Measuring STEM Learning in After-School Summer Programs: Review of the Literature

Erica Mallett Moore

University of Washington
emall@uw.edu

Ari Hock

University of Washington
arihock@uw.edu

Bronwyn Bevan

Wallace Foundation
bronwynbevanresearch@gmail.com

Katie Headrick Taylor

University of Washington
kht126@uw.edu

Abstract

After-school and summer programs provide important opportunities for youth to learn STEM practices and form STEM-related identities. However, there has been limited coordination across these programs to measure effectiveness toward learning outcomes. To better understand the constructs that are used to evaluate these programs, we searched key terms related to out-of-school time STEM learning on several education research databases. Our search uncovered 36 different tools. Across these tools, we discovered 76 measures, which were then grouped into 10 constructs based on similar themes. Constructs included: attitude toward science, career awareness and career interest, curiosity, engagement, home/school environment, interest, motivation, nature of science, self-efficacy, and STEM practices. Each construct is defined and clarified with examples from the tools. The review also considers tensions between attempts to standardize measures for evaluating program success and the need to account for equitable STEM learning pathways and adaptability across diverse communities.

Key words: *out-of-school time, program evaluation, STEM learning, literature review*

Introduction

A recent synthesis of research on human learning demonstrates that people learn through complex, dynamic cultural processes that develop over different timescales and across settings (National Academies of Sciences, Engineering, and Medicine, 2018). The time spent outside of the school classroom is often pivotal in developing interests, identities, and capacities to engage with ideas and subjects students may also encounter in school, including STEM (Banks, 2006; National Research Council [NRC], 2015).

Indeed, several studies have demonstrated the importance of family members, role models, and out-of-school time experiences in developing students' awareness and commitment to pursuing particular STEM learning pathways (Bell et al., 2019; Falk et al., 2017; Halim et al., 2018; NRC, 2015). Out-of-school time (OST) experiences may include structured after-school, weekend, and summertime programs; visits to designed informal learning environments such as nature centers, museums, and libraries; as well as everyday experiences, conversations, and observations that make up much of a young person's daily life. Research suggests these out-of-school time STEM experiences produce the following outcomes:

- **Dispositions** related to choosing to do STEM (e.g., attitude, interest, curiosity, motivation, identity, and self-efficacy),
- **Disciplinary capacity** to productively engage in STEM (e.g., understanding the nature of STEM fields, skills, and concepts),
- **Social capital** in STEM (e.g., role models, mentors, and peer networks), and
- **Commitment** to and pursuit of STEM learning pathways (e.g., career awareness, STEM course and program selection).

This article is focused on two lines of argument in the literature that pertain to the value and potential of after-school and summer STEM programs as they relate to student learning outcomes. The first focuses on issues of equity in STEM learning and the second pertains to ecological perspectives on STEM learning. We examine both of these in turn, to ground a discussion about program goals and how they go about measuring their performance toward each.

After-School and Summer STEM Programs and Equity

For students from communities historically excluded from STEM fields, leveraging OST can be pivotal for advancing more equitable outcomes by improving who has access to and contributes to STEM learning activities and pathways (Calabrese Barton & Tan, 2010; Jenkins et al., 2016;

Nasir et al., 2006). In conceptualizing the role of OST STEM in supporting equity, we draw on Philip & Azevedo (2017), who argue that there are at least four distinct ways that equity in Informal STEM Education (ISE) is commonly characterized within research frameworks, and that each has different implications for program design, implementation, and evaluation:

1. Informal STEM learning can go beyond supporting student achievement in in-school STEM, by creating opportunities for more hands-on and expansive learning.

2. Informal STEM learning can deepen students' interest, identity, and excitement in STEM, which can then be further influenced and developed by school STEM programs.

3. Informal STEM learning can expand students' perceptions of what constitutes STEM, and its relevance for their everyday lives. It may demonstrate how STEM practices are already a part of their own practices as well as the practices of their families and their communities. As such, informal STEM learning can help to break down longstanding cultural barriers of "who does STEM," building a sense of belonging and even ownership in these fields.

