Studies in Educational Evaluation 40 (2014) 18–35

Contents lists available at ScienceDirect

Studies in Educational Evaluation

journal homepage: www.elsevier.com/stueduc

A Common Measurement System for K-12 STEM education: Adopting an educational evaluation methodology that elevates theoretical foundations and systems thinking

Emily Saxton [a,*], Robin Burns [b], Susan Holveck [c], Sybil Kelley [d], Daniel Prince [e], Nicole Rigelman [d], Ellen A. Skinner [f]

[a] Center for Science Education, Portland State University, 1136 SW Montgomery Street, Portland, OR 97201, United States
[b] Intel Corporation, United States
[c] Beaverton School District, 16550 SW Merlo Road, Beaverton, OR 97006, United States
[d] Graduate School of Education, Portland State University, 1136 SW Montgomery Street, Portland, OR 97201, United States
[e] Outdoor School Program, Multnomah Education Service District, United States
[f] Department of Psychology, Portland State University, 1136 SW Montgomery Street, Portland, OR 97201, United States

ARTICLE INFO

Article history:
Received 22 June 2013
Received in revised form 14 October 2013
Accepted 19 November 2013
Available online 21 December 2013

Keywords:
Common measurement system
STEM education
K-12 education
Formative evaluation
College and career readiness

ABSTRACT

Science, Technology, Engineering, and Mathematics (STEM) education is important on a national, regional, local, and individual level. However, there are many diverse problems facing STEM education in the US, one of the most critical is the limitation of current measurement tools and evaluation methodologies. The development of a common measurement system is an important step in addressing these problems. This paper describes the conceptualization stage of the development of a common measurement system. The resulting STEM Common Measurement System includes constructs that span from student learning to teacher practice to professional development to school-level variables. The authors detail the constructs and measurement tools associated with each construct. The interconnections within the STEM Common Measurement System are also discussed.

Introduction

Significance of STEM education

Science, Technology, Engineering, and Mathematics (STEM) education is important on a national, regional, local, and individual level for several reasons. First, multiple stakeholders including government and businesses acknowledge that investments in STEM education are key to the United States' global economic competitiveness (Department of Commerce, 2012; National Academy of Sciences, 2010). To keep that competitive edge, a greater number of students pursuing STEM careers are needed to ensure a STEM capable work force (Carter, 2007). Second, increasing the quality of STEM education is viewed as an important avenue toward creating an informed citizenry that will benefit policy decisions at the national, regional, and local levels (National Research Council (NRC), 2011). Understanding scientific discoveries and methods is relevant for citizenship in a democracy that is dependent on scientific products and technologies (Bencze & Di Giuseppe, 2006). Third, the economic benefits of STEM education to individual students in the form of higher income and greater job security have also been clearly established (Carnevale, Strohl, & Melton, 2011; Zaback, Carlson, & Crellin, 2012). For example, in their analysis of the ten college majors associated with highest median income eight out of the ten majors were engineering majors with the remaining two majors also in STEM fields (Carnevale et al., 2011).

Given the above-described significance, a new emphasis on STEM education as using a coherent, integrated, and multi-disciplinary approach has gained prominence. This new approach to STEM education reflects the ways that STEM concepts and higher-order thinking skills are actually applied in the real world by scientists, engineers, and other professionals in order to recognize, evaluate, and solve complex problems and discover and advance new knowledge (Lewis, 2006; Newmann, Marks, & Gamoran, 1996). Stakeholders from diverse backgrounds recognize the value of high quality STEM education for delivering not only STEM content knowledge, but also the 21st century skills that

* Corresponding author. Tel.: +1 503 725 3190.
E-mail addresses: esaxton@pdx.edu (E. Saxton), robin.burns@intel.com (R. Burns), Susan_Holveck@beaverton.k12.or.us (S. Holveck), sybilkel@pdx.edu (S. Kelley), dprince@mesd.k12.or.us (D. Prince), rigelman@pdx.edu (N. Rigelman), skinnere@pdx.edu (E.A. Skinner).

0191-491X/$ – see front matter
http://dx.doi.org/10.1016/j.stueduc.2013.11.005

all students will need to engage as successful members in our workforce and society. Despite wide-spread recognition of the importance of exemplary STEM education, the United States (US) educational system has found it difficult to deliver on this promise, therefore, change is needed to an education system that is fraught with many problems.

Evidence of the shortfalls of the current US STEM education system

The importance of and increased emphasis on STEM education is not a trend unique to the US (Chubb & Kusa, 2012), but for the purposes of this paper the US education system is critically examined as a case and as the context in which this study took place. There are many problems facing STEM education in the US. In considering the most critical problems facing STEM education, the diversity of these problems is striking. These diverse problems span from student underperformance in STEM disciplines to shortfalls in teacher practice to limitations of current measurement systems.

