Paper Title Exploration of the Alignment of State Data and Infrastructure to Mathematics and Science Success Indicators

Author(s) Ellen B. Mandinach, WestEd; Shandy Hauk, WestEd

Session Title Improving Indicators of K–12 STEM Education Policy and Practice: What Have We Learned?

Session Type Session Paper

Presentation Date 4/28/2017

Presentation Location San Antonio, Texas

Descriptors

Methodology

Unit Division L - Educational Policies and Politics

DOI 10.302/1181686

Exploration of the Alignment of State Data and Infrastructure to Mathematics and Science Success Indicators

Shandy Hauk and Ellen B. Mandinach
WestEd

For decades, state level data gathering and reporting has been evolving from compliance-based warehousing of information to creating and maintaining complex systems that can provide information to shape diverse kinds of educational decision-making. Newer systems curate data for federal and evaluative reporting as well as for instructionally relevant local decision-making. Statewide longitudinal data systems (SLDSs) for education are often designed by database vendors (though some are still "home grown" and can be revised by SLDS employees). In parallel with the technology revolution that allowed more sophisticated management of data, there is now a sharper focus of national attention on science, technology, engineering, and mathematics (STEM) education.

Purpose

Monitoring Progress Toward Successful K-12 STEM Education (National Research Council, 2013) described 14 indicators that could be used as common touchstones, across states, to monitor local and national progress in STEM education and for making evidence-based decisions about improving it. The indicators serve the dual purpose of providing a common language for states to communicate internally and externally about STEM education efforts and as foundation for national-level reporting that could support progress towards the NRC's three goals for education in the STEM disciplines:

Goal 1. Expand the number of students who ultimately pursue advanced degrees and careers in STEM fields and broaden the participation of women and minorities in those fields.

Goal 2. Expand the STEM-capable workforce while broadening the participation of under-represented groups in that workforce.

Goal 3. Increase STEM literacy for all students, including those who do not pursue STEM-related careers or additional study in the STEM disciplines.

Measuring progress in each indicator requires data. The driving question for our work is whether current **SLDSs** collect data sufficiently robust, comprehensive, and aligned to the Indicators to allow monitoring of progress in the areas identified by the Indicators. And, if not, what are next steps to leveraging SLDS structures to evaluate national progress?

The focal STEM Indicators for the project (as numbered in *Monitoring Progress*):

1. Number of, and enrollment in, different types of STEM schools and programs, by district;
2. Time allocated to teach science in grades K-5;
3. Science-related learning opportunities in elementary schools;
6. Teachers' science and mathematics content knowledge for teaching;
7. Teachers' participation in STEM-specific professional development activities; and
8. Instructional leaders' participation in PD on creating conditions that support STEM learning.

Perspective

Accumulation and distribution of data

The primary mechanisms for educational data collection are school- and district-level staff – from teachers and classroom aides who administer tests or complete forms to administrative leaders and support personnel who interact with the local database. Traditionally, these local data flow one way, into the larger stream of data that goes to the state or the massive river of data that flows into federal data caches. Given the information technology that has become available over the last 20 years, a new analogy for data management has emerged, "the data superhighway." Now, data can flow up and down across educational hierarchies, allowing for roadside stops that provide snapshots of the nature of information at a particular place or time. Nonetheless, it is still the case that the superhighway is often seen as one-way, leading from the districts to the states and ultimately to the federal level, with little or no data or information returning back to the districts and schools for locally-relevant use. Many staff at local education agencies are as yet uncertain how the data they are providing in this superhighway have relevance to the kinds of decision-making they need to make.

Each SLDS has hundreds of **data elements** (e.g., hours/week spent by each middle school student in math; number of hours of professional development (PD) and title of each PD course completed by each teacher or leader/principal). Some of these are required by the federal government ED*Facts* system. ED*Facts* contains consistent data across states, but local data systems typically contain unique data elements in addition to federally mandated ED*Facts* elements. Additional data elements may be unique to specific data system vendors and others to states. The SLDSs are populated by district data warehouses that contain even more locally relevant data. However, much local data stay local. That is, it is common that only compliance and accountability data required for collection through ED*Facts* get sent along the data highway.