4. Informal STEM learning can be positioned as a tool for students concerned with broader issues of social justice and community development, thus supporting personal agency among broader communities, and connecting STEM to broader social purposes.

There may be other related benefits, as well. For example, sustained informal STEM engagement may counteract negative school STEM experiences for underserved student groups. Furthermore, engagement in informal settings may serve as an introduction to STEM for students in schools and districts where STEM opportunities are limited.

Some students may be exposed to or express interest in STEM and seek and benefit from academic enrichment opportunities. Other students who identify with social justice or community development may also identify with STEM while using it for broader community purposes. For example, students concerned about environmental contamination near their school in the Bronx learned to code in order to program small robotic toy dogs to find danger spots (see Jeremijenko's Feral Robotic Dog project in DiSalvo, 2012).

Programs may emphasize one or more of these characterizations of equity over others. However, each one can make strong contributions to students' relationship with STEM. They are not mutually exclusive and can instead be complementary. The call from Philip & Azevedo (2017) is not to prioritize one approach over the other, but to understand why students might

opt into or thrive with a given approach, and to develop coherence across program design, implementation, and evaluation.

After-School and Summer Programs in the STEM Learning Ecosystem

Because students' interests in and commitments to STEM fields may fluctuate and change over time (Azevedo, 2015; Barron, 2006), many have argued for the importance of providing multiple, diverse, and repeated opportunities for students to engage in STEM, and for actively brokering awareness and inclusion in these opportunities (Barron & Bell, 2015; Bevan, 2016).

STEM learning ecosystems are made up of the range of STEM-related activities, places, people, and cultural practices that constitute a given community (whether geographical or virtual). Some STEM learning ecosystems—for example, in rural settings—may have access to natural and cultural resources but have less access to institutional opportunities to engage with STEM. Other STEM learning ecosystems—for example, many urban settings—may have a wide range of institutional, cultural, and social STEM learning resources, yet may also be imbued with unique sociocultural histories of exclusion from STEM that effectively communicate to students in those communities who "belongs in STEM" and who does not (Bevan et al., 2018).

Enriching the STEM Learning Ecosystem means providing a range of inclusive and inspiring STEM learning options for students so that—wherever they are in their own personal developmental trajectory and whatever the sociocultural histories that they contend with as to who belongs in STEM or not—they can access programs that match their interests and needs at any moment in their development. This includes opportunities for students committed to social justice to experience academic STEM programs that refine their skills and reflect their interests, as well as for students who participate only in academic STEM to branch out and apply those skills in programs focused on community development (Collins & Bilge, 2016; Collins, 2018). In addition, paying attention to the historical dimensions of STEM learning ecosystems, including the field's marginalization of people of color and women, is essential in ensuring diverse participation in STEM activities.

OST STEM programs typically evaluate the degree to which they are achieving particular learning outcomes through the use of quantitative and qualitative instruments. These instruments are either used by program staff or external evaluators to collect relevant data, assessing the degree and quality of STEM learning taking place in the programs. Evaluation instruments are powerful tools that can both describe and (in)visiblize learning processes,

reifying the kinds of learning that are recognized and valued in educational spaces (Engeström, 1999). The degree to which these instruments acknowledge and respond to important sociocultural and historical dimensions of STEM learning will affect the degree to which the field moves toward a more equitable learning ecosystem (Garibay & Teasdale, 2019).

Conducting the Literature Review

In the sections that follow we outline the range of constructs measured in current tools in use in OST STEM programs. Through our analysis, we aim to describe the state of program evaluation, and advocate for evaluators and researchers to consider measures that support community and educational justice (Bang & Vossoughi, 2016). We draw attention to ecological perspectives as well as the analysis of equity by Philip and Azevedo to ensure that an examination of the OST STEM measurement tools can foster discussion about what after-school and summer STEM programs aim to accomplish, how they do so, and how they document progress in meeting their goals for youth in STEM.

We conclude that the instruments reviewed may encourage two assumptions: (a) that individuals will change based on appropriate programming or instruction and (b) that improvement on these scales represents some type of success. However, the first assumption neglects systemic factors that contribute to inequitable outcomes, and it is unclear in the second assumption whether the kind of success being measured is the kind of success we should be measuring. (See Conclusion for more discussion.)

