

Progress Report on Implementation

Development and Implementation of a STEM K-12 Indicator on State Assessment Policies in Science and Mathematics Education

SEPTEMBER 2017

PRESENTED TO:
National Science Foundation
4201 Wilson Boulevard
Arlington, VA 22230

GRANT #1544123

PRESENTED BY:
Rolf K. Blank
Principal Investigator
NORC at University of Chicago
4350 East-West Highway
Bethesda, MD 20814-4499

Development and Implementation of a STEM K-12 Indicator on State Assessment Policies in Science and Mathematics

NORC at the University of Chicago

Prepared under National Science Foundation PRIME program grant (#1544123)

September 2017

Rolf K. Blank, Principal Investigator
NORC at University of Chicago
Education and Child Development Department
Contact: Blank-Rolf@norc.org
[301 634 9325](tel:3016349325)

In 2015, NORC at the University of Chicago (NORC) submitted a proposal to the National Science Foundation program on Promoting Research and Innovation in Methodologies for Evaluation (PRIME) of the Division of Research on Learning. A three-year grant was awarded to NORC for the project, **STEM Indicators: Implementing a Model for Reporting and Research on State Assessment Policies for K–12 Science and Mathematics Education** (#1544123). The project began in September 2015 and continues for three years. The NORC online reporting system with states and the public website <http://stem-assessment.org> are continuing to operate in 2017-18.

The goal of the project is to develop and implement an online system for reporting state assessment policies across states. The reporting system and website inform mathematics and science educators and leaders at state, local, and national levels about the status of state assessment policies across the nation. The website will provide comparable policy-relevant information that can be tracked as changes are implemented.

Prior Research Informing Development

The STEM Indicator on state assessment policies reflects the strong interest in advancing the quality of student assessments for mathematics and science education. Research and analysis of the content of state achievement tests has yielded concerns about their academic rigor, as the items tend to assess lower-level cognitive skills (such as recall, recognition, and application of procedures) as opposed to higher-level cognitive skills (such as analysis, evaluation, and synthesis of ideas) (Darling-Hammond, et al., 2013; Resnick, et al., 2003; Webb, 2002). Even in the NCLB era, with its emphasis on accountability, researchers have found state tests generally assess lower-level cognitive skills. In a recent study of tests in 17 states that were selected because the tests were reputed to be more cognitively demanding, Yuan and Le (2012) found that fewer than 2 percent of mathematics items assessed a higher level of cognitive demand. Using a different methodology on 19 state tests, Polikoff, Porter, and Smithson (2011) found that only seven percent of mathematics items required students to use higher-order cognitive skills, and fully 80 percent of mathematics items assessed memorization, recall, and use of routine procedures.

The lack of test items assessing more advanced analysis and knowledge has resulted in reforms to state content standards and assessments, with the most prominent reform being the development of the Common Core State Standards (CCSS-M) for mathematics and the Next Generation Science Standards (NGSS) for science (NGA/CCSSO, 2010; Achieve & Lead States, 2013). The CCSS-M and NGSS have the potential to improve the degree to which deeper learning is assessed through state achievement tests. The CCSS-M establishes a single set of educational standards for kindergarten through 12th grade that identifies the concepts and knowledge that students should

acquire to show that they have attained the skills necessary for college and career success (Yuan & Le, 2014). Similarly, the NGSS initiative focuses on coherency and progression of core conceptual science skills from early grades through high school, with an emphasis on preparing students for college, career, and citizenship. A recent study report and recommendations from the National Research Council is informing the development of new designs and systems for science education in the states to support the goals of the NGSS (NRC, 2013). Thus, a key focus for the design of reporting on state assessments is their degree of alignment to content standards, and particularly the extent to which math and science assessments are measuring higher-level cognitive demand or cognitive complexity in student learning. It will be important to track which assessments states have selected and how they are aligned to standards, and additionally how decisions on assessments change over time.

Applying Design Study Results

NORC conducted a design study in 2014-15 to test a model for collecting information on states' policies for state student assessments. The study examined: a) the types of policy information and measures that can and should be reported, and b) tested methods of analyzing and reporting information on alignment of state assessments to content standards, including the Common Core State Standards-Mathematics and the Next Generation Science Standards. The study piloted a design for collection of state policies information from three sources—state websites, a survey of state assessment staff and subject specialists (focusing on design, content, and reporting methods), and recent research studies on trends in mathematics and science assessments. The design options and issues of focus in the pilot study are summarized in a paper, developed from the recommendations of a project expert panel (NORC, 2015).

