of Funds Under Title II	U.S. Dep't Ed.	Annual	National sample of districts
Teaching and Learning International Survey (TALIS)	OECD	Every 5 years	International sample of teachers
Trends in International Mathematics and Science Study (TIMSS)	NCES	Every 4 years	International sample of 4th- and 8th-graders

OECD = Organisation for Economic Co-Operation and Development

About This Report

As outlined in the Statement of Work, this report summarizes the information generated from SRI's data collection and consulting activities to

1. propose more detailed operational definitions for each of the 14 indicators in the NRC *Monitoring Progress* report,
2. describe the sources and necessary actions for producing the needed indicator data for 2014 SEI and 2016 SEI companion reports, and
3. propose research and data collection needed for indicators that require extensive development.

The recommendations related to each indicator distill the range of options proposed by the *Monitoring Progress* authors, the project advisory group, and the stakeholders we interviewed into those that seem most critical, feasible, and promising at present, in the near term, and in the future. This report is not intended to be a comprehensive research agenda related to the indicator system but instead a description of the research needed to improve the measurement of the 14 indicators.

The audience for this report is NSF program officers, as well as the expanded consultancy group, which consists of the DCL awardees, additional NSF-funded researchers with indicators-related projects, and other identified stakeholders in the STEM Indicators community.

Individual road maps for each indicator are first presented, followed by a table summarizing the road map across the indicator system.

Indicator 1: STEM-Focused Schools

Indicator 1 addresses the number of and enrollment in STEM-focused schools and programs in each district. The intent is to measure the extent to which all students have the opportunity to pursue some kind of focused experience in STEM as a part of their K-12 education.

Priority: High

Although Indicator 1 was not identified as a priority indicator in the *Monitoring Progress* report, we recommend moving the priority of the STEM school portion of this indicator (vs. programs) to high on the basis of the recommendation to increase the number of STEM-focused schools in the recent *5-Year Federal Science, Technology, Engineering, and Mathematics (STEM) Education Strategic Plan* produced by the Administration (National Science and Technology Council, 2013). In addition, this indicator most closely aligns with Congressman Wolf's initial request to NSF in 2010 to identify successful STEM schools. A few modest additions to NCES survey items would capture information on this indicator, supporting it as a near-term action.

Operational Definition

SRI proposes to define a STEM-focused school as a school that (1) identifies science, technology, engineering, and/or mathematics as a special emphasis and (2) provides *all* of its students with a more intensive program of instruction in STEM subjects than is required by its district and state. As described in the *NRC Successful K-12 STEM Education* report, there are multiple types of STEM-focused schools that meet this definition. Selective STEM-focused schools can be defined as STEM-focused schools admitting students through a competitive process, usually requiring an entrance examination or using a combination of test scores and prior achievement. Inclusive STEM-focused schools can be defined as STEM-focused schools admitting students primarily on the basis of STEM interest rather than test scores or other measures of prior achievement. At the high school level, selective and inclusive STEM-focused schools share a college preparatory curriculum focus that sets them apart from STEM-focused career technical education schools. We define the latter as high schools with an occupationally focused curriculum designed to prepare students for entry into a STEM-related job, apprenticeship program, or 2-year program of study for a specific career field.

Given the multiple types of STEM-focused schools, we propose capturing Indicator 1 through a series of survey items rather than any single item. First, school leaders can be asked whether or not their school has a special program emphasis or curricular theme that all students participate in and if so, whether it is on science, technology, engineering, or math (or something else). The distinction between inclusive and selective STEM schools can be made using data from existing NCES items on whether or not a school uses test scores in determining admission; researchers can verify reported levels of selectivity by comparing prior achievement scores for a school's incoming students with those of students entering other schools in the same district, region, or state. Among nonselective STEM-focused high schools, we can distinguish between those intending to prepare students for STEM college majors (i.e., inclusive STEM high schools) from those preparing students for industry certification and STEM-related occupations requiring less than a bachelor's degree by examining the school's curriculum. Below, we suggest using the high school's mathematics requirement as a proxy for a college STEM readiness mission.

Currently Available Data

Prior administrations of the CCD, CRDC, and SASS included items on magnet schools and/or schools with a special program emphasis, but it is not possible to identify STEM-focused schools or programs specifically from these data sets because the surveys did not ask about the areas of emphasis. Efforts to examine student achievement data for STEM-focused schools have had to rely in large part on school names and website descriptions, an approach with known deficiencies in terms of distinguishing schoolwide themes from available programs that underserved students may be less likely to participate in.

Recent administrations of the HSLS (2009, 2012) have included a question on the school administrator survey asking whether the school has a magnet program or special focus and whether the focus area is math or science. However, the HSLS currently includes only 10 state-representative samples of 9th-graders and does not distinguish between selective and inclusive STEM-focused schools.

Data source	Administered on a fixed recurring schedule	Produces nationally representative estimate	Produces grade-level estimates	Produces state-level estimates	Allows linking to student achievement data
HSLS09, HSLS12	No	No (representative of 9th-grade students but not schools)	Yes (grades 9 and 11 only)	Yes, but limited to 10 states	Yes

Near-Term Modifications and Activities

To capture the essence of Indicator 1 in the relatively near term, we propose items shown below in italics for possible inclusion on the NTPS. (Those not italicized have been included in the SASS.)

1a. Does the school have a special program emphasis or curricular theme? [Yes/No]

1b. [If so] Is the focus is on science, technology, engineering, or math? [Yes/No]

1c. [If so] Do all students participate in this program or curriculum? [Yes/No]

1d. [If so AND a high school] What is the highest level math course required of all students for graduation? [Select from preset list with write-in option for other]

2a. Does the school use any special requirements when admitting students? [Yes/No]

2b. [If so] What type(s) of special requirements are used to admit students? [Select from preset list]

In the near term, Indicator 1 is well suited to be addressed in future iterations of the NTPS with inclusion of items such as 1a–d and 2a–b. SRI will investigate whether items comparable to the proposed additions have been used on NCES surveys other than the SASS. For STEM high schools, we believe it would be useful also to ask question 1d about the highest level of math all students are expected to take: This question can be used to identify those STEM high schools that are attempting to prepare students for college STEM majors.

If this modification is made, data from the NTPS alone could be used to investigate whether a relationship appears to exist between these kinds of STEM-focused schools and student outcomes. In the interim, this question could be explored using HSLS data. Evidence of a relationship could be used as justification for including these items on the CCD to produce estimates at the state and district level (see Long-Term Modifications and Activities).

Mandinach has received funds through the DCL to determine whether state data collections could address Indicator 1 and what revisions to the state systems would be necessary to do so. This DCL EAGER project may yield additional options.

Additional Research Needs

The NRC report *Successful K-12 STEM Education* called specifically for longitudinal research on the effectiveness of the three models of STEM-focused schools defined above. As research on this topic emerges, the priority that measurement of Indicator 1 deserves will become clearer. Some have suggested that a clearer understanding of the essential elements of each school type is needed in order to measure those elements directly rather than trying to measure school types.

More research is needed addressing the following:

1. How effective are the three models of STEM-focused schools after controlling for characteristics of students who enter these schools? Do different types of STEM-focused schools produce different student outcomes?
2. Are STEM programs within comprehensive schools as effective as STEM-focused schools after controlling for differences in student characteristics?
3. What are the essential features of effective STEM schools and programs?

Long-Term Modifications and Activities

If the near-term activities show relationships between different types of STEM-focused schools and student outcomes, helping to validate the relevance of Indicator 1, we recommend the CCD as the ideal long-term vehicle for measuring this indicator so that states and districts can examine their own data. Adding items to the CCD would entail working with state data coordinators to harmonize definitions and build support for collecting these data at the state level. Building consensus for adding items to the CCD requires an extended process and would most likely require trying out items in a few willing states first before pressing for national data collection.

In addition, there is interest in distinguishing among multiple models of STEM schools characterized by different pedagogies or degrees of occupational specialty. These distinctions will be hard to capture within planned and existing national surveys and are of less immediate priority than obtaining a basic count of STEM schools at different grade levels available to students with different demographic and prior achievement profiles. Century and LaForce have been funded through the DCL to establish a framework for characterizing different kinds of STEM-focused schools and programs. Once it is available, we recommend reviewing their framework to identify any implications for longer term actions that might support future refinements of Indicator 1.

References

Century, J. R., & LaForce, M. (2014). Identifying and measuring STEM schools and programs. National Science Foundation Award Abstract #1445592. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445592

Mandinach, E., & Orland, M. (2014). *An exploration of the alignment of SLDS infrastructure and data highway to relevant success indicators in mathematics and science*. National Science Foundation Award Abstract #1445522. Retrieved from http://nsf.gov/awardsearch/showAward?AWD_ID=1445522

Indicator 2: Time Allocated to Teach Science in Grades K-5

Indicator 2 measures teachers' estimates of the amount of time (instructional minutes per week) that they devote to teaching science.

Priority: High

Indicator 2 is identified as a high priority in the *Monitoring Progress* report. It is both important and feasible to capture in the near term.

Operational Definition

We propose to define this indicator as the number of minutes a week devoted to science instruction at the grade level. So defined, this indicator can be measured using the language in a 2011 TIMSS survey item: time in hours and minutes per week spent on science topics. We recommend this definition because a baseline national estimate for grade 4 is available from the 2011 TIMSS, and this item will be used on the NAEP beginning in 2015. Determining the quality and not just the quantity of science instruction is clearly important, and quality of instruction is the focus of another indicator (Indicator 5). The NRC Monitoring K-12 STEM Education working group framed Indicator 2 in terms of quantity of instructional time per se out of concern that since the advent of No Child Left Behind (NCLB) accountability testing in reading and mathematics, elementary schools have been cutting instructional minutes devoted to science.

Currently Available Data

Data from previous administrations of the NAEP, TIMSS, SASS, and NSSME could be used to provide information on this indicator in the near term.

Data source	Administered on a fixed recurring schedule	Produces grade-level estimates	Produces state-level estimates	Yields write-in responses of time	Allows linking to student achievement data
2011 NAEP Grade 4 Science Teacher Questionnaire	Yes (every 4 years)	Yes (grade 4 only)	Yes, for most states	No, teachers selected among a set range of hours	Yes
2011 TIMSS Grade 4 Teacher Questionnaire	Yes (every 4 years)	Yes (grade 4 only)	No	Yes	Yes
2011-12 SY SASS Teacher Questionnaire	No*	No	Yes*	Yes	No
2011-12 SY SASS Principal Questionnaire	No*	Yes (grade 3 only)	Yes*	Yes	No
2012 NSSME Science Teacher Questionnaire	No	No	No	Yes	No

**The redesigned SASS (now the NTPS) will be administered every 2 years beginning in 2015 and produce national estimates only.*

The currently available data listed in the table above are limited in that they cannot provide comprehensive estimates of time spent teaching science in each elementary grade level. In most cases, data are collected for grade 4 only, which is also the grade at which states most often measure science achievement, most likely leading to more emphasis on science than in other elementary grades. In addition, the NAEP Science survey item about time spent teaching science weekly asked teachers to select among a set of ranges of hours, limiting the analytical power and potentially introducing anchoring effects. Last, existing surveys do not capture science instruction provided to students by other individuals, such as specialists.

Initial recommended target vehicles for collecting data on this indicator are the TIMSS and NAEP because they have most of the desirable characteristics outlined in the table above and allow linking to student achievement data. Although the TIMSS produces results only at the national level, the write-in item offers more analytic power than the 2011 NAEP. Starting in 2015, the 2015 NAEP Science teacher survey item will also be a write-in, identical to TIMSS.

Near-Term Modifications and Activities

Activities and modifications planned for the near term are the following:

- SRI will examine the relationship between teachers' reports of science instruction time from the TIMSS and student achievement to determine whether results are consistent with Blank's (2013), which showed a positive relationship between the amount of class time and student achievement scores in science as measured by the 2009 NAEP Grade 4 Science.
- After conversations with NSF and the project team, the National Assessment Governing Board (NAGB) and NCES have agreed to modify the NAEP survey item for the next administration (2015 Science), replacing the hour-range response options with write-in responses for average weekly science time to match the 2011 TIMSS item.
- Dorph and Hartry have received funds through the DCL to produce recommendations for measuring science instructional time in grades K-5. Mandinach has been awarded funding as a part of the DCL to determine whether existing state data collections could address Indicator 2 and what revisions to the state systems would be necessary to do so. These projects may yield additional options.