Methods

In order to gather a broad collection of OST STEM measurement tools, disaggregated by age and measures we: (a) conducted computer searches on OST STEM databases (Afterschool Impacts Database, Afterschool Matters, California State University, Northridge, Click 2 Science, International Journal of Science Education, The PEAR Institute: Partnerships in Education and Resilience [PEAR] at McLean Hospital and Harvard Medical School, and STEM Ready America), and (b) examined reference lists of subsequent tools, research papers, and websites.

We reviewed these tools over a 3-week period. The following search terms yielded the results presented here: After-School, After-School STEM, After-School Science, Science After-School, OST STEM, OST STEM programs, OST measurement, and OST tools. Tools are categorized by measures and the measures are presented here in alphabetical order. We began our search with the PEAR Institute search engine known as Assessment Tools in Informal STEM (ATIS) and

chose tools that referenced OST STEM. We went on to reference tools that were designed for schools in an effort to include a comprehensive list of STEM tools, many of which mention being adaptable for OST STEM (see Appendix A for a complete list of tools).

Measuring With the Likert Scale

The majority (27) of the 36 tools we identified used student self-reports on Likert scales. The Likert scale was developed to measure attitude reliably and validly, which involves cognitive, affective, and psychomotor components (Joshi et al., 2015). These components make the scale well-conditioned to measure constructs in after-school STEM programs. However, there are some important limitations related to cross-cultural variation (cf. Briggs et al., 2019; Flaskerud, 1988; Lee et al., 2002). Moreover, responses to the scale may actually be misleading. A study with people experiencing homelessness (Ogden & Lo, 2012) found significant contradictions between their answers on a Likert scale and free text responses to questions about their quality of life. The study notes that "when answering questions, different populations may implicitly use very different frames of reference with the focus of the question being interpreted within the context of a different aspect of their lives (Ogden & Lo, 2012, p. 359)." For this reason, some statisticians object to the validity of averaging Likert scales to determine the midpoints within a data corpus (Subedi, 2016). It is essential to keep in mind the importance of context and cross-cultural variation when assessing results from a Likert scale survey. This is particularly important in after-school contexts where youth with a variety of backgrounds, languages, and experiences come together to learn STEM. Likert scales, as with other self-report measures, offer limited insight into youth outcomes.

There have been developments in additional measuring techniques that provide more context on STEM programs while melding more seamlessly into the day-to-day learning activities of the youth involved (e.g., Fu et al., 2019). These techniques should be considered to augment evaluators' repertoires of tools, based on the specificities of activity contexts and program goals.

In our review, some OST STEM programs used qualitative data collection methods, which included observations (three) and interviews (three). Although these methods can require larger time and cost commitments, they also offer a number of benefits. For example, qualitative methods enable researchers to capture more contextual information than a Likert scale allows. Participants are not limited to a numeric response and may explain their beliefs and feelings in

detail. Qualitative data also points to different dimensions of the quality of STEM activities, as opposed to solely focusing on youth outcomes.

Constructs Measured in Existing Out-of-School STEM Evaluation Instruments

In this section we describe the different constructs (see Appendix B for definitions) we encountered in our review of the 36 tools that surfaced in our search. We made a list of the measures within each tool, totaling 76 measures. Those measures were then grouped together based on similar thematic constructs. Below we describe each of the 10 constructs for the 36 tools in alphabetical order.

Most tools measured multiple constructs, such as attitude, engagement, motivation or self-efficacy and learning. Almost every tool mentions at least two constructs, while some tools have a primary focus. For example, the Attitude Toward Science tool focused primarily on attitude, the Engagement survey focused primarily on engagement, and the Views of the Nature of Science Questionnaire focused primarily on the nature of science/views of science. (Appendix A lists all of the tools and provides more detailed information.)