A key issue analyzed in the design study was tracking states use of innovative features of the mathematics assessments being provided through the two state assessment consortia—Smarter Balanced (SB) and PARCC. The consortia assessment systems include performance tasks (with multiple steps and explanation of work), balance of items across the different levels of depth of knowledge, use of computer-based testing to improve turnaround time for scoring and reporting, measures of learning content and skills, use of adaptive testing to measure a full range of student knowledge, benchmark testing to track progress through the year, and a digital assessment library for formative classroom assessment. To work toward an indicator of state assessments alignment to standards especially relative to content coverage and levels of cognitive demand, NORC staff worked gained input from technical advisers and consulted key resources such as the report on

“Criteria for High-Quality Assessments” (Darling-Hammond, et al., 2013) and a report from CCSSO on evaluating student assessments (2014).

A key measure to be reported across states is the alignment of state math and science assessments to state content standards. Many states conducted alignment studies to support and validate the assessment instruments used for accountability reporting as part of the requirements under federal No Child Left Behind funding, and currently all states are preparing submissions to the U.S. Department of Education under ESSA which present a rationale for state student assessments that will provide information for accountability and other uses of data. A recent study provided an analysis of the quality of the state consortia assessments including alignment to the Common Core State Standards and the study included analysis of two other assessments used by states (Doorey & Polikoff, 2016). For reporting on state assessments policies as a STEM K-12 indicator, NORC determined through review with the Technical Advisers that it would not be possible to conduct an analysis of the alignment of each state’s assessment instruments in math and science. Instead, the advisers recommended a design for collecting information on several measures of the degree of alignment that could be compiled and reported across states that would address key differences and similarities in the *breadth* and *depth* of state assessments’ content relative to standards.

Reporting of state policies and trends over time is equally important for science education as part of the state assessment policies STEM K-12 indicator and reporting system. The development of the Next Generation Science Standards (NGSS, 2014) resulted in many state adopting new science education standards and produced a strong interest in new forms of student assessment to match the direction of the standards. Some states are following the recommendations of the NRC committee on NGSS science assessments (NRC, 2014), including moving toward a system of science assessment with varied methods, assessing the three dimensions of science instruction, and maintaining annual reporting of student progress. The design pilot study was carried out with support and participation by staff of ten state departments of education. The results of the design study are available from NORC (<http://www.norc.org/Research/Projects/Pages/developing-a-system-for-tracking-state-assessment-policies-in-science-and-mathematics-education.aspx>).

Measures of State Assessment Policies for a Reporting System

Conceptual Framework. From our review of prior research and the results of the design study, NORC established a conceptual framework for selecting state assessment policies and measures that address improved quality of student assessments in science and mathematics. The framework includes the following elements:

- Measure progression of student knowledge and skills in science and mathematics toward college and career readiness;
- Degree to which assessments have content validity and alignment in relation to state content standards in mathematics and science;
- Assessment instruments include varying levels of cognitive complexity or depth of knowledge, and defined cognitive levels are used in assessment development;
- States provide support and leadership for development of science and mathematics assessment systems, which may include summative assessments, interim/benchmark assessments, and formative assessments in classrooms.

The conceptual framework was applied by NORC in working with a group of technical advisers to carefully consider findings from research on state assessments and to evaluate the results of the design pilot study which included feedback from state assessment leaders and specialists in science and math education. The measures for state reporting were begun with the 2015–16 school year. In addition to designing questions linked to the conceptual framework, a key decision in planning the reporting system was the degree to which state information should be comparable and quantifiable. For purposes of cross-state reporting, a common metric and organization of information in tabular format facilitates comparisons; however, with greater depth of information, key differences in state policies and design of state assessments could be highlighted and shared among users in different states.

Measures of Current State Assessments and Development of New Assessment Systems. The advisers recommended collecting and reporting information for the following measures of the design of state assessment policies:

- Types of state assessments by grade level;
- Intended uses of assessment data, including accountability, high school graduation; improving curriculum/instruction, and college and career readiness;
- Item or task design;
- Timing and methods of administration;
- Methods of assessment reporting and dissemination.

Alignment of State Assessments to Standards:

- State policy on content standards, and relationship of state standards to CCSS and NGSS;
- State alignment studies and technical reports available to public;
- Content topics and levels of cognitive demand by grade level of assessment;
- Progression of assessed content and expectations across grades;
- State plans and strategies for development of new assessments and state assessment system.

Online Survey and Reporting Process: Year One

NORC staff designed and implemented the reporting instrument as planned for spring 2016. The questions were developed by NORC with input from the technical advisers in fall 2015, with the results of the design pilot study providing evidence for design. The survey questions were written in a web software program and the online instrument was pre-tested with respondents from two participating states (Washington and North Carolina). Cognitive interviews were carried out with four SEA staff who participated in the pre-test to determine if they had problems with understanding or interpretation of survey questions, or had any difficulties in accessing or moving through the instrument which was sent to them as a NORC weblink. Based on responses question edits were made and directions clarified. The online instrument (which has unique web addresses for each participating state), was sent via email to SEA contacts in mid-May 2016. NORC requested that the online reporting be completed and submitted by mid-June 2016. Twenty state education agencies completed the report for the first-year implementation.