Additional Research Needs

The quality of science instructional time and the amount of time spent teaching science in the context of other subjects are broadly recognized as important long-term goals. However, recognizing the difficulties inherent in trying to capture these elements of science instruction, we defined Indicator 2 more narrowly. To address the current, narrowly defined indicator more accurately, we need research on the following in areas:

1. How accurate are teachers' self-reports of time devoted to science instruction? To what extent is the science instructional time provided to students over- or under-represented when relying on teachers' self-reports of instruction? For example, is instructional time double-counted when classes are team taught? Do teacher self-reports miss instruction provided by other individuals, such as science specialists?
2. Building on the work of Dorph et al. (2011), how should time spent on multidisciplinary activities be allocated if one of the disciplines is science? What is the nature of science instruction that takes place during interdisciplinary instruction? What counts as adequate depth when science is combined with other disciplines?

3. If instructional time devoted to science is measured only at selected grade levels (e.g., grade 4), how representative are those data of instructional time at other elementary grade levels?

Long-Term Modifications and Activities

In moving toward capturing quality as well as quantity of science instructional time, there may be additional opportunities for collecting data on this indicator with the NSSME or by proposing additional items addressing the quality of science instructional time (alone and in the context of other subjects) for inclusion on future NAEP Math and Science surveys. A number of our project advisors and DCL PI (principal investigator) meeting participants urged consideration of applying this indicator to middle and high school as well as elementary school.

References

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,* 830-847.

Dorph, R., & Hartry, A. (2014). *Defining, measuring, & monitoring adequate instructional time and resources for science in grades K-5*. National Science Foundation Award Abstract #1445621. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445621

Dorph, R., Shields, P., Tiffany-Morales, J., Hartry, A., & McCaffrey, T. (2011). *High hopes—Few opportunities: The status of elementary science education in California*. Sacramento, CA: The Center for the Future of Teaching and Learning at WestEd.

Mandinach, E., & Orland, M. (2014). *An exploration of the alignment of SLDS infrastructure and data highway to relevant success indicators in mathematics and science*. National Science Foundation Award Abstract #1445522. Retrieved from http://nsf.gov/awardsearch/showAward?AWD_ID=1445522

Indicator 3: Science-Related Learning Opportunities in Elementary Schools

Indicator 3 concerns the range of in-school but nonclassroom science-related learning opportunities that elementary schools may offer, arrange, or help broker.

Priority: Medium

Indicator 3 was not identified as a high priority in the *Monitoring Progress* report, but there is strong interest among researchers and philanthropists in out-of-class STEM learning opportunities as strategies for engaging more students and specifically a more diverse group of students in STEM.

Operational Definition

We propose to define this indicator as activities that occur outside the regular classroom place and/or time that are intended to contribute to students' STEM learning, interest, or engagement. The intention of Indicator 3 is to capture learning opportunities that would not be included in the instructional time reported as a part of Indicator 2. There was some disagreement among advisors and experts as to whether the emphasis should be on measuring those opportunities that are provided by or brokered by schools as opposed to all STEM-related learning opportunities that students experience. Advisors and PI meeting participants recommended measuring this indicator in middle and high school as well as elementary school.

The experts we consulted indicated that important features to measure about STEM-related learning opportunities are

1. whether participation is required,
2. timing of the activity (within or outside the regular school day), and
3. the kinds of resources available to students.

Although a suggestion was that information about this indicator be gathered from school administrators, it was generally agreed that asking students about their STEM learning opportunities would probably be more informative. The experts we consulted also noted that concrete types of experience, such as working in a lab, would be easier to measure accurately via surveys than aspects of experience related to practices or processes (e.g., experiences developing your own line of inquiry). The downside of this approach is that it would require responding to items about many possible specific activities, making data on this indicator less feasible than others to capture on recurring national surveys.

The considerable uncertainty about how to define this indicator operationally and about the best respondent population and data collection strategy means that we are not well positioned to collect national data on this indicator in the near term.

Currently Available Data

Limited data are available from the 2012 NSSME about whether a school offers certain science-related learning opportunities, but the NSSME did not collect information about student participation in available activities. Research by the Tai Research Group (Dabney et al., 2011) included a student survey on participation in out-of-school STEM activities in four districts. Additional data will be available in 2015 from the 2014 NAEP TEL, which has questionnaire items on opportunities provided by schools as well as on out-of-class activities students engage in. Because each source addresses a different aspect of Indicator 3, we recommend collecting and examining data from all three for possible inclusion in an SEI companion report.

Data source	Administered on a fixed recurring schedule	Nationally representative	Produces state-level estimates	Produces grade-level estimates	Measures opportunities school provides	Measures opportunities students engage in	Allows linking to student achievement data
2012 NSSME Science Teacher Questionnaire	No	Yes	No	No	Yes	No	No
Tai Research Group	No	No	No	No	No	Yes	No
2014 NAEP TEL	Yes (every 4 years)	Yes	No	Yes (grade 8 only)	Yes	Yes	Yes

Near-Term Modifications and Activities

Recommendations for near-term modifications activities are as follows:

- Survey responses from Dabney et al. (2011) may provide insight into what is important and feasible to capture with respect to Indicator 3. Questions developed for the HSLS study about high school students' participation in extracurricular STEM activities could be modified or adapted for the elementary grades. The Organisation for Economic Co-operation and Development (OECD) is also developing items for the 2015 PISA that may be relevant to modify or adapt for this indicator.
- Dorph and Hartry have been funded though the DCL to produce recommendations for how to measure the adequacy of resources for science learning in grades K-5 in terms of time, type, and substance for out-of-class as well as in-class activities. We expect that their projects may provide additional options for near-term modifications.
- Near-term data opportunities include the 2019 NAEP Science or TEL survey, but there is a long lead time for adding questions. NCES indicates that items could also be considered for NTPS. These opportunities have associated costs for the development and testing of new items.

Additional Research Needs

More research is needed addressing the following:

1. How can student survey items be developed that are understandable and generate accurate data?
2. What core set of science-related learning opportunities appear to be most important for future STEM learning and career interests?

Long-Term Modifications and Activities

A longer term need is to prioritize appropriate recurring survey vehicles and sets of items for measuring Indicator 3.

References

Dabney, K., Tai, R. H., Almarode, J. T., Miller-Friedman, J. L., Sonnert, G., Sadler, P. M., & Hazari, Z. (2011). Out of school time science activities and their association with career interest in STEM. *International Journal of Science Education*, Part B. doi:10.1080/21548455.2011.629455

Dorph, R., & Hartry, A. (2014). *Defining, measuring, & monitoring adequate instructional time and resources for science in grades K-5*. National Science Foundation Award Abstract #1445621. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445621

Mandinach, E., & Orland, M. (2014). *An exploration of the alignment of SLDS infrastructure and data highway to relevant success indicators in mathematics and science*. National Science Foundation Award Abstract #1445522. Retrieved from http://nsf.gov/awardsearch/showAward?AWD_ID=1445522

Indicator 4: Adoption of Standards-Aligned Instructional Materials

Indicator 4 concerns the adoption of instructional materials in grades K-12 that embody the Common Core State Standards (CCSS) and Science Framework underlying the Next Generation Science Standards (NGSS).

Priority: High

Indicator 4 is identified as a high priority in the *Monitoring Progress* report. This indicator could also have a high impact on student learning for a relatively low cost because research has shown that many teachers closely follow textbooks in their instruction.

Operational Definition

We define Indicator 4 as the extent to which the content and learning activities in a set of instructional materials cover the full content in a set of standards, are consistent with the performance level expectations of those standards, and reflect the same conception of learning and instruction (e.g., the intertwining of disciplinary core ideas, science and engineering practices, and crosscutting concepts in the Science Framework). This definition is much more encompassing than the approach typically used, which merely maps the content in the instructional materials to topics in standards statements for the relevant subject and grade level. The experts we interviewed regard such judgments of relevance as insufficient because they do not reveal the extent to which the full content of each standard is covered in the materials, the extent to which the level of performance expectation in the materials and the standards are the same, or whether the materials are consistent with the standards' conception of learning and instruction. Given the challenges and potential for controversy in making alignment judgments, we recommend forming a team of experts to propose processes for judging the alignment of instructional materials with the CCSS-Mathematics and the Science Framework, consistent with the definition provided above.

Currently Available Data

In previous work, Polikoff used data from the Survey of Enacted Curriculum to investigate the alignment of fourth-grade math textbooks with the CCSS-Mathematics (CCSS-M). These methods could be used to study textbook alignment in other grades and in other states and could possibly be modified to address NGSS alignment in science textbooks. Schmidt is coding textbooks and teacher daily logs to compare textbook coverage and teacher coverage with the CCSS-M. Smithson has developed a process for collecting, processing, and reporting detailed, valid, and replicable descriptions of curricular practice, materials, standards, and assessments.

Data source	Produces grade-level estimates	Produces state-level estimates	Addresses CCSS-M alignment	Addresses NGSS alignment	Includes practices as well as content
Morgan Polikoff, University of Southern California	No (grade 4 only)	No (Florida only)	Yes	No	No
William Schmidt, Michigan State University	Yes	No	Yes	No	Yes
John Smithson, University of Wisconsin-Madison	Yes	No	Yes	Yes	Yes

Near-Term Modifications and Activities

Data from the 2012 NSSME on commonly used instructional materials could be used in a standards-alignment exercise using Polikoff's or others' methods.

Other recommendations for near-term modifications activities are the following:

- Useful information could be obtained from EdReports.org, a nonprofit organization that plans to post online reviews of major textbooks and curricula purportedly aligned with the CCSS-M.
- Information about commonly used materials at selected grade levels could be collected by adding questions to the NAEP Math and Science teacher questionnaires.

Additional Research Needs

Clarifying the scope of “instructional materials” and identifying appropriate data collection methods were recognized as important preliminary work for measuring Indicator 4. Specifically, more research is needed to address the following:

1. Beyond textbooks, what kinds of resources should count as instructional materials in K–12 mathematics and science teaching?
2. What methods are appropriate for analyzing the nontextbook instructional materials used by students (e.g., videos, websites, printed worksheets)?
3. What are the pros and cons of different techniques for measuring the alignment of instructional materials with standards?
4. How can we determine the degree to which different tools and methods produce the same results (i.e., how sensitive alignment conclusions are relative to the measure used)?

A number of researchers from the DCL PI Meeting have advocated for smaller scale research that links instructional materials to student achievement before working to embed qualities aimed at measuring indicators on national surveys. Other researchers, however, felt this has already been accomplished via the Survey of Enacted Curriculum.

Long-Term Modifications and Activities

The following researchers have received DCL awards for Indicator 4, which we expect to yield additional options:

- Gardner and colleagues will convene a summit meeting to develop guidelines and criteria to determine how well widely used STEM instructional materials embody the national standards and to develop evaluation tools and processes for the indicator. Their work will also explore how curriculum developers consider equity when developing standards-aligned curricula.
- Schmidt will convene a group of experts to review the current evidence base for curriculum adoption and contribute to conceptual and operational definitions related to the alignment of instructional materials with national standards.

- Polikoff will gather and make publicly available the largest-ever database of mathematics and science textbook adoption data and analyze the impact of textbook adoption on student achievement in the five largest U.S. states.

Pilot or initial work on this indicator can be used as a time point for starting a trend line of the instructional materials that districts are adopting. As work on this indicator develops, expert groups may have an opportunity to recommend an appropriate schedule for applying techniques for aligning instructional materials with standards (e.g., teacher logs and instructional artifacts do not require routine analysis).

References

Gardner, A., Mohan, A., & Bintz, J. (2014). *Developing consensus guidelines for tools and processes that assess science instructional materials*. National Science Foundation Award Abstract #1445675. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445675

Polikoff, M. (2014, March). *How well aligned are textbooks to the Common Core Standards in Mathematics?* Paper presented at the Annual conference of the Association for Educational Finance and Policy, San Antonio, TX. Retrieved from http://stemindicators.sri.com/archive/resources/Polikoff_AERA_2014.pdf

Polikoff, M. (2014). *An online system for the collection of textbook adoption data*. National Science Foundation Award Abstract #1445654. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445654&HistoricalAwards=false

Schmidt, W. (2014). *Classroom implementation Indicators for K-12 Common Core State Standards in Mathematics (CCSSM)*. National Science Foundation Award Abstract #1445583. Retrieved from http://nsf.gov/awardsearch/showAward?AWD_ID=1445583

Indicator 5: Classroom Coverage of Content and Practices

Indicator 5 concerns the coverage of content and use of practices in the CCSS-M and the NGSS.

Priority: High

Indicator 5 is identified as a high priority in the *Monitoring Progress* report. Like Indicator 4, Indicator 5 could have a considerable impact on student learning.

Operational Definition

We define Indicator 5 as the extent to which the instruction and learning activities students experience in a classroom cover content in a set of standards, are consistent with the performance-level expectations of those standards, and reflect the same conception of learning and instruction (e.g., the intertwining of disciplinary core ideas, science and engineering practices, and crosscutting concepts in the Science Framework). Whereas Indicator 4 captures the intended curriculum, Indicator 5 is meant to capture the enacted curriculum: What CCSS-M and NGSS content and practices are teachers actually implementing in their instruction? Unlike a set of instructional materials or an assessment (the focus of Indicator 12), the entire set of instructional experiences in a class is rarely available in its entirety. For this reason, measurement of Indicator 5 is likely to have to occur at a somewhat coarser grain size than measurement of Indicator 4.