We had access to different levels of detail on how the toolmakers theorized, defined, or validated their constructs. In some cases, we had peer-reviewed papers that carefully described how constructs were defined and measured. In other cases, we had only the tools themselves or web-/report-based accounts of what the tools were measuring at a broad level. By and large, the instruments did not distinguish between "science" as specifically pertaining to in-school or OST learning. As a result, in many cases, the descriptions below are meant to generalize across multiple tools or even across scales or items on tools. We alluded to specific definitions when they were available.

Attitude Toward Science

Attitude toward science was identified as a core component in seven of the 36 tools we discovered. Blosser (1984) describes an attitude toward science as how a person feels or behaves with respect to "scientists, scientific careers, methods of teaching science, scientific interests, parts of a curriculum, or the subject of science in the classroom" (quoted in Germann, 1988, p. 690).

One tool, the 4-H Science Initiative, measured aspects of participant attitudes towards and opinions about science and the 4-H science program itself. All of the instruments that measure

attitudes use student self-reports with multiple choice and/or Likert scale-like questions. Tools were used with learners ranging from 8 to 18 years of age.

Questions ranged from addressing students' emotional feelings about science as a subject matter (Attitude Toward Science Survey), to its utility (CARS), and their desire to learn more (Modified Attitudes Towards Science Inventory).

Example 1. Attitudes Toward Science Survey

(5-pt. Likert scale from *Strongly agree* to *Strongly disagree*)

(Q1). Science is fun.

(Q5). If I knew I would never go to science class again, I would feel sad.

Example 2. Changes in Attitudes About the Relevance of Science (CARS)

(5-pt. Likert scale from *Strongly agree* to *Strongly disagree*)

(Version A. Question 4). Science class helps me to evaluate my own work.

(Version C, Question 45). Using scientific methods helps me think things through.

Example 3. Modified Attitudes Towards Science Inventory

(5-pt. Likert scale from *Strongly agree* to *Strongly disagree*)

(Q24). I have a real desire to learn science.

Career Awareness and Career Interest

Nine of the measurement tools we found have at least one question related to students' awareness of or interest in STEM careers or future employment. Six of those have questions that are STEM-specific, while the remaining three explore questions about the environment and obstacles one may face in relation to STEM. All but one of the measurement tools in this section use a self-report Likert-type scale. The Exploring Youths' Interest-Related Pursuits tool involves an interview protocol where participants are interviewed by an adult in the program. The majority (eight) of tools were reported to be used with learners from age 7 to 18, while the Science Motivation Questionnaire II (SMQ-II) was designed for college students.

Questions ranged from directly asking students about their interest in pursuing science careers (Science Opinion Survey), to asking about the types of future employment they imagined (ROSE), to (in the interview) probing their thoughts about connections between their STEM activity and their future career trajectories (Exploring Youths' Interest-Related Pursuits).

Example 1. Science Opinion Survey

(5-pt. Likert scale from *Strongly agree* to *Strongly disagree*)

(Q2). I would dislike being a scientist after I leave school.

(Q14). A career in science would be dull and boring.

Example 2. The Relevance of Science Education (ROSE)

(4-pt. Likert scale from *Not important* to *Very important*)

How important are the following issues for your potential future occupation or job?

    Coming up with new ideas

    Becoming famous

Example 3. Exploring Youths' Interest-Related Pursuits

(Semi-structured interview protocol)

(Q24). As a result of engaging in [activity], have you gotten any new ideas about things you might want to do in the future?

[Prompt, if needed: it could be something you want to do as a hobby (like a sport), for school, for work, or to make the world a better place.]

b. (if no) - can you think of specific jobs [paid work] this might be preparing you (or other young people like you) for?

Curiosity

Curiosity in relation to STEM was measured in seven of the 36 tools found in Appendix A. One tool, the Children's Science Curiosity Scale (CSCS) measures curiosity specifically. The remaining tools include questions about curiosity but have overarching themes around fascination, innovation, attitude, connected learning, and motivation. Here we define curiosity as a positive reaction to new stimuli, expressing the desire to know more about oneself or one's environment, and examining/exploring stimuli in order to learn more about them (Harty & Beall, 1984). These tools have been used with learners between the ages of 9 and 18 with one, the SMQ-II, designed for college students.