Several issues were identified regarding the online reporting process via feedback from state contacts and NORC review of data. The design needed to have a search function that would allow respondents to move to a specific question or section, especially with multiple respondents involved and for a respondent who saves entries and then wants to return to the system to complete. A guide that indicates what has been completed/remaining could appear along the margins of a computer screen to help respondents track their progress. Review of responses by NORC indicated that definitions of types of assessments, e.g., EOC vs. comprehensive assessment, should be clarified further because some of the entries were misplaced.

Validation of State Policy Information Reported. The policies information reported by states was validated through several methods. The responses to the online survey for each state education agency were submitted by multiple staff within an agency according to their specialization. The individual responses were reviewed internally by the agency contact person and submitted as one agency online survey. The survey responses were reviewed by NORC project staff and checked against other state-level information sources (e.g, that was downloaded from state websites or reported through a third source such as other 50-state reporting sites (Any inconsistencies identified are checked with the state education contacts and edits are made in the data file. The cross-state policies information was reviewed by the Technical Advisers and any questions were forwarded to the NORC principal investigator who determined if follow-up with state contacts were needed. Finally, the policies and SEA responses were organized in tables

with policy information comparable by state, and each contact person in the state education agencies reviewed the reporting tables and submitted edits to the principal investigator.

The State Assessment Policies website for reporting system (<http://stem-assessment.org>) is organized for two primary audiences and users of policy information: A) policy-makers, B) science and mathematics educators. The use of information organized for these two types of audiences is likely to overlap. The policy-makers focus provides summary information on types of assessments being implemented under state policies and the major intended uses of the assessments. The S-M educators focus section provides more detailed information on state policies and further cross-state analysis of patterns and trends among the states.

Findings from Year-one Implementation: 20 states. One summary finding from year one was that state policies cover both state-required assessments in math and science and support of student assessments that can be selected for use by districts or schools. Benchmark assessments are supported by states primarily in mathematics. About half the states have policies supporting formative classroom assessment initiative in math and science.

At the high school level, 11 of the 20 states have adopted the ACT or SAT as a student assessment requirement, with some of these states reporting the intended use for school accountability as well as use for college and career readiness. Several of the states reported their state had shifted policy from requiring the SBAC or PARCC high school assessments to the ACT or SAT. Some of these states also require an end of course (EOC) high school assessment for school accountability reporting. In total, 17 states require a comprehensive high school mathematics assessment (usually at grade 11, testing more than one math content area), and 13 states require a comprehensive science assessment (more than one science content area). EOC assessments are given by 13 states for high school science, and by nine states for high school math courses.

The measures of mathematics assessment content alignment to standards provided useful findings. Fifteen of the 20 reporting states in Year 1 reported on the types of items/tasks used in math and science assessments, and the data show that in mathematics a majority of states have incorporated multiple types of items and tasks. The reporting indicates that states working with the assessment consortia and states with computer-based instruments have assessments with multiple item/task types and thus rely less on multiple choice item formats. All of the states reported that at least three cognitive levels are defined for development of the state-administered assessments, however only five states use the cognitive levels in reporting assessment results. A review of the reporting on grade 5 math assessment content topics showed that six of the selected topics drawn from the CCSS-M are assessed in each of the reporting states. States with computer-adaptive assessments reported a range for percentage of assessment points per topic since students

are assessed differently according to their performance. Review of reporting on high school math topics showed wide variation across states in coverage of six selected high school math topics. Several states reported a substantial portion of the assessment points were in additional topics. End-of-course assessments were not included in the reporting on math high school topics.

The information reported on science assessment content focused primarily on state plans and timelines for development of new science assessments that will be aligned to state science standards. The online survey asked about whether state standards developed or revised since 2013 and whether they were by the NRC Framework (2012) or the NGSS (2013). States also reported on plans for assessing three dimensions of science education content and strategies for assessment design to meet the standards content and for sampling across standards. More than half the reporting states indicated that decisions were not complete regarding new standards-aligned assessments, with the school year 2018–19 frequently reported as the target operational date.

Online Survey and Reporting: Improvements in Year Two

In the second year of the state policies online reporting of state assessment policies, the objective was to report on math and science assessment policies in all 50 states and DC. The web-based survey instrument tailored to each state were sent to contact persons in all 50 states, and the website for reporting was revised to include information for all states.