In the near term, we propose using data from teacher self-reports of instructional practices to describe their coverage of content and practices in mathematics and science. This recommendation is based on studies of the Survey of Enacted Curriculum suggesting that teachers can reliably report on their coverage of content during instruction. Alternative methods that correlate self-reports with observational data are lengthy and place a heavy burden on teachers and thus are not easily embedded in an existing multitopic national data collection.

Another approach that is the subject of ongoing research is the analysis of artifacts of classroom instruction (e.g., end-of-unit tests) and samples of student work as a proxy for teachers' coverage of content and practices based on their expectations of what students have learned (e.g., Gitomer & Bell, 2013; Gitomer et al., 2014; Hill, 2005; Newman, Smith, Allensworth, & Bryk, 2001).

Currently Available Data

Self-report data are available from administrations of the NAEP, TIMSS, and ECLS-K related to teachers' coverage of content and instructional practices, but they are not specific to new standards. Existing items would need to be mapped to CCSS-M and NGSS to determine which data would be relevant to include in an SEI companion report. Research on Indicator 4 by Schmidt and Smithson will also inform teachers' coverage of content and practices as outlined in the CCSS-M and NGSS using teacher reports.

Data source	Produces grade-level estimates	Produces state-level estimates	Addresses CCSS-M alignment	Addresses NGSS alignment	Allows linking to student achievement data
2009 NAEP Math and Science Student Questionnaire	Yes (grades 4, 8, and 12)	Yes	Partially	Partially	Yes
2013 Math NAEP Student Questionnaire	Yes (grades 4, 8, and 12)	Yes	Partially	Partially	Yes
2011 TIMSS Math and Science Teacher Questionnaire	Yes (grades 4 and 8)	No	Partially	Partially	Yes
ECLS-K11 Teacher Questionnaire	Yes (grade K only)	No	Partially	Partially	Yes
William Schmidt, Michigan State University	Yes	No	Yes	No	Yes
John Smithson, University of Wisconsin-Madison	Yes	No	Yes	Yes	Yes

Near-Term Modifications and Activities

- Examine the 2009 NAEP, 2011 TIMSS, and ECLS-K to identify items that address the CCSS-M and the NGSS. If items on are identified as relevant to the new standards, they could be used to track changes in instruction to see whether traditional instruction is being replaced with other strategies after the introduction of the new standards.

Additional Research Needs

Developing new items for measuring coverage, including those that measure content in concert with math and science practices, was recognized as an important long-term goal for Indicator 5. More research is needed regarding the following:

1. How can we develop survey items that measure the coverage of content and practices during math and science instruction? How can we assess the reliability and validity of survey reports that measure coverage of content and practices? To what extent can we rely on teacher self-reports to accurately measure their coverage of content and practices during instruction?
2. What is the feasibility of using a nonsurvey approach for collecting data at scale with a nationally representative sample (e.g., Hill, 2005)?
3. Is there value in integrating survey questions into benchmark and formative assessments developed by the Partnership for Assessment of Readiness for College and Careers (PARCC) or the Smarter Balanced Assessment Consortium that ask students whether the content covered by the assessment was also covered by their teachers in class? If so, how feasible would this approach be for measuring this indicator?
4. What implications does students' use of adaptive or personalized learning technologies have for measuring the degree to which the content and practices they are learning embodies the standards? How should this indicator be measured in such contexts when students are exposed to different content and practices?

Long-Term Modifications and Activities

We recommend that the same group of experts who convene to discuss standards-aligned instructional materials (Indicator 4) be involved in the ongoing work of framing improved survey items for Indicator 5. The following researchers have received DCL awards for Indicator 5, which we expect to yield additional options for measuring this indicator in the long-term:

- Banilower and colleagues will convene an expert panel to identify key aspects of science and engineering practices and will develop and validate new survey items for measuring the coverage of content and practices.

- Hamilton and colleagues will explore existing measures and opportunities for collecting evidence about classroom coverage of math and science content and practices in standards for college and career readiness.
- Working with a group of experts, Schmidt will develop mechanisms for measuring the alignment of instructional materials and classroom instruction with key elements of the CCSS-M through extensive use of teacher classroom log data and CCSS-M–coded instructional materials.
- Gitomer will develop protocols that support the use of classroom artifacts (e.g., teacher lesson plans, assignments, and assessments as well as the student work produced from them) as measures of the range and quality of instructional practices, as they relate to CCSS-M and NGSS.

Based on results of these studies, consider proposing to NAGB new standards-related items for the NAEP about content coverage and instructional practices that could be used to track national progress on student learning with CCSS-M and NGSS in selected grade levels (grades 4 and 8).

References

Banilower, E., Pasley, J., & Trygstad, P. (2014). *Operationalizing the science and engineering practices*. National Science Foundation Award Abstract #1445543. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445543

Gitomer, D. (2014). *Classroom artifacts as indicators of quality STEM education*. National Science Foundation Award Abstract #1445632. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445632

Gitomer, D., & Bell, C. (2013). Evaluating teaching and teachers. In K. F. Geisinger (Ed.), *APA handbook of testing and assessment in psychology* (Vol. 3, pp. 415-444). Washington, DC: American Psychological Association.

Gitomer, D. H., Phelps, G., Weren, B., Howell, H., & Croft, A. J. (2014). Evidence on the validity of content knowledge for teaching assessments. In T. J. Kane, K. A. Kerr, & R. C. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the Measures of Effective Teaching project* (pp. 493-528). San Francisco, CA: Jossey-Bass.

Hamilton, L., Stecher, B., & Yuan, K. (2014). *Measuring classroom coverage of content and practices in the new generation of mathematics and science standards*. National Science Foundation Award Abstract #1445670. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445670

Hill, H. (2005). Content across communities: Validating measures of elementary mathematics instruction. *Education Policy, 13*(3), 447–475.

MET Project. (2013, January). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET Project's three-year study.* Retrieved from http://www.metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf

Newman, F., Smith, B., Allensworth, E., & Bryk, A. S. (2001). *School instructional program coherence: Benefits and challenges.* Consortium on Chicago School Research Report. Retrieved from https://ccsr.uchicago.edu/sites/default/files/publications/p0d02.pdf

Schmidt, W. (2014). *Classroom implementation indicators for K-12 Common Core State Standards in Mathematics (CCSSM)*. National Science Foundation Award Abstract #1445583. Retrieved from http://nsf.gov/awardsearch/showAward?AWD_ID=1445583

Schoenfeld, A. H., & Burkhardt, H. (2013). *Content specifications for the summative assessment of the Common Core State Standards for Mathematics*. Retrieved from http://www.smarterbalanced.org/wordpress/wp-content/uploads/2011/12/Math-Content-Specifications.pdf

Indicator 6: Teachers' Science and Mathematics Content Knowledge for Teaching

Indicator 6 addresses teachers' science and mathematics content knowledge for teaching.

Priority: High

Indicator 6 was a high priority for the *Monitoring Progress* committee.

Operational Definition

We define "content knowledge for teaching" as the combination of knowledge and skills required for the tasks performed in teaching students effectively; it goes beyond knowledge of subject matter per se to include an understanding of difficulties students often have in mastering the subject and techniques for overcoming those difficulties. While a direct assessment of science and mathematics teachers' content knowledge for teaching (CKT) would be ideal (e.g., Kersting, 2008; Hill, Rowan, & Ball, 2005), there are doubts about the feasibility of administering such assessments at scale. Interviews of experts focused on potential proximal indicators presumed to be related to teacher content knowledge for teaching. One approach that has been used in the past but that was deemed inadequate by a number of experts is whether teachers formally studied the subjects they teach in college and how prepared they feel to teach those subjects.

Interviews also surfaced some measurement concerns associated with these proxies. Teachers who are more prepared may be more likely to question their own preparedness and rate themselves lower. Also, some advisors and experts noted that degrees and credit hours in specific subjects often show little or no relationship to student outcomes. Others, however, indicated there may be value in collecting more nuanced data about teachers' degrees. For example, research by Monk and King-Rice suggests the degree teachers have matters more for some subjects (e.g., higher level science) than for others. Consensus existed among experts that teacher content knowledge data can be collected more accurately and cost-effectively from teachers rather than principals.

Currently Available Data

Data could be gathered from mathematical knowledge for teaching (MKT) studies, including those by Hill, and those used in the Measures of Effective Teaching project (MET Project, 2013). However, the data may be out of date, and a concern is that such measures are more appropriate for elementary rather than secondary grades.

Data related to teachers' perceptions of their preparedness to teach certain topics can be collected from past administrations of NAEP, TIMSS, and NSSME surveys. These data are limited, however, because none of the large-scale surveys contain direct assessments of teacher content knowledge (see table below). The Baccalaureate and Beyond Survey includes information about courses taken in college, although it is relevant only for recently hired teachers. The HSLS also has information about courses taken in college. The TALIS includes information on teacher training and education, as well as teachers' reported feelings of self-efficacy. As a whole, advisors believed these currently available data had face validity as measures for Indicator 6. In the next phase of the project, data will be collected and examined for possible inclusion in a companion report to the SEI to determine which ones offer the best information on this indicator.

Data source	Administered on a fixed recurring schedule	Produces grade-level estimates	Produces state-level estimates	Teacher perceptions of preparedness	Teacher college back-ground	Direct assessment of teacher content knowledge	Allows linking to student achievement data
2011 NAEP Science Teacher Background Questionnaire	Yes (every 4 years)	Yes (grades 4 and 8 only)	Yes, for most states	No	Yes	No	Yes
2011 TIMSS Study Grade 4 Teacher Questionnaire	Yes (every 4 years)	Yes (grades 4 and 8 only)	No	Yes	Yes	No	Yes
2012 NSSME Science Teacher Questionnaire	No	No	No	Yes	No	No	No
Baccalaureate and Beyond Survey	Yes (about every 3 years)	No (post-secondary students)	No	Unclear	Yes	No	No
HSLS09, HSLS12	Recurring; next administration unknown	Yes (just grade 9)	Partial (for 10 states)	No	Yes	No	Yes
2013 TALIS	Yes (every 5 years)	No	No	Yes	Yes	No	No

Near-Term Modifications and Activities

- SRI will investigate the extent to which teachers' perceptions of preparedness correlate with direct measures of content knowledge for teaching.
- The SRI team will assemble and compare data on teacher content knowledge for teaching and its proxies from the most recent administrations of the TIMSS, HSLS, and NSSME, as well as the MET and Hill's MKT surveys. SRI will review and synthesize what is known about the correlations between these measures and student achievement data.

Additional Research Needs

The relationship between teachers' science and math content knowledge and student achievement was recognized as an area where more research is needed as a long-term goal for measuring this indicator. Specifically, more research is needed addressing the following:

1. How well do teachers' self-reports of preparedness relate to their content knowledge?
2. How accurate are teachers' college study backgrounds in predicting student achievement?
3. What are some more cost-effective measures for direct assessments of teacher content knowledge that can be explored?

Additional research is also needed to develop teacher CKT instruments for science and high school math where teachers have content expertise, as well as nonsurvey measures to get at knowledge in use.

Long-Term Modifications and Activities

Through the DCL awards, NSF is investing in activities that address ways of measuring teacher content knowledge for teaching that may have implications for measuring this indicator in the long term:

- Goldhaber has received DCL funding to investigate the validity of the use of licensure tests as an indicator of teacher quality. This project may reveal another option for measuring this indicator.
- Through a joint award, Howell, Phelps, and Lai will test whether the assessment design theory for elementary MKT assessment items extends to secondary MKT item design

and develop recommendations for the research agenda needed to develop valid and reliable indicators of secondary teachers' mathematical knowledge for teaching.

- Mikeska and Phelps are examining how practice-based CKT items and subject matter knowledge items assess distinct aspects of elementary teachers' knowledge and how these measures relate to other indicators of teachers' professional preparation and experience.
- Kersting will develop an extension of the Classroom Video Analysis instrument that is aligned with the CCSS and then assess its feasibility as a scalable indicator of MKT.
- Mandinach and Orland will examine the feasibility of using the Statewide Longitudinal Data Systems (SLDS) infrastructure to collect some of the data elements related to Indicator 6.

References

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, *59*(5), 389–407.

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(1), pp. 14-17, 20-22, 43-46.

Goldhaber, D. (2014). *Assessing the use of licensure tests as an indicator of teachers' science and mathematics content knowledge for teaching*. National Science Foundation Award Abstract #1445618. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445618

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, *42*(2), 371–406.