Each survey uses a self-reporting Likert-type scale to examine curiosity. Questions ranged from asking about broad topics such as interest in the weather (Harty & Beall, 1984), to direct questions about interest in "science" (CSCS), to questions about making and learning (Innovation Stance in STEM).

Example 1. Children's Science Curiosity Scale (CSCS)

(5-pt. Likert scale from *Strongly agree* to *Strongly disagree*)

(Q2). I like to watch television programs about science.

(Q8). I want to know what causes wind.

Example 2. Fascination in STEM

(4-point Likert scale:1 [*YES!*], 2 [*yes*], 3 [*no*], 4 [*NO!*])

(f2). I like to figure out how things work

(f4). I want to learn as much as possible about math

Example 3. Innovation Stance in STEM

(4-point Likert scale:1 [*YES!*], 2 [*yes*], 3 [*no*], 4 [*NO!*])

(IS01). I like making new things even if I am not very good at it

(IS04). I try to learn new things even if I might make mistakes

Engagement

Engagement was measured in 11 of the 36 tools we identified. The Science Learning Activation Lab (2016b) defines engagement "as one's focus, participation, and persistence on a task (p. 1)." There are cognitive and behavioral dimensions of engagement, which may involve movement and gestures (Bell et al., 2019).

There were two tools that focused exclusively on engagement (the Engagement Survey and the Engagement Observation Protocol) and nine that measured some form of engagement.[1] Overall these tools focus on constructs such as attitude, success, and STEM learning. These tools were reported as being used with learners between the ages of 7 and 18.

Most of the tools (seven) used self-report Likert-type scales, while two of the tools were observation protocols. Questions ranged from asking about broad topics such as level of focus or attention (Science Activation Lab, 2016a, 2016b), to choices to continue to practice skill development (Organisation for Economic Co-operation and Development, 2006), to documenting the nature of student engagement in the science activities.

---

[1] The following tools listed are two of nine that measured some form of engagement. The Common Instrument was validated as a self-report measure of engagement. Also, Dimensions of Success is an observational measure of engagement in STEM activities.

Example 1. Engagement in Science Learning Activities

(4-point Likert scale:1 [*YES!*], 2 [*yes*], 3 [*no*], 4 [*NO!*])

(E01). During this activity: I felt bored.

(E05). During this activity: I was focused on the things we were learning most of the time.

Example 2. Survey of Principles of Connected Learning

(*Yes/No* responses)

(IP2). Please tell us if you have done the following things since you started participating in the activity:

c. Looked for things to do where you could get better at the activity?

(Select one: Never or Hardly Ever; 1-3 times a month; Once a week; More than once a week)

(PC2). When making or designing things while you are engaged in this activity, how often do you:

f. Try to influence what people think about an issue you care about?

Example 3. Engagement Observation Protocol

(Open-ended responses; observers record six dimensions of learner engagement)

- Sequential segments (different engagement type, points of child's transition, science content changes, activity structure)
- With whom? (adult, facilitator, peer, self)
- What was done? (ask, answer, connect, describe, discuss, etc.)
- Done with what? (metacognition, ideas, procedure, challenge/problems, artifacts, etc.)
- How does the learner participate? (active: takes initiative, passive+: listening, attentive, alert, etc.)
- Affect (+aroused, amazed, joyful, fun, happy, etc.)

Home/School Environment

Attributes related to home/school environment were measured in many (16) of the instruments in our review. None of these instruments listed this as a construct in its own right. Rather, we arrived at this category by grouping together several pre-existing measures. These included: access to resources (1), access to technology (1), books/resources at home (2), bullying (1), class climate (1), collaboration (2), environmental issues/topics (1), general habits (1), obstacles (1), opportunity (6), parental involvement/involvement from others (2), parent/guardian education (2), parent/guardian work (1), reading habits (1), relationships (2),

schedule (2), school habits (1), science courses (8), support (6), teacher/adult perceptions (2), and technology use (1).