Remark: ED*Facts* and Outcomes Data

As noted on its website, the U. S. Department of Education ED*Facts* initiative combines performance data from K-12 state education agencies with other data assets, such as financial and grant information (http://www2.ed.gov/about/inits/ed/edfacts/index.html). The purpose of ED*Facts* is to:

- Place the use of robust, timely performance data at the core of decision and policymaking in education.
- Reduce state and district data burden and streamline data practices.
- Improve state data capabilities by providing resources and technical assistance.
- Provide data for planning, policy, and management at the federal, state, and local levels.

States spend time and effort on data submissions that comply with ED*Facts* data element descriptions. This is good for many reasons. However, the focus on outcomes and performance does not include attention to educational and instructional process data.

The goal of the STEM Indicators is to specify aspects of the educational experience of students, teachers, and administrators in order to attend to and improve those experiences. The outcomes reported to ED*Facts* are necessary but not sufficient for that specification. Additional information is needed on educational *processes,* such as: how courses in a STEM school or program are structured to provide greater access and deeper learning about STEM topics (as

compared to a usual comprehensive curriculum, Indicator 1), the amount of time a Grade 4 class spends on science learning each year (Indicator 2). Thus, the ED*Fact*s data elements built into most educational data systems are valuable components in determining whether data exist to address an Indicator, but are insufficient in and of themselves for addressing questions about educational access and process.

Emergence of the *Data Element* for Communication and Collaboration

A data element consists of the name and description of a piece of data (fact or statistic). Because a data element is information for classifying or organizing data, it is often called *metadata*. Each SLDS data system has some form of *data dictionary* (i.e., a list of the kinds of data tracked by the system). Sometimes the dictionary is implicit in the design of the state database. In these older systems, there is the challenge of inclusion: what data are included? Each field is a way of reporting data and only those data that have a named field can be entered into the database. In these systems, without an explicit data dictionary, one decides where to put the data based on what may be idiosyncratic interpretation of the name of the field. For example, absent any additional information, which of the following data are appropriate for a field named *Minutes of Science Per Class*: The average number of minutes/day spent in science instruction by one teacher? Minutes per week? Or does it refer to a school-level value of accumulated time spent by all teachers in the school?

The last decade has seen several federally funded efforts to support development of SLDSs and to create a shared dictionary. The data elements (names and definitions) for a few elements in the U.S. Department of Education's *Common Education Data Standards* (CEDS) handbook, Version 6, are shown in Table 1. The CEDS data element collection was developed over the last decade, contributed to by the vendors and SLDS experts in many states.

Table 1. Examples of Data Elements from CEDS V.6

Element Name	Element Description/Definition
Course Title	The descriptive name given to a course of study offered in a school or other institution or organization. In departmentalized classes at the elementary, secondary, and postsecondary levels (and for staff development activities), this refers to the name by which a course is identified (e.g., American History, English III). For elementary and other non-departmentalized classes, it refers to any portion of the instruction for which a grade or report is assigned (e.g., reading, composition, spelling, and language arts).
Instructional Minutes	The total number of instruction minutes in a given session, as determined by time in class, time on task (e.g., engaged in a class), or as estimated by a qualified course designer.
Activity Title	The title for a particular activity, such as a co-curricular or extra-curricular activity.
Activity Identifier	A unique number or alphanumeric code used in the local system to identify an activity, such as a co-curricular or extra-curricular activity that is offered at an education institution.

Activity Time Involved	The amount of time the student participated in the events and procedures of an activity, such as a co-curricular or extra-curricular activity that is offered at an education institution.
Teacher Education Credential Exam Type	The type of examination used to assess teacher candidate's knowledge and skills. *Option Set:* PraxisI; PraxisII; ACTFL; StateExam; Other.
Years of Prior Teaching Experience	The total number of years that a person has previously held a teaching position in one or more education institutions.

The specific objectives of the work reported here:
(1) Identify data elements currently in use by SLDSs.
(2) Categorize existing data elements according to their association with the data needed to monitor progress in the focal indicators (1, 2, 3, 6, 7, 8).
(3) Identify alignments and possible extensions to the national Common Education Data Standards (CEDS) data elements as a foundation for understanding changes needed in SLDS data elements.