Howell, H., & Phelps, G. (2014). *Building understanding and measurement of secondary mathematical knowledge for teaching (MKT)*. National Science Foundation Award Abstract #1445630. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445630&HistoricalAwards=false

Kersting, N. (2008). Using video clips as item prompts to measure teachers' knowledge of teaching mathematics. *Educational and Psychological Measurement, 68,* 845–861.

Kersting, N. (2014). *Adapting the Classroom Video Analysis approach as a feasible and scalable measure of Common-Core-aligned mathematics knowledge for teaching*. National Science Foundation Award Abstract #1445431. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445431

Mandinach, E., & Orland, M. (2014). *An exploration of the alignment of SLDS infrastructure and data highway to relevant success indicators in mathematics and science*. National Science Foundation Award Abstract #1445522. Retrieved from http://nsf.gov/awardsearch/showAward?AWD_ID=1445522

Mikeska, J., & Phelps, G. (2014). *Moving beyond subject matter knowledge: Assessing elementary teachers' content knowledge for teaching science*. National Science Foundation Award Abstract #1445641. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445641

Monk, D., & King-Rice, J. (1994). Multi-level teacher resource effects on pupil performance in secondary mathematics and science: The role of teacher subject matter preparation. In R. Ehrenberg (Ed.), *Choices and consequences: Contemporary policy issues in education* (pp. 29-58). Ithaca, NY: ILR Press.

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, *44(*4), 921–958.

Penuel, W., Fishman, B. J., Gallagher, L. P., Korbak, C., & Lopez- Prado, B. (2009). Is alignment enough? Investigating the effects of state policies and professional development on science curriculum implementation. *Science Education*, *93*(4), 656–677.

Roth, K. J., Garnier, H. E., Chen, C., Lemmens, M., Schwille, K., & Wickler, N. I. Z. (2011). Video-based lesson analysis: Effective science PD for teacher and student learning. *Journal of Research in Science Teaching*, *48*, 117–148. doi:10.1002/tea.20408

Lai, Y.-J. (2014). *Building understanding and measurement of secondary mathematical knowledge for teaching (MKT)*. National Science Foundation Award Abstract # 1445551. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445551

Indicator 7: Teachers’ Participation in STEM-Specific Professional Development

Indicator 7 concerns teachers’ participation in professional development activities targeted for one of the STEM fields or STEM in general.

Priority: Medium

The *Monitoring Progress* committee did not identify Indicator 7 as a high priority. In addition, a number of experts SRI interviewed regarded it as a relatively low-priority indicator because participation in professional development (PD) is generally not a significant predictor of student achievement except when the PD is closely tied to a content-specific instructional intervention. The advisory group, however, disagreed, believing there was enough information about the aspects of PD that impact instructional practices to heighten the priority of this indicator.

Operational Definition

We define Indicator 7 as in-service teachers’ participation in formal or informal activities designed to enhance their ability to deliver effective instruction in one or more of the STEM subject areas. Note that informal activities such as mentoring programs or communities of learners around STEM instruction would be included in this definition; general training or mentoring on becoming a better teacher in general would not. Indicator 7 deals with quantitative rather than qualitative aspects of professional development. Clearly, the quality of professional development is important and research clarifying the quality dimensions of STEM PD that relate to better student outcomes may provide a basis for future extension of Indicator 7 to incorporate those dimensions.

The research literature suggests several attributes that appear to be associated with more effective professional development (e.g., Darling-Hammond et al., 2009; Desimone, 2009; Garret et al., 2001; Penuel et al., 2007): PD that is ongoing, includes more than one teacher per school, and is tightly aligned with curriculum. In the short term, we propose an operational definition of whether a teacher participated in STEM (inclusive of just science, just math, etc.) professional development in the past year and if so, for how much time. In the long term, we propose an operational definition that also reflects whether that professional development included any of the known best practice characteristics as determined by Garret et al. (see Near-Term Modifications for language suggested by the advisory group). Prior survey items on STEM PD were not framed with the new standards and their emphasis on practices and crosscutting concepts in mind. Some advisors suggested that instructional approaches and therefore PD approaches might vary across grade levels.

Currently Available Data

The Education Commission for the States' 50-State Analysis has descriptive information from 2008 on high school-level STEM initiatives related to whether states offered or required targeted PD for STEM teachers. However, these data may be out of date, and they do not provide information on teachers' participation in the PD activities that were offered.

Data from previous administrations of the NAEP, TIMSS, SASS, and NSSME could be used for this indicator regarding the amount of PD received in science and mathematics. Past administrations of the SASS and NAEP have included questions about common topics for professional development. The 2012 NSSME included many questions about teacher PD that were specific to math and science, including some about coaching and professional learning communities. The ECLS-K asks about teachers' participation in math or the teaching of math and science or the teaching of science. The OECD TALIS survey data recently became available, and the survey included PD questions for principals and teachers.

These data are limited, however, in that most cannot provide aspects of PD, such as application in the classroom, method of delivery, and perceived usefulness (see table below). Nonetheless, advisors felt that the currently available data offer powerful descriptive information on whether there is investment in PD at the state level. In the next phase of the project, we will collect and examine data for possible inclusion in a companion report to the SEI to determine which sources offer the best information on this indicator. Advisors also indicated that an SEI companion report could be an opportunity to discuss the challenges in measuring the quality of PD.

Data Source	Administered on a fixed recurring schedule	Produces grade-level estimates	Produces state-level estimates	Math & Science PD items	Includes PD best practices	Allows linking to student achievement data
2011 NAEP Grade 4 Science Teacher Background Questionnaire	Yes (every 4 years)	Yes (grade 4 only)	Yes, for most states	Yes	No	Yes
2011 TIMSS Grade 4 Teacher Questionnaire	Yes (every 4 years)	Yes (grades 4 and 8)	No	Yes	No	Yes
2011-12 SY SASS Teacher Questionnaire*	Yes (every 4 years)	No	Yes	Yes	No	No
2011-12 SY SASS Principal Questionnaire*	Yes (every 4 years)	No	Yes	Yes	No	No
2012 NSSME Science Teacher Questionnaire	No	No	No	Yes	Yes	No
2013 OECD TALIS	TBD	No	No	Yes	Yes	No

**The redesigned SASS (now the National Teacher and Principal Survey, NTPS) will be administered every 2 years and produce national estimates only.*

Near-Term Modifications and Activities

- The 2017–18 NTPS will include a teacher professional development module to look at aspects of professional development such as application in the classroom, method of delivery, and perceived usefulness. The SRI team will review the existing PD items on the NAEP, TIMSS, SASS, and NSSME and develop a set of proposed items for consideration by the NCES team developing items for the NTPS. Potential NTPS item topics include:
  - Did you receive PD in this past year?
    - Did it focus on content, instructional strategies, or a combination?
    - Was it a one-time PD event or was it ongoing?
    - How long did the PD last?
    - Did the PD include a chance to practice the new approach?
    - To what extent was the PD you received connected to evaluation of you as a STEM teacher?
    - Was the PD targeted for STEM teachers, or was it for all teachers?
- Data on the quantity of professional development may also be available from districts. Mandinach has received funding as a part of the DCL to determine whether existing

state data collections could address Indicator 7 and what revisions to the state systems would be necessary to do so. This project may yield additional options for near- and long-term modifications and activities.

- In 2015, data from the 2014 NAEP TEL will be available related to common topics for professional development. The PD items used on the NAEP TEL could be modified/added for the 2017 NAEP Math.
- Several experts we spoke with suggested additional options for defining this indicator—such as whether teachers have professional development opportunities to learn about the content their students are learning—that could be explored as measures of or proxies for this indicator.
- SRI will examine the results of recent and upcoming meta-analyses related to research on PD, including
    - Blank's meta-analysis of teacher PD and its relationship to student achievement. The study examined some of the more detailed aspects of PD and primarily focused on science and math.
    - Taylor's NSF grant for a meta-analysis examining the treatment effects of various teacher professional development programs.
    - Desimone's review of randomized controlled trials (RCTs) on PD, which indicates that short, targeted PD can work if it is focused on high-leverage practices. PD also must be linked to what is happening in the classroom in terms of instructional strategies or the curriculum. Additionally, changes in practice are more likely if teachers are accountable for the instructional strategies they were taught in the PD. In the near term, the group suggested developing survey questions to determine whether the PD teachers receive includes these components.

Additional Research Needs

Research efforts are needed to determine the relationship between teacher participation in PD and student achievement in math and science. Specifically,

1. What are the characteristics of PD activities that impact classroom practice? To refine the set of characteristics that relate to classroom impact, are there features other than those identified by Garret et al. (2001) that relate to student outcomes and/or changes in teacher practice? What about those that affect student outcomes in math and science?
2. What impact do features such as application in the classroom, method of delivery, and perceived usefulness of PD have on student outcomes?
3. How should the quality of STEM PD be measured?
4. How could formal and/or informal mentoring experiences of teachers be measured?

Advisors also indicated that the field lacks information on what kind of approach to instruction is effective for students at different grade levels in different subjects. More information on this would help support the development of effective STEM PD.

Long-Term Modifications and Activities

- Consider modifications based on recommendations, once available, from Mandinach's DCL award.
- One aspect of Indicator 7 that may be important to explore in the future as more PD includes an online component is whether it is possible to leverage back-end learning system data about the quantity and quality of STEM PD as an alternative to relying on surveys.

References

Blank, R. *What research tells us – Common characteristics of professional development that leads to student achievement*. (2013, February). Learning Forward. Retrieved from http://learningforward.org/docs/default-source/jsd-february-2013/blank341.pdf?sfvrsn=2

Desimone, L. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38,* 181–199.

Garret, M., Porter, A., Desimone, L., Birman, B., & Kwang, S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38,* 915–945.

Mandinach, E., & Orland, M. (2014). *An exploration of the alignment of SLDS infrastructure and data highway to relevant success indicators in mathematics and science*. National Science Foundation Award Abstract #1445522. Retrieved from http://nsf.gov/awardsearch/showAward?AWD_ID=1445522

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, *44*, 921–958.

Wei, R. C., Darling-Hammond, L., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad.* Dallas, TX: National Staff Development Council.

Indicator 8: Instructional leaders' Participation in PD to Support STEM Learning

Indicator 8 concerns instructional leaders' participation in professional development that helps principals create school conditions supportive of STEM learning. Some of these conditions may be specific to STEM, whereas many are aspects of a positive school climate that supports all kinds of academic learning.

Priority: Low

The Successful K-12 STEM Education committee considered this a key indicator to monitor, but the Monitoring Progress committee did not. Considerable research supports the idea that principals are very important in shaping the climate and allocation of resources within a school that enable good STEM education. However, there is very little research on what kinds of PD enable principals to do this better and even less that suggests specific ways in which principals could enhance STEM teaching and learning. Project advisors saw this as a low- priority indicator as it is currently framed because very little evidence exists suggesting that PD changes principals' effectiveness or student outcomes in STEM.

Operational Definition

We define Indicator 8 as the participation of in-service principals and other school leaders (such as department chairs) in formal or informal activities designed to enhance their ability to serve as instructional leaders in one or more of the STEM subject areas. Our project advisors pointed to emerging research on the importance of teacher evaluation and feedback (Steinberg & Sartain, forthcoming; Taylor & Tyler, 2012). Asking principals whether they have received training on evaluating and providing feedback specific to math and science instruction would be one aspect of measuring Indicator 8.

The *Monitoring Progress* report states that for Indicator 8, professional development for leaders should be defined broadly and could include such activities as coaching and time to discuss work with peers. The NRC committee was initially interested in examining PD on creating a school climate for learning in general as well as PD on becoming an instructional leader for STEM. We have proposed an operational definition focused on just the latter purpose because our project advisors believed this to be the most pressing need with respect to enhancing STEM instruction.

Currently Available Data

No data sources were identified for this indicator. The SASS asks about participation in general PD for school principals but does not ask about its characteristics or quality or about PD relating to instructional leadership in specific subjects.

Near-Term Modifications and Activities

- Synthesizing theories about principal leadership for instructional improvement would be helpful before measuring Indicator 8 more thoroughly. Goldring and Porter's VAL-ED work on principal leadership may contribute insights into the quality and nature of principal PD that is related to student achievement. While not a tool for assessing PD directly, this leadership assessment could be used as an outcome measure to examine the efficacy of principal PD. Spillane's work on principal policy and practice research may also provide insights.
- Data on developing a positive school climate could be collected on the future NCES School Climate Survey (currently under development) that will be administered to schools, districts, and states and will include questions for students, teachers, staff/principals, and possibly parents. In addition, NCES will conduct a one-time national school climate benchmark, possibly as early as 2014–15 or 2015–16, that may contain helpful data.
- Review 2013 TALIS questionnaires to determine whether any items provide relevant data for this indicator. The TALIS includes teachers' reported job satisfaction and the climate in the schools and classrooms they work in. Relating these reports of climate to principal behaviors may be possible.
- The 2017–18 NTPS principal survey will include a principal professional development module. NCES has indicated that it would welcome input on the content of this module. This is most likely the best fit to provide some type of data on the topic.