Because home/school environment encompasses so many different categories, the relevant questions asked were diverse. Some tools, like the Program for International School Assessment (PISA), included multiple questions that articulated various dimensions of home/school environment. Others, like the Test of Science-Related Attitudes, focused on just one. Regardless, all 16 of these tools measured at least one aspect of how the students' STEM experiences may be mediated by their time at home or at school, based on the physical and social setting. Of these, 13 utilized a Likert-type scale, two incorporated interviews, and one used an interview protocol.

Items ranged from *learners enjoy pursuing enjoyable activities* (Survey of Principles of Connected Learning), to the *education level of parents* (Modified Attitudes Towards Science Inventory), to the *number of books available at home* (ROSE).

Example 1. Survey of Principles of Connected Learning
(*Yes/No* responses; short answer responses)
Think of an activity that:
- You enjoy doing
- You do with other people
- You get better at doing, the more you engage in the activity

(IP1). Where are all the places you pursue the activity?
- At home?
- At my school?

Example 2. Modified Attitudes Towards Science Inventory
(Multiple choice)
(Q3). The adults(s) with who I live have completed
  a. Elementary school
  b. Middle school
  c. Trade/vocational school
  d. 2-year college
  e. 4-year college
  f. I do not know

Example 3. Relevance of Science Education Questionnaire (ROSE)

(Multiple choice)

(J). How many books are there in your home?

There are usually about 40 books per meter of shelving. Do not include magazines.

(Please tick only one box.)

□ None

□ 1-10 books

□ 11-50 books

□ 51-100 books

□ 101-250 books

□ 251-500 books

□ More than 500 books

Interest

Fourteen of the 36 instruments measured some aspect of interest in STEM. Because none of them define interest, we rely on Hidi and Renninger's conceptualization as "a motivational variable [that] refers to the psychological state of engaging or the predisposition to reengage with particular classes of objects, events, or ideas over time" (2006, p. 112).[1] While 13 of the 14 tools have one or more questions about learners' interests in STEM, another tool, Exploring Youths' Interest-Related Pursuits, measures interest-related pursuits. Twelve of the 14 tools use a self-reported Likert-type scale, one is an interview (Exploring Youths' Interest-Related Pursuits), and one is an observation (Engagement Observation Protocol). The tools were reported as being used with learners between the ages of 8 and 19. Questions ranged from broad areas of interest such as "how things work" (Emerging STEM Learning Activation Survey), to more direct questions about interest in each of the four STEM disciplines (STEM-Related Scales), to interest in technology at school and in general (CATS).

Example 1. Emerging STEM Learning Activation Survey

(5-pt. Likert scale with frowning to smiling faces)

(Q9). I wish I could build things more often.

(Q14). I like to know how things work.

---

[1] For more on defining interest, see Bell et al. (2019).

Example 2. STEM-Related Scales

(5-pt. Likert scales from *Not interested* to *Very interested*)

How interested are you in science, technology, engineering and/or math (STEM)?

- a. Science
- b. Technology
- c. Engineering
- d. Math

Example 3. Children's Attitude Toward Technology Scale (CATS)

(4-pt. Likert scale from *Strongly agree* to *Strongly disagree*)

(Q9). I would like to learn more about technology at school.

(Q10). I am NOT interested in technology.

Motivation

Motivation was measured in just three tools, one of which solely focused on this construct. Motivation has been theorized extensively and has been operationalized in a variety of ways. Here, we offer the definition of Glynn et al. (2011), using social cognitive theory, as "An internal state that arouses, directs, and sustains goal-oriented behavior. By extension, the motivation to learn science can be defined as an internal state that arouses, directs, and sustains science-learning behavior" (p.1160).

Each of the three measurement tools include at least one question about motivation. The Science Motivation Questionnaire II (SMQ II) is designed to measure science majors' motivation to learn science in college. SMQ II is the only tool that measures motivation specifically, while the other two tools include one or more questions about motivation. Each tool uses a self-report Likert scale and can be used with learners from ages 10 to 18. Questions addressed topics such as the relevance of science and future career paths.