Methods and Data Sources

Informed by the latest research on educational structures, such as what makes something a "STEM program," the study sought information on SLDS content and processes. We combined qualitative (document analysis and interview) with quantitative (descriptive statistics) methods. Project STEM experts identified classes of data needed to address focal indicators then examined the national Common Education Data Standards (CEDS) for such data elements. We obtained and analyzed 15 states' data dictionaries and categorized state elements by indicator. Additionally, we talked with SLDS and STEM program staff in six states (a total of 10 individual and group interviews) about current use and future viability of SLDSs as an indicator-supportive repository.

Results

The presence of indicator-aligned data elements in CEDS and sample states varied widely (particularly for Indicators 2 and 3). Combining data elements might lead to useful composites for tracking progress in an Indicators. SLDS leaders and STEM directors demonstrated eagerness to talk with colleagues, across professional subcultures, to see a big picture that "supports children in learning – that's what we are all concerned with, really." Some states had SLDS managers who were "stars" in the development of the CEDS dictionary, some were interested parties whose data management vendors were involved with CEDS development, and some SLDS managers noted they had kept as far away from CEDS as they could ("it was a make-work project and you got involved to look good for a little while, we didn't have the time for that").

After a quick look at an example, Tables 2, 3, and 4 below summarize counts of data elements that could provide information about each indicator. To create the counts, we analyzed the content of each data element (name and description/definition) and noted the relevant category for the indicator.

An Example – Behind the Numbers in Table 2

Row 1 in Table 2 summarizes counts from our review of the Arkansas SLDS data dictionary. It included 4 data elements directly related to Indicator 1 – identify existence of, or enrollment in, STEM schools and/or programs:

Arkansas – Indicator 1 Data Elements

1.1. "CL0052 course_desc" – The name given to the course from the Master Schedule.
1.2. "CL0055 digital_learn" – If a class is conducted primarily through the utilization of digital learning technology, indicate the type of technology used from the Option Set: CV = Compressed/ Interactive Video, WB = Web Based, NA = Not Applicable
1.3. "SC0350 accred_ncentral" – Indicates whether the school participates in the AdvancedED/North Central Program.
1.4. "SC0360 magnet" – Elementary or secondary school program within the public education system that offers exemplary learning environments or courses that are not available within the regular school curriculum. The programs are designed to attract voluntary students from all parts of a school district without reference to the usual attendance zone rules.

In addition, we found in Arkansas' data dictionary: 3 data elements about teacher STEM preparation (Indicator 6), 2 data elements providing detail on teacher professional development in STEM (Indicator 7) and 1 element related to each of the other indicators. The counts are a proxy measure that provide a glimpse into the potential of existing SLDS designs to address the call for data-driven decision-making in each Indicator area.

Table 2. Summary Count of Data Elements by State and for CEDS for Each Indicator

Data System	Indicator 1	Indicator 2	Indicator 3	Indicator 6	Indicator 7	Indicator 8	Total
AR	4	1	1	3	2	1	12
AZ	4	1	1	3	2	1	12
CA			2	1	1		4
CO	1		2	2	1		6
GA	8			3	4	3	18
KS	12		3	3			18
KY	3			1			4
NC	10			3	2	2	17
NH	1			7			8
NV	17		1	1			19
OR	8						8
RI	4	2	1	2	1		10
VA	19		1	1			21
VT	4	1	1	3	1	1	11
WA	17		1	1			19
CEDS	25	3	3	16	25	22	94

It is worth noting here that a larger count is not necessarily "better" in terms of usefulness for examining progress in an Indicator. Questions about the richness of the data element as a source and quality of data stored in that data field were beyond the scope of this study.

Tables 3 and 4 (next page) slice the data in a different way, giving categories of data elements and the counts across all states (Table 3) and CEDS (Table 4) of elements in those categories – this gives a sense of the variability of the data pool across topics, at a smaller grain size than Indicator. The row "(blanks)" in Table 3 provides a count of the number of states for which NO data element existed for an Indicator. Table 4 provides parallel information for the national data dictionary proposed by CEDS.