Reporting and Validation of State Policies (2016-17). The survey responses were reviewed by NORC project staff and checked against information that was downloaded from state websites or reported through a third source, including other projects with 50-state information on standards and assessments (CSAI, 2017; C-SAIL, 2017). If states reported policies for the first year (2015-16) there were re-sent and states could revise or edit responses for year two (2016-17). The project is voluntary. NORC re-contacted state education agency staff for any missing items, and the membership lists of the State Mathematics Supervisors, State Science Supervisors, and State assessment directors were used to provide follow-up information and requests for completion. Tabular information for each policy measure reported was sent to state education contact persons for review and editing. The state information was collected from March 2017 through June 2017, and all the responses were checked and validated by September 2017.

Website Design and Format. During year two, NORC presented the first year online assessment reporting results to state science supervisors and state math supervisors for review, feedback, and comment. These key user groups were asked to report on policy measures they found useful and informative, and they were also asked to review the design and format for reporting and make suggestions about the usefulness of the website. The same questions were

posed to the Technical Advisers. A number of the recommendations were used in revisions and improvements to the <http://stem-assessment.org> webpage for reporting on the 50-state information for 2016-17. The categories by which the policies are reported by state were revised to simplify and clarify the organization. The access to each table and return to the text for description and explanation was improved. A Summary of findings was added as a separate section. The Home page was reorganized to provide a clear message of purpose, and further information was provided in the appendix to provide explanation of the process of development, data collection and reporting. Graphics were added to each section of the site and to the Home page, and the survey of users was clarified for intent.

References

Achieve & Lead States (2013) *Next Generation Science Standards: By States, For States*. Washington, D.C.: National Academies Press. <http://www.nextgenscience.org/next-generation-science-standards>

Center on Education Policy (2012) *State High School Exit Exams: A Policy in Transition*. George Washington University, Washington, DC. <http://www.cep-dc.org/displayDocument.cfm?DocumentID=408>

CSAI (2017) Center on Standards and Assessment Implementation Center led by WestEd and UCLA/CRESST under a contract with the U.S. Department of Education <https://www.csai-online.org>,

C-SAIL (2017) Center on Standards, Alignment, Instruction and Learning, C-SAIL, led by researchers at Graduate School of Education, University of Pennsylvania, University of Southern California, and AIR. gse-csail@gse.upenn.edu

CCSSO's (2014) *Criteria for Procuring and Evaluating High-Quality Assessments* <http://www.ccsso.org/Documents/>.

CCSSO/ATP (2013) *Operational Best Practices for Statewide Large-Scale Assessment Programs*. http://www.ccsso.org/Resources/Publications/Operational_Best_Practices_for_Statewide_Large-Scale_Assessment_Programs.html

Darling-Hammond, L., Herman, J., Pellegrino, J., Abedi, J., Baker, E., Bennett, R. et al. (2013). *Criteria for high-quality assessment*. Stanford, CA: Stanford Center for Opportunity Policy in Education.

Doorey, N. & Polikoff, M. (2016) *Evaluating the Content and Quality of Next Generation Assessments*. Washington, DC: Fordham Institute. <http://edexcellence.net/publications/evaluating-the-content-and-quality-of-next-generation-assessments>

National Governors Association Center for Best Practices and Council for Chief State School Officers (2010) *Common Core State Standards*. Washington, DC: National Governors Association Center for Best Practices and Council of Chief State School Officers. <http://www.corestandards.org/>

National Research Council (2012) *Framework for K–12 Science Education*. Washington, DC: National Academy Press. http://www.nap.edu/openbook.php?record_id=13165

National Research Council (2013) *Monitoring Progress toward Successful K–12 STEM Education: A Nation Advancing?* Washington, DC: National Academy Press.

National Research Council. (2013). *Developing Assessments for the Next Generation Science Standards*. James W. Pellegrino, Mark R. Wilson, Judith A. Koenig, and Alexandra S. Beatty, Editors; Committee on Developing Assessments of Science Proficiency in K–12. Board on Testing and Assessment. Washington, DC: The National Academies Press.

NORC (2015) Study Design Paper: Developing a System for Tracking State Assessment Policies in Science and Mathematics Education Instrument Design and Pilot Study Plans, Bethesda, MD. March.

Polikoff, Porter, and Smithson (2011) How well aligned are state assessments of student achievement with state content standards? *American Educational Research Journal*, 48(4), 965–995.

Resnick, L.B., Rothman, R., Slattery, J.B., & Vranek, J.L. (2003). Benchmark and alignment of standards and testing. *Educational Assessment*, 9, 1–27.

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

Webb, N. L. (2002). *Alignment study of language arts, mathematics, science, and social studies of state standards and assessments in four states*. Washington, DC: Council of Chief State School Officers.

U.S. Dept. of Education's Peer Review Guidance (issued 9/25/15)
<http://www2.ed.gov/policy/elsec/guid/assessguid15.pdf>

Yuan, K. & Le, V. (2012). *Estimating the percentage of students who were tested on cognitively demanding items through the state achievement tests*. Santa Monica, CA: RAND Corporation.