Example 1. The Science Motivation Questionnaire II (SMQ II)

(5-pt. Likert scale from *Never* to *Always*)

04. Getting a good science grade is important to me.

05. I put effort into learning science.

Example 2. Program for International Student Assessment (PISA) Student Questionnaire (2006)

(4-pt. Likert scale from *Strongly Agree* to *Strongly disagree*)

(Q18). How much do you agree with the statements below?

(e) I will use science in many ways when I am an adult

Example 3. Innovation Stance in STEM

(4-point Likert scale:1 [*YES!*], 2 [*yes*], 3 [*no*], 4 [*NO!*])

(IS03). I try to find new ways of doing things even if they might not work out

(IS04). I try to learn new things even if I might make mistakes

Nature of Science/Views on Science

Of the 36 measurement tools discussed here, 13 discuss some aspect of the construct of the nature of science or views on science (NOS/VOS). Four tools measure views of science specifically while the nine remaining tools have one or more questions about students' views of science. Views on science is defined as one's views and attitude on the nature of science (Chen, 2006). The nature of science as evidence-based reasoning and engaging with science to develop an evidence-based understanding, is seen as an important goal in science improvement efforts (e.g., NRC, 2009). These NOS/VOS tools described here have been used with a range of learners between the ages of 5 and 18. Views of science are thought to be linked to learners' engagement with and achievement in science (Khishfe & Abd-El-Khalick, 2002). There are two short-answer surveys, one interview, and the remaining tools are self-report surveys using Likert-type scales.

Example 1. Views of Scientific Inquiry—Primary School Version (VOSI-P)

(Open-ended interview questions)

(Q1). What kinds of work do scientists DO?

(Q2). Explain HOW scientists do their work. I'm not asking what they do but <u>how they do the work you just described for the last question?</u>

Example 2. Views on Science and Education Questionnaire (VOSE)

(5-pt. Likert scale from *Strongly disagree* to *Strongly agree*)

(Q2). Scientific investigations are influenced by socio-cultural values (e.g., current trends, values).

   Yes, socio-cultural values influence the direction and topics of scientific investigations.

   Yes, because scientists participating in scientific investigations are influenced by socio-cultural values

   No, scientists with good training will remain value-free when carrying out research.

   No, because science requires objectivity, which is contrary to the subjective socio-cultural values.

Example 3. Views on Science Technology Society (VOSTS)

(Multiple-choice)

Defining science is difficult because science is complex and does many things. But MAINLY science is:

> A. a study of fields such as biology, chemistry and physics.
>
> B. a body of knowledge, such as principles, laws and theories, which explain the world around us (matter, energy and life).
>
> C. exploring the unknown and discovering new things about our world and universe and how they work.
>
> D. carrying out experiments to solve problems of interest about the world around us.
>
> E. inventing or designing things (for example, artificial hearts, computers, space vehicles).
>
> F. finding and using knowledge to make this world a better place to live in (for example, curing diseases, solving pollution and improving agriculture).
>
> G. an organization of people (called scientists) who have ideas and techniques for discovering new knowledge.
>
> H. No one can define science.
>
> I. I don't understand.
>
> J. I don't know enough about this subject to make a choice.
>
> K. None of these choices fits my basic viewpoint.

Self-Efficacy (Competency Belief)

No measurement tool in our search explicitly examined self-efficacy. However, 17 of the 36 tools measured some aspect of it. Self-efficacy is commonly defined in the literature as individuals' beliefs about their ability to perform (Bandura, 1994). Many scholars posit that these beliefs determine how people motivate themselves and behave in life (cf. Schunk, 1991). Three of the 17 tools measure ability, seven measure confidence, and 11 measure metacognition.[1] The majority (14) of tools used a self-report Likert-type scale. One tool uses an observation protocol (Engagement Observation Protocol), one is an interview (Exploring Youths' Interest-Related Pursuits), and one is a multiple-choice survey (Test of Science-Related Attitudes). Tools are designed for learners between the ages of 9 and 18. Questions addressed issues such as reflection on learning approaches, thinking, and academic self-assessment .

---

[1] Self-efficacy has been shown to mediate metacognition in physics achievement (Yerdelen-Damar, S. & Peşman, H. 2013).