Table 3. Summary Count: Data Elements by Category across **States** for Each Indicator

Date Element Category	Ind. 1 STEM School/ Program	Ind. 2 Time in K-5 on Science	Ind. 3 Access to K-5 Science	Ind. 6 Teacher Know-ledge	Ind. 7 Teacher STEM PD	Ind. 8 Admin. STEM PD
Course Description	37					
Digital Learning	6					
School/Program Detail	62					
STEM Performance	5		1			
SPED Content detail			1			
SPED Time detail			1			
Extra-curric. program detail			1			
K-5 Class Time on Science		4				
K-5 Course Description			7			
K-5 Content Description	1		4			
Professional Certificate - Content detail	1			11		
Professional Certificate - Outcome				9		
Professional Content Expertise/Degree				5	1	
Professional Degree Level				1		
Professional Experience				5		
PD Content detail					3	1
PD Outcome detail					1	
PD Process detail					6	5
PD Time detail				1	1	1
(blank)	1	12	5	1	8	11
Column Totals	113	16	20	33	20	18

Table 4. Summary Count of Data Elements by Category across **CEDS** for Each Indicator

Date Element Category	Ind. 1 STEM School/ Program	Ind. 2 Time in K-5 on Science	Ind. 3 Access to K-5 Science	Ind. 6 Teacher Knowledge	Ind. 7 Teacher STEM PD	Ind. 8 Admin. STEM PD	Total
Course Title/Description	10						10
Digital Learning	1						1
School/Program Detail	9		1	1			11
STEM Performance	3						3
Extra-curr. program detail	1	1					2
K-5 Classtime on Science	1	2					3
K-5 Content Description			2				2
Professional Certificate - Content detail (e.g., math)				4			4
Professional Certificate – Outcome (e.g., HQ)				3			3
Professional Certificate – Process (e.g., degree)				6			6
Professional Degree Level				1			1
Professional Experience				1			1
PD Content detail					2	2	4
PD Outcome detail					6	4	10
PD Process detail					17	16	33
Column Totals	25	3	3	16	25	22	94

Review of Results by Indicator

Indicator 1: Number of, and enrollment in, different types of STEM schools and programs
The largest set of data elements across SLDS collections was under Indicator 1. Related recent NSF-EAGER-funded work, by Century, LaForce, and Noble (2015), focused on specifying what "STEM schools and programs" means at the secondary level. Developed with regional/district educational experts but without consulting SLDS or national/CEDS data systems, Century and colleagues have offered a taxonomy for determining whether a school or program might be called a "STEM school or program" (see Figure 1, next page).

Using their carefully developed and regionally vetted criteria, we find that data elements exist in CEDS for making the determination of whether a school is a STEM school. However, data elements in SLDS coffers are not *sufficient* for making the determination – additional information would be required. Of the 37 *course description* data elements identified across the sampled states, 32 of them are specific to high school advanced placement or International Baccalaureate courses. The other 5 include descriptors for middle school STEM courses. None identify STEM content in elementary schools (which do not have "courses" in the way that secondary schools do).