- Mandinach has received funding through the DCL to determine whether existing state data collections could address Indicator 8 and what revisions to the state systems would be necessary to do so. This project may yield additional options for this indicator.

Additional Research Needs

Experts noted concerns regarding the viability of collecting accurate and interpretable data for this indicator. Additional research is needed to gain insights into the qualities of principal PD that impact school climate or teacher instructional practices to effectively support STEM teaching and learning. Specifically,

1. What are the qualities of principal PD that make a difference for STEM teaching and learning practices?
2. What types of principal PD are related to student achievement?

Long-Term Modifications and Activities

In moving toward capturing the aspects of principal PD that affect school climate and STEM teaching and learning practices, there may be opportunities for collecting data on this indicator on future NAEP Math and Science, NTPS, and/or NSSME surveys.

References

Bryk, A. S., Sebring, P. B., Allensworth, E., Easton, J. Q., & Luppescu, S. (2010). *Organizing schools for improvement: Lessons from Chicago*. IL: University of Chicago Press.
Last chapter includes additional information on what sort of school leadership qualities affect teaching and learning.

Elmore, R. F., & Burney, D. (1997). Investing in teacher learning: Staff development and instructional improvement in Community School District #2, New York City. New York, NY: National Commission on Teaching & America's Future, Teachers College.

Grissom, J. A., Loeb, S., & Master, B. (2013). Effective instructional time use for school leaders: Longitudinal evidence from observations of principals. *Educational Researcher*, 42(8), 433–444.

Mandinach, E., & Orland, M. (2014). *An exploration of the alignment of SLDS infrastructure and data highway to relevant success indicators in mathematics and science*. National Science

Foundation Award Abstract #1445522. Retrieved from http://nsf.gov/awardsearch/showAward?AWD_ID=1445522

Steinberg, M., & Sartain, L. Does teacher evaluation improve school performance? Experimental evidence from Chicago's Excellence in Teaching Project. Forthcoming in *Education Finance and Policy*.

Taylor, E. S., & Tyler, J. H. (2012). The effect of evaluation on teacher performance. *American Economic Review*, *102*(7), 3628–51.

Indicator 9: Inclusion of Science in Federal and State Accountability Systems

Indicator 9 addresses the inclusion of science in federal and state accountability systems.

Priority: High

Indicator 9 is identified as a high priority in the *Monitoring Progress* report.

Operational Definition

Indicator 9 is defined as a policy of holding districts, schools, and/or teachers responsible for students' academic progress in science by linking such progress to funding, autonomy, or other incentives. To count as being part of an accountability system, science must not only be assessed, but student performance on the assessment must also have stakes attached. The NRC Successful K-12 STEM Education working group was concerned that science instruction would always take a backseat to instruction on English language arts (ELA) and mathematics as long as accountability systems stressed those two subjects and not science.

Currently Available Data

- Data are available from the Council of Chief State School Officers (CCSSO) from the 2009–10 school year on science assessments by state. However, these data do not indicate whether science is included in states' accountability systems. These data still appear to be the best source of information on state subject area assessments, although they are also most likely out of date.
- Blank (2013) summarizes current policies, indicating that science must be tested and reported publicly at three grades, at least one each at the elementary, middle, and high school levels. Additionally, though student science achievement scores are not required to be a part of school annual progress determinations, 12 states have elected to include science scores as an additional accountability indicator.

- Judson (2010, 2012) examined effects on student performance of states having a policy to include science in the accountability system.

Near-Term Modifications and Activities

- EDFacts is collecting survey data on states' science assessment participation and science assessment results. Although it is not currently collecting information on the use of assessment data in state accountability systems, its data collection system has the ability to do so. EDFacts' Office of Management and Budget (OMB) data collection budget may cover the costs of expanding the survey to address accountability.
- SRI plans to determine what information can be gleaned for this indicator from the Consolidated State Application Accountability Workbooks, which include addenda that describe any Elementary and Secondary Education Act (ESEA) flexibility that has been approved. These could be analyzed to understand
    - Past and current state policies for types and uses of science assessments (i.e., formative for improvement, summative for student outcomes) and for the frequency of assessment (e.g., once a year, midyear, multiple times a year)
    - Policies 2 years before analyzing assessments (e.g., review policies that are in place in 2016 and then review assessments in 2018).
- Blank's EAGER DCL study will also examine state tests and accountability in science; these data may be available in time for a companion to the SEI 2016.

Additional Research Needs

Identifying relevant and appropriate data sources for this indicator is an important first step. More research is needed related to the following:

1. What kinds of information can be compiled from states' submitted NCLB reports (including waiver applications) and from existing policy websites? What kind of coding system would be appropriate to ascertain whether science is included as part of states' accountability system?
2. What features of the accountability system are important to measure for states that include science in their accountability system (e.g., nature of stakes attached to science achievement)?

Long-Term Modifications and Activities

- Blank's EAGER DCL study examines what core information on state student assessments *should be* collected and reported across states and tracked over time. Designing and implementing a database with information about accountability structures and legislative mandates across states will facilitate reporting on science accountability systems on a national scale.
- Making a policy recommendation about the frequency of information gathering for this indicator will also be important, as will be identifying a suitable mechanism for collecting the data.

References

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,* 830-847.

Blank, R. (2014). *Developing a system for tracking state assessment policies in science and mathematics education*. National Science Foundation Award Abstract #1445610. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445610

Judson, E. (2012). When science counts as much as reading and mathematics: An examination of differing state accountability policies*. Education Policy Analysis Archives, 20*. Retrieved from http://epaa.asu.edu/ojs/article/view/987

Judson, E. (2010). Science education as a contributor to adequate yearly progress and accountability programs. *Science Education, 94*(5), 888–902.

U.S. Department of Education. (2012). *Education data express, State performance and accountability*. Retrieved from http://www.eddataexpress.ed.gov/

Indicator 10: Inclusion of Science in Major Federal K-12 Education Initiatives

Indicator 10 examines the proportion of major federal K-12 education initiatives that include science.

Priority: Low

Indicator 10 was not identified as a priority in the *Monitoring Progress* report and was also viewed as a low priority by the advisory group. This indicator may be more relevant to address once the Elementary and Secondary Education Act (ESEA) is reauthorized. However, there is a desire to be more specific about the minimization of science in advance of an ESEA reauthorization. During the 2014 meeting, the advisory group supported the selection of this topic for a concept paper to spur discussion leading up to the 2016 election.

Operational Definition

This indicator concerns the inclusion of science in federal education initiatives and the level of funding relative to ELA and math. We propose to define the following as major federal K-12 education initiatives:

1. ESEA reauthorization
2. Investing in Innovation and Improvement
3. America COMPETES
4. NAEP Assessments.

The authors of the *Monitoring Progress* report also included Race to the Top in the definition for this indicator, but we recommend excluding it now that the initiative has concluded. We could also look at NCLB waivers on the grounds that they represent the de facto federal policy, but given the labor intensity of this task, we recommend waiting until there is a reauthorized ESEA to collect a baseline measure for this indicator. Reporting meaningfully on this indicator also requires establishing clear definitions for how program dollars are allocated to subject areas, which would be complicated by the fact that funding is often targeted to the student population served rather than the content addressed.

Currently Available Data

Several information sources for Indicator 10 exist, but most have not been coded for this indicator nor has the information been compiled. Two more readily available data sources were identified:

1. Information about the frequency of NAEP science assessments relative to math and reading is available from NCES.
2. The 2010 Federal STEM Education Inventory Data Set is available from OSTP.

SRI will need to reanalyze the OSTP data to restrict the analysis to be specific to K-12 and examine funds allocated by STEM discipline (science, technology, engineering, or math).

Near-Term Modifications and Activities

The priority near-term activity is to develop a report for a general audience on the emphasis of science relative to ELA and math to generate interest in the topic. Longer term activities should wait until after the new ESEA reauthorization, since the new administration may change policy and reallocate funds for science education initiatives (e.g., Department of Energy, NASA).

Effort is needed to analyze publicly available budget data, federal solicitations for the largest grant programs and other competitive funding, and the language in the reauthorized ESEA to review funds allocated to science compared with reading and math. ESEA reauthorization has yet to occur, however, and state education policies are in flux with respect to adoption and support of new standards. We therefore recommend waiting to conduct this task until there ESEA is reauthorized.

Additional Research Needs

Additional research could examine the role of STEM in federal initiatives over the years.

Long-Term Modifications and Activities

It will be important to determine an efficient mechanism for collecting and analyzing these data on a recurring basis. Policymakers can provide direction on whether this is an appropriate activity for the Congressional Budget Office or the Congressional Research Service.

References

Federal Coordination in STEM Education Task Force Committee on STEM Education National Science and Technology Council. (2012). *Coordinating federal science, technology, engineering, and mathematics (STEM) education investments: Progress report*. Retrieved from http://www.whitehouse.gov/sites/default/files/microsites/ostp/nstc_federal_stem_education_coordination_report.pdf

Indicator 11: State and District Staff Dedicated to Supporting Science Instruction

Indicator 11 is intended to measure the human resources available at the state and district levels for science as compared with math and reading for all grades.

Priority: Low

The advisory group considered Indicator 11 to be a lower priority indicator. In addition, it was not a priority indicator for the *Monitoring Progress* committee.

Operational Definition

This indicator is defined as the number of full-time employees in state and district education agencies who are devoted solely to science instruction, curriculum, professional development, or assessment. Discussions with project advisors and other experts suggested that it is premature to provide an operational definition for this indicator. One limitation of this definition of "state and district staff devoted to science instruction" is that some states have math and science centers between the district and state level. By design, these intermediate units supplement professional development and resources provided by local districts and state education agencies. There is also ambiguity regarding issues such as whether staff developing science assessments or science teachers who mentor colleagues should be included in the definition. Background work on this indicator revealed that one aspect of Indicator 11 that could be measured in the relatively near term is the prevalence of science specialists teaching science at an elementary level.

Currently Available Data

Data relevant to Indicator 11 could be collected from the National Council for Teachers of Mathematics (NCTM) *Elementary Mathematics Specialist's Handbook* regarding mathematics specialists in Virginia. Pat Campbell at the University of Maryland has a similar study (MACMTL

Quantitative Study of Teacher Knowledge and Student Achievement), but the sample is limited to 350 teachers. In the 2014 meeting, the advisory group believed that it would not be difficult to measure whether there are science and math specialists, but it may be difficult to count coaches or other staff with more ambiguous roles. However, the original NRC group wanted to focus first on counting staff at state education agencies (SEAs) who are devoted to science because this may be a clearer measure.

Additional available sources include data from the SASS and ECLS-K that ask teachers to report their main teaching assignment field and whether they are in contained classrooms or are elementary subject specialists. The ECLS-K also has an item on the School Administrator survey capturing the full-time equivalent for math and science specialists. Additionally, the 2012 NSSME included questions about coaching on the program questionnaire. These data are limited in that they cannot provide comprehensive estimates of state and district staff dedicated to supporting science instruction because that position varies across districts (see table below). In the next phase of the project, we will examine the data sources for possible inclusion in a companion report to the SEI to determine which ones offer the best information on this indicator.

Data source	Administered on a fixed recurring schedule	Produces grade-level estimates	Produces state-level estimates	Provides info from teachers	Provides info from principals/admins	Allows linking to student achievement data
NCTM's The Elementary Mathematics Specialist's Handbook	No	TBD	No	Yes	TBD	TBD
MACMTL Quantitative Study of Teacher Knowledge and Student Achievement	No	Ranges	No	Yes	No	Yes
2011-12 SY SASS Teacher Questionnaire	Yes (every 4 years)	No	Yes	Yes	No	No
2012 NSSME Science Teacher Questionnaire	No	No	No	Yes	No	No
ECLS-K11 School Admin Survey	No	Yes (K)	No	No	Yes	Yes
ECLS-K11 Teacher Survey	No	Yes (K)	No	Yes	No	Yes

Near-Term Modifications and Activities

- SRI will contact CCSSO, National Science Teachers Association (NSTA), Association of State Supervisors of Mathematics (ASSM), American Association for the Advancement of Science (AAAS), the National Science Supervisors Association (NSSA), the Board on Science Education (BOSE), and/or SEAs/LEAs (e.g., Kansas State Department of Education, Palo Alto Unified School District) to see whether they have information on organizations at the intermediate level that have staff supporting science instruction or are aware of ongoing research in this area.
- SRI will make a recommendation for which existing survey, such as the CCD, would be the best for counting staff at SEAs who are devoted to science. The draft question is: How many full-time employees do you have in the SEA who are devoted solely to science instruction, curriculum, professional development, or assessment?