Student underperformance in STEM

Both national and international assessments of student mathematics and science performance reveal shortfalls in STEM education. Only 39% of 4th grade students and 34% of 8th grade students score as proficient or advanced in mathematics on the National Assessment of Educational Progress (NAEP) (National Center for Educational Statistics, 2012a). In science, only 31% of 8th grade students score as proficient or advanced on the NAEP (National Center for Educational Statistics, 2012b). The Trends in International Mathematics and Science Study (TIMSS) provides further evidence of poor student performance in STEM. These findings are especially relevant when the percentage of students meeting TIMSS *advanced* benchmarks are considered, given the close association between the high level of content and skills described in the advanced benchmarks and college readiness indicators (Conley, 2007a, 2007b; National Center for Educational Statistics, 2013). Nationally, only 13% of 4th grade students and 7% of 8th grade students score at advanced benchmarks on the mathematics version of TIMSS (National Center for Educational Statistics, 2013). Similarly, only 15% of 4th grade students and 10% of 8th grade students score at advanced science benchmarks (National Center for Educational Statistics, 2013). Taken together, recent NAEP and TIMSS scores provide clear evidence that US students' are underperforming in STEM.

Shortfalls in teacher practice

There is an increasing awareness that current K-12 teaching practices tend to isolate STEM disciplines, emphasize rote memorization of STEM content, and neglect higher-order thinking skills (NRC, 2001; Niemi, Baker, & Sylvester, 2007). The purpose of this article is not to review the current findings on teaching practices in STEM; however, a few examples from the literature are provided to illustrate the need for improvement of teacher practice in STEM. In their study of elementary mathematics instruction, Rowan, Harrison, and Hayes (2004) note the repetition of content across grade levels, emphasis on breath of curriculum coverage over depth of coverage, and inconsistency of both "content coverage and teaching practice among teachers within the same school, even when these teachers work at the same grade level" (p. 121). Findings from research on science education convey similar trends. Gallagher (2000) notes a predominance of teaching for lower-order skills and a failure to target deep conceptual understanding. Typical science instructional practice at the elementary level is depicted as "...teacher centered, textbook driven, and geared toward achieving lower level knowledge and comprehension objectives... [with] at best, few student-centered, inquiry-based activities" (Schwartz, Abd-El-Khalick, & Lederman, 2000, p. 191).

Moreover, national accountability policies further compound the shortfalls in STEM teaching practices. In a review of research literature on test-based accountability policies in science education, Anderson (2011) finds a conflict between research grounded, reform-based efforts and accountability policies based on standardized testing. In this research review, the conclusion from teacher and administrator perspectives is clear: "accountability limits time and effort spent on science, drives the remaining science instruction toward memorization of facts, and constrains student learning" (Anderson, 2011, p. 121). Given that mathematics is also a subject associated with high stakes testing, it is reasonable to infer that the conflict between reform efforts and accountability policies is relevant to K-12 mathematics reform efforts as well.

Limitations of measurement systems

Current measurement systems in STEM education are associated with multiple limitations that hinder efforts to improve outcomes for students and teachers and thus are a contributing factor to the persistence of the problems described above. Generating and sharing information about how to improve STEM teaching and learning depends on the quality of measurement systems, therefore, addressing these limitations is key to transforming STEM education. Measurement of key constructs in STEM education are difficult or problematic at nearly every level of the education system.

The current capacity for the measurement of key student outcomes for college and career readiness in STEM education is associated with significant limitations. Many scholars advocate for changes to the format, goals, and frequency of assessment of student learning (Anderson, 2011; Black & William, 1998, 2009; Gallagher, 2000; Yeh, 2006). Similarly, Ball and Hill (2009) characterize the measurement of teacher practice, in general, as under-developed. Further, the measurement of teacher practice in STEM education specifically is an area that is relatively new, and as a result, associated with few measurement tools that work across the STEM disciplines. Pianta and Hamre (2009) emphasize the need for better teacher practice measurement instruments, "...we need more evidence on why and how classrooms, and teachers, matter; the need for evidence is not trivial..." (p. 110). The need for teacher practice measurement instruments that are effective, feasible, valid, and allow for the investigation of links between specific teaching strategies and student learning outcomes is also a need that is common across STEM disciplines (Ball & Hill, 2009; Hiebert & Grouws, 2007; Lewis, 2005).

Moving beyond assessment of individual students and teachers, measurement strategies for programs and systems are also lacking and underdeveloped. There is a clear need for appropriate measurement of teacher professional development programming. Crow (2011) stressed "using [teacher practice] data for professional learning at the individual level has not yet been systematic or widespread" (p. 26). Finally, the limitations of measurement instruments extend into organizational variables that are of relevance to the quality of STEM education. Matsumura, Garnier, Pascal, and Valdes (2002) demonstrate the extension of measurement needs to the school level "New indicators that help schools, districts, and states monitor and support efforts to improve the quality of instruction are clearly needed. These indicators are important for providing feedback to schools and districts about their interim progress towards reform goals" (p. 208). Taken together, these sources demonstrate the limitations of current measurement tools across many of the constructs that are most relevant to STEM education.