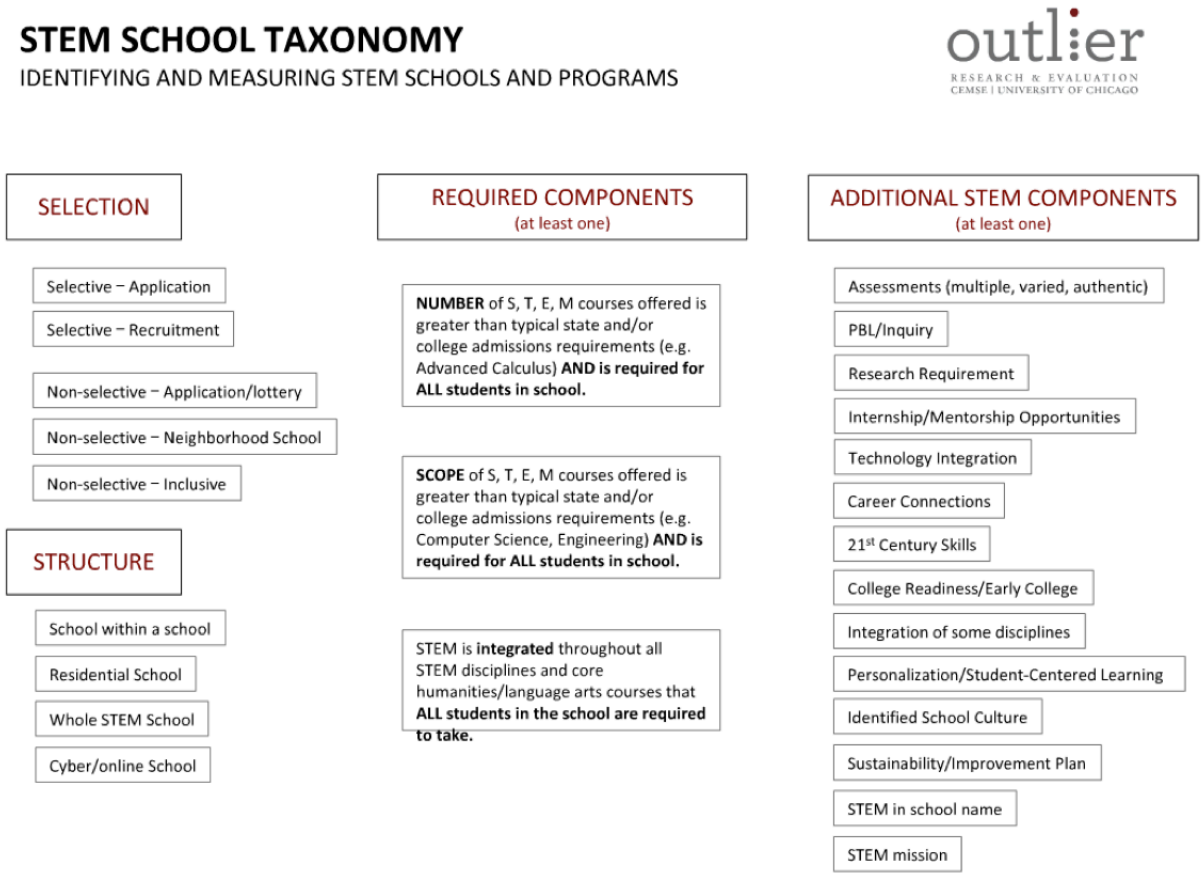

Figure 1. "Types of STEM schools and programs" for Indicator 1 (Century & Noble, 2015)

The CEDS elements could provide data that would allow determining the number and nature of STEM schools/programs **at the secondary level** in a region, state, or the nation. As an example, the 10 data elements in CEDS related to identifying course content in ways recommended by recent work by Century et al. (2015) for Indicator 1 are:

• K12 -> K12 Course
• K12 -> Course Section -> Directory
• K12 -> K12 Course -> High School Course Requirement
• K12 -> K12 Course -> Course Title
• K12 -> K12 Course -> Course Description
• K12 -> K12 Course -> Course Department Name
• K12 -> K12 Course -> Course Aligned with Standards
• K12 -> K12 Course -> Core Academic Course
• Career and Technical -> Course Section -> Session Type
• Career and Technical -> Course -> Course Title

Indicator 2 and Indicator 3: Grades K-5 time on and access to science learning
We remarked in the original proposal that information on these indicators might be hard to come by – federal reporting to date has not required description of elementary grades content related to science. The focus of past data gathering and reporting at the state level has been reading (curriculum, instruction, and assessment), with some attention to mathematics (assessment). Consequently, there were no data elements found in national or state data dictionaries that addressed Indicator 2: Time allocated to science in K-5 instruction. Some data elements did exist that address Indicator 3: Science learning opportunities in elementary schools, most were related to special programs (e.g., for children with special needs or individual education plans, Title 1 programs).