Additional Research Needs

The relationship between state and district staff dedicated to supporting science instruction and student achievement was recognized as an area for additional research. Specifically, more research is needed addressing the following:

1. Is there any relationship between allocation of resources for science support staff and student achievement outcomes?
2. Has the number of staff supporting science increased or decreased in the past 5 years? Is there a relationship to test scores in corresponding areas?

Long-Term Modifications and Activities

If the initial research shows relationships between this indicator and student achievement, beyond the NTPS, there may be additional opportunities for collecting data on this indicator with the NSSME or by proposing NSSME-like items regarding coaching for inclusion on future NAEP Math and Science surveys.

References

Campbell, P., Ellington, A., Haver, W., & Inge, V. (2013). *The elementary mathematics specialist's handbook*. National Council for Teachers of Mathematics. Retrieved from http://www.nctm.org/catalog/product.aspx?id=14302

Levy, A. J., Pasquale, M., & Marco, L. (2008). *Models of providing science instruction in the elementary grades: A research agenda to inform decision makers.* National Science Teachers Association. Retrieved from http://www.nsta.org/elementaryschool/connections/201007SummerResearchLevy.pdf

Indicator 12: States' Use of Standards-Aligned Assessments

Indicator 12 looks at states' use of assessments that measure the core concepts and practices of science and mathematics disciplines.

Priority: Medium

The decision to place Indicator 12 at medium priority was based on the recommendation of the advisors and other experts we consulted that it be measured in tandem with standards-aligned instructional materials (Indicator 4), which is identified as a high priority in the *Monitoring Progress* report.

Operational Definition

We define a "standards-aligned assessment" as a set of tasks or items administered to students that (1) call on the same concepts and practices as the CCSS-M and NGSS science standards and (2) provide scores relative to the level of performance expectations in those standards. As was the case for instructional materials, to be fully standards-aligned, an assessment would need to cover the full content in the standards for the relevant grade level, be consistent with the performance-level expectations of those standards, and reflect the same conception of learning and instruction (e.g., the intertwining of disciplinary core ideas, science and engineering practices, and crosscutting concepts in the Science Framework).

The Smarter Balanced and PARCC Mathematics assessments were designed to be consistent with the CCSS-M standards, but an independent evaluation of the degree of alignment of those tests with the standards is still needed. For states that do not adopt CCSS-M assessments or the NGSS, data for this indicator will be more difficult to collect. For these states, analysts still could look at the extent to which their assessments do align with the content, practices, and performance expectations of the national standards. Applying this process to many different state assessments would be labor intensive, but it could provide information that is useful in comparing student assessment data across states as well as capturing Indicator 12 at a national level. Experts agreed that more work is needed before a process for measuring Indicator 12 consistent with this operational definition can be fully established.

Currently Available Data

As for Indicator 4 (standards-aligned instructional materials), data are available on the most commonly used assessments for K-12, but no data exist on the alignment of the assessments with standards.

- The CCSSO 50 State report detailed states' assessments by grade, by subject, and how they were conducted (e.g., open response items and/or multiple choice), as well as when the testing occurred. This work was produced annually from 2000 to 2009.
- Some descriptive information can also be obtained from the *State of State Science Standards 2012*, produced by the Thomas E. Fordham Institute.

Near-Term Modifications and Activities

A number of efforts are currently under way to produce data on the alignment of assessments with the new standards:

- The National Center for Research on Evaluation, Standards, and Student Testing (CRESST) is engaged in research on the development of K-12 assessments that measure the CCSS. CRESST researchers are also conducting a formative evaluation of embedded assessment modules developed by WestEd and providing guidance on reliability studies of simulation-based assessments and assisting with validity studies.
- RAND is conducting research to understand the relationship between new CCSS-aligned assessments that are designed to measure higher order skills and how assessments influence classroom practice.
- Blank's EAGER DCL study examines how assessments in math and science education are aligned with state-adopted content standards.

SRI will review the RAND report (Faxon-Mills, Hamilton, Rudnick, & Stecher, 2013) and consider interviewing the CRESST researchers to ascertain when their data will be available and to obtain details of their methodology.

Additional Research Needs

Information will also be needed on assessments that states adopt in the coming years that have not been reviewed by RAND or CRESST to determine the extent to which they align with the CCSS-M and the NGSS or determine the extent to which the assessment systems otherwise emphasize math and science practices in addition to concept mastery. Specifically,

1. What procedures or mechanisms exist for assessing the alignment of assessments with state standards? How sensitive are conclusions about alignment to the measure used?
2. How can we leverage the methodologies used by RAND and CRESST to gather and analyze non-CCSS state math assessments and states' science assessments?
3. How can RAND’s and CRESST’s methodology be adapted to assess science practices in the NGSS, where students’ competencies are developed over time?
4. To what extent are states also supporting the implementation of aligned interim, benchmark, or formative assessments?

Long-Term Modifications and Activities

Rolf Blank received a DCL award for studying the alignment of state assessments with the core concepts in the current college and career readiness standards adopted by states. We expect this project will provide additional insights and relevant data.

References

Blank, R. (2014). *Developing a system for tracking state assessment policies in science and mathematics education*. National Science Foundation Award Abstract #1445610. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445610

Faxon-Mills, S., Hamilton, L. S., Rudnick, M., & Stecher, B. M. (2013). *New assessments, better instruction? Designing assessment systems to promote instructional improvement*. Los Angeles, CA: RAND Corporation. Retrieved from http://www.rand.org/pubs/research_reports/RR354.html

Lerner, L. S., Goodenough, U., Lynch, J., Schwartz, M., Schwartz, R., & Gross, P. R. (2012). *The state of state science standards 2012.* Washington, DC: Thomas B. Fordham Institute.

Indicator 13: Expenditures for Improving the K-12 STEM Teaching Workforce

Indicator 13 concerns state and federal expenditures dedicated to improving the K-12 STEM teaching workforce.

Priority: Medium

Neither the Monitoring Progress committee nor our advisory group considered this indicator a priority. Experts consulted indicated that obtaining data would require considerable time and resources. However, the recent *5-Year Federal Science, Technology, Engineering, and Mathematics (STEM) Education Strategic Plan* produced by the Obama Administration states that one of the Administration's top goals is to "provide additional resources to meet specific national goals, such as preparing and recruiting 100,000 high-quality K-12 STEM teachers" (National Science and Technology Council, 2013). At this time, there are three types of activities that SRI can engage in, and the development of items that are easy to measure for future collections is the top priority of those activities.

Operational Definition

Indicator 13 is defined as state and federal expenditures targeted for one of the STEM fields or STEM in general that are designed to increase the supply of K-12 STEM teachers, reduce attrition from that workforce, or improve the quality of K-12 STEM teaching. On advice from experts, we recommend concentrating near-term efforts to monitor this indicator on professional development activities supported by federal and state funding. Any information collected should distinguish preservice preparation from in-service professional development. More work is needed to define what other components of workforce development (such as investments in recruitment costs, differential pay, alternative credentials) should be measured for Indicator 13 longer term.

Currently Available Data

Although they are not comprehensive, several sources of data on components of this indicator were identified:

- The Survey of the Use of Funds Under Title II provides annual data from a nationally representative sample of districts about how funds were used, including the proportion spent on science PD.
- OSTP has 2010 data on federal STEM expenditures; these data are limited in that they are available for only a single point in time. SRI will also need to reanalyze the data to be specific to K-12 and examine funds allocated by STEM discipline (science, technology, engineering, or math).
- The Education Commission for the States' 50 State Analysis provides information on high school-level STEM initiatives, including STEM-targeted professional development.

Data source	Administered on a fixed recurring schedule	Produces grade-level estimates	Produces state-level estimates	Includes spending amounts	Allows linking to student achievement data
Survey of the Use of Funds Under Title II	Yes (annual)	No	No	Yes	No
OSTP data on federal STEM expenditures	No	No (grade ranges)	No	Yes	No

Near-Term Modifications and Activities

SRI can engage in three types of activities and modifications: seek more sources of existing relevant data, develop items that are easy to measure for future collections, and research existing sources of qualitative and quantitative information.

Seek More Sources of Relevant Data

- Efforts are needed to consult with OMB to determine whether it has data relevant to this indicator. NSF is willing to help coordinate this.
- Data on expenditures may also be available from districts. Because some states have more established systems for reporting on PD, states could be surveyed regarding their definitions and level of detail for their reporting. We also propose exploring whether Mandinach's DCL award determining whether existing state data collections could

address the indicators and what revisions would be necessary to do so be expanded in include Indicator 13.

- The National Academy of Science had a meeting on designing data systems for teacher preparation data (one form of expenditures), organized by Judith Koenig. A small group has been trying to develop indicators in response to this, and it could be contacted to see whether this effort could provide any insights into measuring this indicator.
- One type of data that would be relatively easy to obtain is any differences in salary levels for STEM teachers (versus non-STEM teachers) across states.

Develop Items That Are Easy to Measure

- To add to the available data on this indicator, develop two items for proposed inclusion in the NTPS:
  - An item asking teachers whether they are compensated for their STEM-focused service.
    - Do you receive differential pay for teaching in the STEM field?
  - An item asking principals whether they have unfilled STEM teaching jobs. Determine whether the SASS had a similar question on past administrations.
- Because collecting data for this indicator will be challenging given differences in state structures and definitions for reporting expenditures, we recommend working with the NCES Data Forum and/or Common Education Data Standards (CEDS) Stakeholder Group to develop common expenditure definitions and reporting structures.

Researching Existing Sources of Qualitative and Quantitative Information

- Determine whether the OSTP standing committee on STEM has additional information and/or could play a role in coordination.
- Compile information on major specific federal and state programs aimed at improving the STEM teaching workforce. Additional information may be available from the OSTP Committee on STEM and from OMB and the Department of Defense. Indirect funding information could be gained by determining the number of STEM credentials given to education school graduates versus people in other fields and trying to calculate the state and federal subsidies that are going into the production of those alternatively certified teachers.

- Consider reviewing data on differential pay by state and district from the Bureau of Labor Statistics and compare elementary and secondary teachers with other professions with the same degree. Review the recent National Bureau of Economic Research (NBER) study that found a relationship between level of per-pupil spending and achievement gain and that also includes information on the level of spending on teachers. The NBER data may specify the teachers' subject area, allowing comparisons of science and math teachers with those of other subjects.
- Efforts are needed to identify and total the grants or grant programs that support K-12 STEM teacher training and professional development at NSF, the Department of Education (Title II), and other agencies such as NASA and the Department of Energy.

Additional Research Needs

1. How should STEM expenditures be classified, particularly at the elementary level where this can be more challenging because teachers typically teach multiple subjects?
2. How can we convince developers of state systems that the benefits of a comparable system (e.g., comparable PD cost reporting) would be worth the time, cost, and effort it would take for the development work?
3. Are the professional development activities supported by federal and state funding consistent with best practices?

Long-Term Modifications and Activities

It would be beneficial to discuss with policymakers what data they would like to have on this topic, as they may be especially interested in international comparisons. This indicator may be especially useful in the policy advocacy context. OSTP may make long-term recommendations regarding STEM spending in future reports.

References

National Science and Technology Council. (2013). *5-Year federal science, technology, engineering, and mathematics (STEM) education strategic plan.* Retrieved from http://www.whitehouse.gov/sites/default/files/microsites/ostp/stem_stratplan_2013.pdf

OSTP. (2010). Federal STEM Inventory data set. Retrieved from http://www.whitehouse.gov/sites/default/files/microsites/ostp/Guide%20to%20STEM%20Ed%20Data%20Set.pdf

OSTP. (2011). *The federal science, technology, engineering, and mathematics (STEM) education portfolio*. Retrieved from http://www.whitehouse.gov/sites/default/files/microsites/ostp/costem__federal_stem_education_portfolio_report_1.pdf

Picus, L., Monk, D., & Knight, D. (2012). Measuring the cost-effectiveness of rich clinical practice in teacher preparation: Part one, understanding the problem. Retrieved from https://caepnet.files.wordpress.com/2012/12/picusmonk.pdf

Indicator 14: Federal Funding for STEM Education Research

Indicator 14 calls for new research to fill critical gaps in knowledge about programs and practices that contribute to student learning and to the other goals of STEM education. This indicator is based on the recommendation for a robust and strategic research agenda for K-12 STEM education that would (1) disentangle the effects of school practice from student selection, (2) recognize the importance of contextual variables, and (3) allow for longitudinal assessments of student outcomes.