Indicator 6: Teachers' science and mathematics content knowledge for teaching
All data dictionaries had a (federally mandated) record of the type of certification a teacher had. Additionally, in most states, there were one or more elements with detail that coded content specialization for "highly qualified" status (e.g., highly qualified in mathematics or biology). However, data elements for the processes or interim experiences (e.g., professional development) of teachers that contributed to the highly qualified designation were absent from most SLDS collections.

Indicator 7: Teachers' participation in STEM-specific professional development activities
It appeared that sharing of data between the office of teacher certification in each state and the SLDS was limited to the federal-reporting compliance categories. What the educational SLDSs kept track of regarding professional learning content, processes, and outcomes varied. Some states used content, outcome, or process metrics and one used a metric of time (e.g., time spent in PD) as a proxy for the quality or content of professional learning. On a related note, indication of the qualification of staff (teachers, assistants, administrators) to provide STEM-rich experiences is absent from Century and colleagues' (2015) taxonomy of what is necessary to identify a STEM school or program.

Indicator 8: Instructional leaders' participation in PD on creating conditions that support STEM learning.
Just as "types of STEM schools" needed a taxonomy for Indicator 1, "conditions that support STEM learning" needs description in order to identify potential data categories for Indicator 8. The categories identified by the state representatives and our research group helped clarify what was, and was not, available in CEDS or in the SLDS dictionaries. After examining the literature and discussing various existing frameworks for describing the work of instructional leaders, a draft taxonomy informing analysis for Indicator 8 was based on Cohen et al.'s (2009) four dimensions of school climate that inform an instructional leader's work:

- Safety - ensuring physical and social-emotional well-being;
- Teaching & Learning - supporting the quality of instruction, professional development, leadership, and ethical learning;
- Relationships - safeguarding respect for diversity, school community and collaboration, morale and connectedness
- Environment & Structure - orchestrating adequate space and resources, aesthetic quality, curricular and extra-curricular offerings.

With these in mind, we examined CEDS and identified a set of elements for Indicator 8 for the aspects of Environment & Structure – this was the area in which most CEDS/SLDS data elements appeared to fall. We would have to invent new data elements for things not currently measured in CEDS or SLDSs in order to address the other three components in any substantive way. Further development, perhaps in the direction of a taxonomy like that created by Century and colleagues for Indicator 1, is a clear next step.

Conversations with State-Level Staff

We conducted ten individual and focus group interviews with SLDS staff and state-level STEM staff (e.g., state director of Career & Technical Education, state director of STEM teacher professional development) in six states. These discussions indicated that local SLDS cultures differed in their communication structures and degree of cross-talk with STEM-specific state-level staff. We learned from the interviews that for most there was a high staff churn rate: turnover in staffing on an annual, quarterly, even monthly basis. At the same time, senior management was fairly stable. The type of flux in staffing meant that many offices relied on senior management for institutional memory with little opportunity for continuity in development issues. We asked SLDS staff about CEDS. Responses ranged from ignorance to dismissal to passionate support. "STEM indicators" was news to all but one interview group.

Given these circumstances, with the exception of one western state, the SLDS offices had in common a highly compliance-based approach, with little room for attention to improvement-based efforts in data curating. As a result, SLDS data generally did not systematically include process details that might be available from other state-level offices. As an example, consider Indicators 6, 7, and 8 – about the professional preparation and development of teachers and school leaders. The state office of teacher certification may have a great deal of data on who has what credentials and the details of the courses or professional development activities engaged in by teachers or principals (e.g., when a state has a 5-year cycle of teacher certification renewal through course and PD credits) but those detailed data are not part of the SLDS stewardship for compliance-based federal reporting. Instead, the SLDS data elements most commonly found were summative outcomes: type of certification of teacher (e.g., highly qualified or not) and number of years of teaching experience (e.g., as listed in Table 3).

Two state interviews included comments from staff on "glaring holes in the practice of gathering and paying attention to data from the classroom and district level." With state-level STEM professional staff noting: "There is a culture of top-down, tell me what to do, in schools and the principal plays a major role." SLDS and STEM staff in each state echoed the concerns of one state level STEM staffer who asked several practical questions: "We are already dealing with data and people are already building local data dashboards – what more do we need? What is the payoff, for principals and districts, of doing this extra data collection needed for this indicator thing? Who would be in charge of entering that data? Cleaning the data?"