Priority: High

Indicator 14 is identified as a high priority in the *Monitoring Progress* report. In addition, the recent *5-Year Federal Science, Technology, Engineering, and Mathematics (STEM) Education Strategic Plan* produced by the Administration (National Science and Technology Council, 2013) recommends investing in breakthrough research on STEM teaching and learning.

Operational Definition

Additional discussion is needed before an operational definition of Indicator 14 can be solidified. Consensus is needed on the boundaries of what constitutes "research" as opposed to services, demonstration, or advocacy and whether research activities in all federal agencies or only those of the NSF, Department of Education, and Smithsonian should be examined. Finally, agreement is needed on whether to try to estimate the total investment in K-12 STEM education research or to focus more narrowly on research exhibiting the three characteristics advocated by the Successful K-12 STEM Education working group. Attempts to measure the latter would require a labor-intensive examination of project abstracts and reports for funded projects. Experts we interviewed proposed focusing on generalizability, attention to context, and longitudinal assessment as criteria for defining a robust and strategic research agenda in K-12 STEM education. Such a definition should not be construed as not valuing exploratory work.

As a part of the DCL, Kowalski and Taylor are producing baseline data on the extent to which federal agencies have funded the kinds of research called for in the NRC report. They will review the current evidence base on the indicator and its measurement and may produce recommendations for the creation of measures for the indicator.

Currently Available Data

NSF has funded evaluations of its DRK-12, ITEST, and ISE portfolios. However, these evaluations have not generally included coding for the operational definition for Indicator 14, therefore limiting their use for measuring it. Whitehurst (2003) provided information on rigorous experimental methods, but it is not specific to STEM education and is now out of date.

Near-Term Modifications and Activities

Depending on current procedures, ongoing Institute of Education Sciences (IES)/NSF portfolio analyses may need to be modified to capture more relevant data for this indicator and repeated to include recent award activity. SRI met with IES and NSF to explore the existing portfolio analyses that may be relevant for this indicator and recommend modifications that would yield enhanced data for this indicator. Kowalski and Taylor's DCL work will support this work by reviewing project abstracts and reports from the federal portfolios (NSF, IES, National Institutes of Health) and published articles in 22 journals (2001–present).

Additional Research Needs

Work is needed to develop a more precise operational definition for this indicator and an affiliated coding scheme that could be applied to the portfolio analyses and research reviews. Kowalski and Taylor will be undertaking this research as part of this DCL grant.

Long-Term Modifications and Activities

Effort is needed to conduct the portfolio analyses and research review on a reoccurring basis (perhaps every 5 years) to generate longitudinal data and assess educational impact.

NSF is working on articulating how it can characterize its portfolio to make this process easier in the future.

References

Kowalski, S., & Taylor, J. (2014). *Analyzing funding and publication patterns in STEM education research: Establishing a baseline for improved monitoring of research to advance STEM education*. National Science Foundation Award Abstract #1445540. Retrieved from http://www.nsf.gov/awardsearch/showAward?AWD_ID=1445540

Whitehurst, G. (2012, August). The Institute of Education Sciences: New wine, new bottles. Retrieved from http://ies.ed.gov/director/speeches2003/04_22/2003_04_22b.asp

Whitehurst, G. J. (2003). Research on mathematics education. Washington, DC: US Department of Education.

Currently Available Data → Near-Term Activites → Long-Term Activities | Additional Research Needs

Summary

Indicator 1 **Number of and enrollment in different types of STEM schools and programs in each district** **DCL Awards: 2**	• Upper estimate of number of STEM schools from CCD, CRDC, and/or SASS • Percentage of 9th-graders in school with special program emphasis in STEM from HSLS09 and HSLS12 • Schools with AP programs from CRDC	• Add items to NTPS to explore different features of STEM schools and programs • Examine relationship between STEM-focused schools and student outcomes using HSLS and SLDS	• Review recommendations stemming from DCL awards • Work with state data coordinators to harmonize definitions and build state buy-in to collect via CCD	• Effectiveness of different STEM-focused school models • Effectiveness of STEM programs within comprehensive schools vs. STEM-focused schools • Essential characteristics of STEM schools and programs
Indicator 2 **Time allocated to teach science in grades K-5** **DCL Awards: 2**	• Limited data on class time from previous administrations of the NAEP, TIMSS, SASS, and NSSME	• NAGB modified the 2015 NAEP science time item to match TIMSS write-in • Use NAEP Data Explorer tools to review time spent on science for high- and low- poverty schools before and after NCLB (compared with trends in math and ELA)	• Review recommendations stemming from DCL awards • Propose items for the NSSME or NAEP to measure instructional quality	• Accuracy of teacher self-reports (over- or under-reports) and conditions that lead to most reliable estimates • Allocation of multidisciplinary instructional time • Representativeness of selected grade level estimate(s) of other elementary grades
Indicator 3 **Science-related learning opportunities in elementary schools** **DCL Awards: 2**	• Limited data from NSSME about opportunities provided by schools • Data from four districts from Robert H. Tai Research Group on student participation in out-of-school activities	• Data available from NAEP TEL in 2015 on school offerings and student participation in activities • Modify survey items from Tai Research Group, HSLS, and/or PISA to propose for future NAEP or NTPS; include companion questions about criteria for student participation (e.g., offered to all students or selected grades)	• Review recommendations stemming from DCL awards • Prioritize appropriate recurring survey vehicle(s) and items	• Understandable survey items for students that generate accurate data • Core science-related learning opportunities most important for future STEM learning and interests

	Currently Available Data	Near-Term Activites	Long-Term Activities	Additional Research Needs
Indicator 4 **Adoption of instructional materials in grades K-12 that embody the CCSS-M and NGSS** **DCL Awards: 4**	• Textbook alignment research by Polikoff (USC), Schmidt (MSU), and Smithson (UW-Madison) • Partial data available from NSSME on most commonly used math and science instructional materials; data have not been coded for alignment with the CCSS-M or NGSS	• Code NSSME data on most commonly used instructional materials for alignment with the CCSS-M and NGSS • Propose items about most commonly used instructional materials for NAEP • Obtain information from EdReports.org on its textbook and curricula alignment work	• Review recommendations stemming from DCL awards • Convene group of experts to recommend alignment methods and appropriate schedule for applying	• Definition for nontextbook instruction materials and methods to analyze alignment of these materials • Pros and cons of different alignment tools and analysis techniques, including consistency of results across different approaches • Smaller scale research studies that examine link between instructional materials and student achievement
Indicator 5 **Classroom coverage of content and practices in CCSS-M and NGSS** **DCL Awards: 3**	• Data from NAEP, TIMSS, and ECLS-K on coverage of content and practices, but survey items not written to be specific to new standards • Enacted curriculum research by Porter (U Penn), Schmidt (MSU,) and Smithson (UW-Madison)	• Identify relevant items from NAEP, TIMSS, and ECLS-K that best address the CCSS-M and NGSS • Propose new items written for CCSS-M and NGSS for NAEP about coverage of content and practices	• Review recommendations stemming from DCL awards • Involve same experts who convene for Indicator 4 in framing of new survey items for Indicator 5	• Survey items for teachers that measure content coverage in concert with practices • Value and viability of approaches other than teacher self-report (e.g., student surveys, assessment review, instructional artifacts) at scale • Impact of adaptive/personalized learning on content and practices coverage
Indicator 6 **Teachers' science and mathematics content knowledge for teaching** **DCL Awards: 6**	• Data from Hill (Harvard) and the Measures of Effective Teaching (MET) project • Teacher perceptions of preparedness from NAEP, TIMSS, and NSSME • Baccalaureate & Beyond, HSLS, and TALIS data on college coursework	• Assemble and compare existing survey data and data from Hill and MET studies • Review and synthesize what is known about correlations between these measures and student achievement • Review proposed NTPS items re: PD and propose modifications if necessary	• Review recommendations stemming from DCL awards • Develop instruments to measure teachers' content knowledge for teaching science and high school math • Develop nonsurvey measures to get at knowledge in use	• Relationship between college backgrounds and self-reports of preparedness and direct assessments of content knowledge for teaching • Cost-effective measures for direct assessments at scale

	Currently Available Data	Near-Term Activites	Long-Term Activities	Additional Research Needs
Indicator 7: Teachers' participation in STEM-specific professional development activities **DCL Awards: 1**	• 2008 descriptive data on states' offerings of targeted PD from Education Commission • Data from ECLS-K, NAEP, NSSME, TALIS, TIMSS, and SASS about amount of science/math PD	• Data available in 2015 from NAEP TEL on common PD topics • Review results of Blank, Desimone, and Taylor meta-analyses as available • Propose items for NTPS to get at specific aspects of PD shown to impact classroom practice • Determine whether more PD data are available from districts or SLDS	• Review recommendations stemming from DCL award • With increasing role of online PD, explore use of back-end system data to measure quantity and quality of STEM PD	• Characteristics of PD that impact classroom practice • Impact of STEM-specific PD on student outcomes • Measuring the quality of STEM PD • Measuring formal and informal mentoring experiences
Indicator 8 Instructional leaders' participation in professional development on creating conditions that support STEM learning **DCL Awards: 1**	• No data sources were identified to inform this indicator	• Synthesize theories about principal leadership for instructional improvement • Review 2013 TALIS and new NCES School Climate Survey (currently under development) for relevant items • Provide input on NTPS principal survey	• Review recommendations stemming from DCL award • Consider the NAEP, NSSME, or NTPS as a potential target vehicle for collecting data on this indicator	• Qualities of principal PD that impact STEM teaching and learning practices • Types of principal PD related to student achievement
Indicator 9 Inclusion of science in federal and state accountability systems **DCL Awards: 1**	• Older CCSSO data from 2009–10 on whether science is assessed by states but not on inclusion in accountability system	• Obtain survey data from EDFacts on states' use of science assessments • Contact CCSSO, NGA, Education Week, and Assoc. of State Science Supervisors to see whether they are planning or willing to survey states on this topic • Examine Consolidated State Application Accountability Workbooks	• Review recommendations stemming from DCL award • Establish national-level database with information on state accountability systems • Produce a policy recommendation for frequency of data collection and suitable mechanism for collecting data	• Analysis of past and current state policies for types, uses, and frequency of science assessments (formative/summative) • Identification of important features of state accountability systems and coding system to determine science inclusion • Design of survey to collect data from state science coordinators

	Currently Available Data	Near-Term Activites	Long-Term Activities	Additional Research Needs
Indicator 10 **Inclusion of science in major federal K-12 education initiatives** **DCL Awards: 0**	• Information about frequency of NAEP science assessments vs. math and reading • OSTP 2010 Federal STEM Education Inventory Data set, limited to K-12 and cut by subject	• Develop a report for a general audience on the emphasis of science relative to ELA and math to generate interest • Wait until there is a reauthorized ESEA to analyze budget data, federal solicitations, and language in reauthorized ESEA	• Determine efficient mechanism for collecting and analyzing data on recurring basis	• Definitions for how program dollars are allocated by subject area (not by population served) • Role of STEM in federal initiatives of the years
Indicator 11 **State and district staff dedicated to supporting science instruction** **DCL Awards: 0**	• NCTM handbook on math specialists in VA • Campbell research on elementary math specialists • Data from SASS and ECLS-K on use of elementary subject specialists • Data from NSSME on coaching	• Determine whether CCSSO, SEAs, or science supervisors associations have data on intermediate-level supports • Determine whether NSTA, ASSM, AAAS, NASSS, Krehbiel at BOSE, and PAUSD Supt McGee are aware of ongoing research • Propose items for the NTPS about coaching from instructional specialists	• If related to student achievement, consider adding items about coaching to the NAEP	• Relationship between resources allocated for science support staff and student achievement • Trends in number of science support staff over last 5 years
Indicator 12 **States' use of assessments that measure the core concepts and practices of science and mathematics disciplines** **DCL Awards: 1**	• Descriptive data from 2009 on states' assessments from the CCSSO 50 State report and the Fordham Institute's State of Science Standards 2012	• Research by CRESST and RAND on CCSS-aligned assessments • Contact CCSSO, NGA, Education Week, and Assoc. of State Science Supervisors to see whether planning or willing to survey states on this topic	• Review recommendations stemming from DCL award	• Pros and cons of different alignment tools and analysis techniques, including consistency of results across different approaches • Use RAND and CRESST methods for non-CCSS math and science assessments • Adapt RAND and CRESST's methods for science practices

	Currently Available Data	Near-Term Activites	Long-Term Activities	Additional Research Needs
Indicator 13 **State and federal expenditures dedicated to improving the K-12 STEM teaching workforce** **DCL Awards: 0**	• Survey of the Use of Funds under Title II data on district funds spent on science PD • 2010 OSTP data on federal STEM expenditures • Ed. Commission 50 State Analysis info on high school-level STEM initiatives, including STEM-specific PD	• Propose items for NTPS about unfilled STEM teaching jobs and compensation for STEM-focused service • Seek additional data sources: Determine whether districts have data and/or SLDS infrastructure that might be used to capture this indicator • Research existing quantitative/qualitative information, such as from OSTP or BLS	• Work with NCES Data Forum and CEDS Stakeholder Group to develop common expenditure definitions and reporting structures • Determine what kind of data would be most useful to policymakers	• Other aspects of workforce development to measure besides PD and associated classification scheme • Strategies to build buy-in among state system developers • Consistency of funded PD with best practices
Indicator 14 **Federal funding for the research identified in *Successful K-12 STEM Education*** **DCL Awards: 1**	• NSF has funded evaluations of DRK-12, ITEST, and ISE portfolios, but these evaluations have not coded for this indicator • Whitehurst data from 2003 on rigorous experimental methods but not specific to STEM and now outdated	• Meet with IES and NSF to explore/modify existing portfolio analyses and support the DCL award efforts	• Review recommendations stemming from DCL award • Conduct portfolio analyses and research review on reoccurring basis to generate longitudinal data and assess educational impact review recommendations stemming from DCL awards	• Refined definition for indicator and affiliated coding scheme for portfolio analyses and research review • Programs and practices that contribute to student learning and to other goals of STEM education

Appendix A: SRI Data Collection and Consulting Activities

The initial information-gathering and consulting phase consisted of two primary activities, each of which is described in additional detail below:

1. Convening an external advisory group of experts who specialize in science education policy, education databases, ongoing education data collections, and data collection methodologies
2. Interviewing stakeholders from NSF and other key organizations concerning available data and desired data uses for the 14 indicators.

Advisory Group

SRI convened an advisory group of national experts in STEM education policy and practice, education databases, data analysis, indicator systems, and education staffing and finance (see table below). In collaboration with NSF staff, invited experts, and SRI staff working on the indicators, the advisory group is guiding indicator system data gathering, collaboration, and reporting efforts.

Advisory Group Member	Affiliation
Adam Gamoran	William T. Grant Foundation
Alan Friedman (now deceased)	Consultant
Bill Schmidt	Michigan State University
Bill Tate	Washington University of St. Louis
Dan Goldhaber	University of Washington
Drew Gitomer	Rutgers University
Irwin Kirsch	Educational Testing Services
Joe Krajcik	Michigan State University
Laura Desimone	University of Pennsylvania
Mike Smith	Carnegie Foundation
Natalie Nielsen	National Research Council
Peggy Carr	Institute of Education Sciences

SRI hosted an advisory group meeting on October 24, 2013, which included full-group discussions and breakout activities related to the project's history and scope, the set of 14 indicators recommended in the *Monitoring Progress* report and their operational definitions and potential data sources, prospective stakeholders to interview to inform the development of the indicator system and interview protocols, and plans for ongoing collaboration and future meetings.

SRI held a virtual advisory group meeting on April 17, 2014, to review progress to date, solicit feedback on the results on the stakeholder interviews, and receive guidance on next steps, including prioritizing potential activities. Advisors were also invited to participate in a follow-up feedback activity to collect additional information.

SRI hosted a second in-person advisory group meeting on October 2–3, 2014, to collect feedback on the current, near-term, and long-term activates for measuring the indicators outlined in the draft road map and to devise strategies for broadening the impact of NSF's STEM Indicators efforts, including potential topics for concept papers.

Dear Colleague Letter Grantees

Knowing that **some** of the Indicators require further conceptual work and some preliminary research before a data collection plan can be solidified, NSF issued a Dear Colleague Letter (DCL) to fund a series of EAGER grants to advance knowledge about how to measure the indicators. The EAGER funding mechanism may be used to support exploratory work in its early stages on untested but potentially transformative research ideas or approaches. Requests may be for up to $300,000 and of up to 2 years in duration. Fifteen EAGER grants were awarded through the DCL (see table below).

EAGER Award	Indicator(s) Addressed
Identifying and Measuring STEM Schools and Programs PI: Jeanne Rose Century, University of Chicago Co-PI: Melanie LaForce	1
Defining, Measuring, & Monitoring Adequate Instructional Time and Resources for Science in Grades K-5 PI: Rena Dorph, University of California, Berkeley Co-PI: Ardice Hartry	2, 3
Developing Consensus Guidelines for Tools and Processes that Assess Science Instructional Materials PI: April Gardner, Biological Sciences Curriculum Study Co-PIs: Audrey Mohan, Jody Bintz	4
An Online System for the Collection of Textbook Adoption Data PI: Morgan Polikoff, University of Southern California	4
Classroom Implementation Indicators for K-12 Common Core State Standards in Mathematics (CCSSM) PI: William Schmidt, Michigan State University	4, 5
Classroom Artifacts as Indicators of Quality in STEM Education PI: Drew Gitomer, Rutgers University New Brunswick	4, 5
Measuring Classroom Coverage of Content and Practices in the New Generation of Mathematics and Science Standards PI: Laura Hamilton, RAND Corporation Co-PIs: Brian Stecher, Kun Yuan	5
Operationalizing the Science and Engineering Practices PI: Eric Banilower, Horizon Research Inc. Co-PIs: Joan Pasley, Peggy Trygstad	5
Assessing the Use of Licensure Tests as an Indicator of Teachers' Science and Mathematics Content Knowledge for Teaching PI: Dan Goldhaber, University of Washington	6
Building Understanding and Measurement of Secondary Mathematical Knowledge for Teaching (MKT) PI: Heather Howell, Educational Testing Service Co-PI: Geoffrey Phelps	6

EAGER Award	Indicator(s) Addressed
Building Understanding and Measurement of Secondary Mathematical Knowledge for Teaching (MKT) PI: Y.-J. Yvonne Lai, University of Nebraska-Lincoln	6
Adapting the Classroom Video Analysis Approach as a Feasible and Scalable Measure of Common-Core-Aligned Mathematics Knowledge for Teaching PI: Nicole Kersting, University of Arizona	6
Moving beyond Subject Matter Knowledge: Assessing Elementary Teachers' Content Knowledge for Teaching Science PI: Jamie Mikeska, Educational Testing Service Co-PI: Geoffrey Phelps	6
An Exploration of the Alignment of SLDS Infrastructure and Data Highway to Relevant Success Indicators in Mathematics and Science PI: Ellen Mandinach, WestEd Co-PI: Martin Orland	1, 2, 3, 6, 7, 8
Developing a System for Tracking State Assessment Policies in Science and Mathematics Education PI: Rolf Blank, National Opinion Research Center	9, 12
Analyzing Funding and Publication Patterns in STEM Education Research: Establishing a Baseline for Improved Monitoring of Research to Advance STEM Education PI: Joseph Taylor, Biological Sciences Curriculum Study Co-PI: Susan Kowalski	14

Stakeholder Interviews

The research team has been meeting with stakeholders from key organizations that collect, report, or use K-12 STEM education data. See the table below for a list of completed interviews.

Researchers

SRI interviewed experts with experience measuring one or more of the indicators to help inform operational definitions for terms in that need further definition (e.g., STEM schools), identify what sources of information for each indicator currently exist or could be adjusted to obtain the necessary information, and determine what additional research might be needed to ensure a robust indicator. One to five researchers have been interviewed per indicator.

Policymakers

SRI interviewed two congressional staffers and three additional policymakers to ascertain the level and locus of interest in the indicators, determine how to make the indicator information most useful to policymakers, and collect recommendations for engaging policy groups. Additionally, SRI held a brown bag session at NSF to raise awareness about the indicators work and discuss potential ways to support foundational research needed to further develop the indicators.

Data Collection	Completed Interviews
Indicator 1	Jeanne Century Robert Tai Sharon Lynch
Indicator 2	Carla Zembal-Saul Karen Worth
Indicator 3	Beth Warren Bill Penuel Carla Zambal-Saul Karen Worth Robert Tai
Indicator 4	John Smithson Mary Ann Huntley Melanie Cooper Nancy Songer
Indicator 5	Anthony Bryk Ed Silver John Smithson Mary Ann Huntley
Indicator 6	Bill Penuel Heather Hill
Indicator 7	Joseph Taylor
Indicator 8	Anthony Bryk Ed Silver
Indicator 9	John Easton
Indicator 10	John Easton
Indicator 11	Ed Silver Jane Hannaway
Indicator 12	Ed Silver Jim Pellegrino Joan Herman Rolf Blank
Indicator 13	David Monk
Indicator 14	David Monk John Easton
Policymakers	Chris Swanson Deborah Ball Jane Hannaway Soncia Coleman (Harkin staff) Thomas Culligan (Wolf staff)

Data Collection	Completed Interviews
Statistical Agencies	NAGB: NAEP NCES: CCD, CEDS, CRDR, NAEP, SASS/NTPS, SLDS Horizon: NSSME

Statistical Agencies

SRI organized meetings with key statistical entities to discuss the feasibility and costs of expanding or modifying data collections to collect more relevant data for the indicators. Meetings were held with the National Assessment Governing Board (NAGB), Horizon Research, and the National Center for Educational Statistics (NCES). SRI also led a webinar for the NCES Forum to disseminate information about the indicators work to its members and solicit advice from state and district data experts for leveraging these data systems to measure the indicators at the national and local levels.

Appendix B: Priority Recurring Data Collection Summary

Data Source	Agency	Sampling	Grades	Representative of…	Respondent(s)	Frequency	Last Admin.	Next Admin.	Level(s) of Estimates	Student Outcomes	Informs Indicator(s)
Baccalaureate and Beyond Survey, 2008 Cohort	NCES	Sample	Postsecondary	Recent recipients of bachelor's degrees	Graduates	Longitudinal	2012	n/a	National	No	6
Common Core of Data (CCD)	NCES	Census	K-12	Schools	State education agencies	Annual	2014	2015	National State District School	No	1
Civil Rights Data Collection (CRDC)	NCES	Varies	K-12	Schools	Schools Districts	Every 2 years	2014	2016	National State District School	No	1
Early Childhood Longitudinal Program (ECLS) Kindergarten Class of 2010-11 (ECLS-K:2011)	NCES	Sample	K followed through grade 5 or 8	Kindergarten class of 2011	Student Teacher Parent Sch. admin. Care providers	Longitudinal	2014	2015	National	Yes	1, 2, 5, 6, 7, 11
High School Longitudinal Study of 2009 (HSLS)	NCES	Sample	Grade 9 followed through secondary and postsecondary	9th-grade class of 2009	Student Teacher Parent Sch. admin. Sch. counselor	Longitudinal	2012	2016	National State (10)	Yes	1, 6, 13
National Assessment of Educational Progress (NAEP) Math Contextual Variables	NAGB	Sample	Grades 4, 8, and 12	4th-, 8th-, and 12th-graders	Students Teachers Schools	Every 2 years	2013	2015	National State (most)	Yes	5, 6, 7
NAEP Science Contextual Variables	NAGB	Sample	Grades 4, 8, and 12	4th-, 8th-, and 12th-graders	Students Teachers Schools	Every 4 years	2011	2015	National State (most)	Yes	2, 5, 6, 7

Data Source	Agency	Sampling	Grades	Representative of…	Respondent(s)	Frequency	Last Admin.	Next Admin.	Level(s) of Estimates	Student Outcomes	Informs Indicator(s)
NAEP Technology and Engineering Literacy (TEL) Contextual Variables	NAGB	Sample	Grade 8 (2014) Grades 4, 8, and 12 (future)	4th-, 8th-, and 12th-graders	Students Schools	Every 4 years	2014	2019	National State (most)	Yes	3
National Survey of Science and Mathematics Education (NSSME)	Horizon Research	Sample	K-12	Teachers	Teachers Program coord.	Varies	2012	n/a	National	No	2, 3, 4, 6, 7, 11
National Teacher and Principal Survey (NTPS; formerly SASS)	NCES	Sample	K-12	Schools	Teacher Principal	Every 2 years	n/a	2015	National	No	1
Programme for International Student Assessment (PISA)	OECD	Sample	15-year-olds	Countries	Students Principals	Every 3 years	2012	2015	National	Yes	3
Schools and Staffing Survey (SASS)	NCES	Sample	K-12	Schools	Teacher Principal School	Every 4 years	2012	n/a	National State	No	1, 2, 6, 7, 11
Survey of the Use of Funds Under Title II	ED	Sample	K-12	Districts	Districts	Annual	2014	2015	National	No	13
Teaching and Learning International Survey (TALIS)	OECD	Sample	Lower secondary	Teachers	Teachers Schools	Every 5 years	2013	2018	National	No	6, 7
Trends in International Mathematics and Science Study (TIMSS)	NCES	Sample	Grades 4, 8	Countries	Student Teacher School	Every 4 years	2011	2015	National	Yes	2, 5, 6, 7

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

http://www.sri.com/education