Scholarly (and Practical) Significance

As noted above, the SLDSs and CEDS have the most data elements that pertain to Indicator 1, some for Indicators 6 and 7, a few for Indicator 8, and the fewest for Indicators 2 and 3. It may be possible to combine some data elements to address the indicators, but certainly this process is not straightforward and depends on the specific state's definition of particular data elements.

In fact, because of the differences across states, it is difficult to make worthwhile comparisons. We went into this study with a driving question: To what extent might the SLDSs serve as a repository of data that could be used to measure progress in the STEM indicators? An ancillary question arose as the project progressed: To what extent do any standardized, national data repositories address the STEM indicators? The federal data store through ED*Facts* is limited and the CEDS database has not (yet) gained the state and local attention and "buy-in" needed.

Going back the driving question. Clearly, states differ in the extent to which the current data dictionaries address even a small subset of the indicators and the degree to which professional preparation data sets are included in the SLDS. At the local level, it is unclear the extent to which district data might address the STEM indicators more effectively than what is currently in place at the state or federal level.

Certainly, if all states aligned their data dictionaries to the CEDS dictionary, a national picture, if sparse, would be possible. The detail would, necessarily, be meager because at this time in K-5 science there do not appear to be any state-level mandates like those in reading and mathematics. No professional structures (much less data elements) exist that allow local or state-level staff to gauge where and how there are time and opportunity to learn science, in each grade (much less how this is situated in the context of the other goals for each grade).

National policy documents like the *Next Generation Science Standards* (2013) can inform *what* is taught, but building a culture of evidence-based decision-making about progress regarding who is learning and under what conditions, is nascent. For example, there is no articulation between elementary, middle, and high school in science (e.g., placement processes for entering middle school or high school that are like the teacher- and data-driven placement decisions in mathematics). Thus, the future of the effort is intimately linked to Indicator 9: Inclusion of science in federal and state accountability systems.

Parallel to the question of whether the information in the data systems is sufficient for measuring progress is the question: How do we define "progress"? In particular, where are learner access to science, teacher knowledge, and instructional leaders' professional growth in the *accountability* systems? Data to support examination of progress do not have to come from written tests. Good proxy measures for professional growth can include feedback mechanisms (e.g., an audit trail for teacher development of an online professional portfolio; whether or not an online professional course for leaders has *Quality Matters* certification or not; and other characteristics of professional learning processes, in addition to outcomes).

Two national policy agendas intersected for the work reported here: CEDS and STEM indicators. SLDS and state STEM program staff can contribute complex, thoughtful, and rich ideas. Carefully orchestrated conversations that clearly value those contributions can be fruitful (they were for us!) in pinpointing requisite data elements and framing motivation for data collection. We also saw how SLDSs can function both as an essential tool for ensuring equity and excellence through compliance and as a means for supporting educational equity and excellence through the quality of the STEM indicator data they curate.

Acknowledgment

This project is supported by a grant from the National Science Foundation (NSF), DGE-1445522. Any opinions, findings, conclusions or recommendations are those of the authors and do not necessarily reflect the views of the NSF.

References

Century, J., LaForce, M., & Noble, L. (2015, October). *Finding Common Ground: A Taxonomy for Identifying and Describing STEM Schools.* MSPnet Academy Webinar <http://hub.mspnet.org/index.cfm/stemschools>.

Cohen, J., McCabe, L., Michelli, N. M., & Pickeral, T. (2009). School climate: Research, policy, practice, and teacher education. *The Teachers College Record*, *111*(1), 180-213.

Common Education Data Standards (2016, Version 6). Accessed online through the ceds.ed.gov website: <https://ceds.ed.gov/whatisCEDS.aspx>

National Research Council (2013). *Monitoring Progress Toward Successful K-12 STEM Education*. Washington DC: National Academies Press.

Next Generation Science Standards Lead States. (2013). *Next generation science standards: For states, by states*. National Academies Press. Retrieved from <https://www.nap.edu/read/18290/chapter/1>.