Example 1.Science Students' Metacognition, Self-Efficacy and Learning Processes (SEMLI-S)

(5-pt. Likert scale from *Never or almost never* to *Always or almost always*)

(AW1). I am aware of when I am about to have a learning challenge.

(CC3). I seek to connect what I learn in my life outside of class with science class.

Example 2. Test of Science-Related Attitudes (TOSRA)

(5-pt. Likert scale from *Strongly agree* to *Strongly disagree*)

(Q53). I Am unwilling to change my ideas when evidence shows that the ideas are poor.

(Q38). I would rather find out about things by asking an expert than by doing an experiment.

Example 3. Wareing Attitudes Toward Science Protocol (WASP)

(5-pt. Likert scale from *Strongly agree* to *Strongly disagree*)

(Q2). I am a persistent student.

(Q17). I do not feel I am bright enough for science.

Example 4. 4-H Science Youth Survey

(4-pt. Likert scale from *Never* to *Always*)

(Q10). When I have decisions to make . . .

b. I think before making a choice

STEM Practices

While most of the constructs described above are generally about students' attitudes and feelings about STEM, some questions related specifically to some or all of the eight specific practices of STEM (NRC, 2012). These include developing questions, designing experiments, developing explanations, using models, communicating results, et cetera (NRC, 2012). A majority of tools (25) have questions related to STEM practices. Questions include topics related to understanding, cross-curricular connection, difficulty, general knowledge, problem solving, skills, STEM knowledge, STEM learning, STEM skills, thinking processes, and working with others. Of these measurement tools one is an observation tool, one is an interview, and the rest are learner self-report surveys using Likert-type scales. Learners can range from ages five to 18. Examples of questions measuring STEM learning include:

Example 1. Science Process Skills Inventory (SPSI)

(4-pt. Likert scale from *Never* to *Always*)

6. I can use data to create a graph for presentation to others

7. I can create a display to communicate my data and observations

Example 2. Simpson–Troost Attitude Questionnaire—Revised (STAQ-R)

(5-pt. Likert scale from Strongly agree to *Strongly disagree*)

3. We learn about important things in science class.

Example 3. Views on Science Education (VOSE)

(5-pt. Likert scale from *Strongly agree* to *Strongly disagree*)

12. Students should understand that scientific knowledge may change.

    A. Yes, so they realize the real nature of science.

Discussion

From surveying 36 STEM measurement tools, we found a variety of constructs that programs used to assess participant learning and development. Most tools measured more than one construct, encouraging a multi-faceted analysis of program success. In our review we identified 10 constructs overall: attitude toward science, career awareness and career interest, curiosity, engagement, home/school environment, interest, motivation, nature of science/views on science, self-efficacy, and STEM practices.

In addition to these, the field of social and emotional learning (SEL) has been gaining popularity as a means to incorporate relevant measures of success in after-school STEM programs. Noam & Shah (2014) note that for young adolescents, STEM participation is often motivated by a sense of belonging, learning with friends, and caring adults. To become avid STEM learners, youth will need to cultivate 21st century skills related to emotional regulation and well-being, like the ability to manage stress and work in groups (Greenberg et al., 2017; see Chalkiadaki, 2018 for a literature review of 21st century skills in primary education).

However, in our survey just one tool explicitly measured the connection between STEM and SEL, the Common Instrument Suite (Allen et al., 2019; Sneider & Noam, 2019). There were also five tools that included questions about participants' feelings. Out of these tools, four of them use a self-reported Likert-type scale, and just one uses an observation protocol to gauge learners' SEL behavior. They are all designed to be used with learners between the ages of 7 and 18 years. There may be other tools as well, but assessment developers do not always explicitly state that they are designed for measuring SEL. One barrier to more rapid SEL measurement adoption is cost; the SEL tools were among the only ones in our survey that

required payment. This poses challenges for community-based STEM education spaces with limited financial resources.

Despite the explicit STEM focus of the programs surveyed, the measurements tended to center on just the "S" (science), leaving out the "TEM" (technology, engineering, and mathematics). It is unclear why this is the case. Perhaps this finding could highlight the disconnect between policymakers and educators in the United States; "STEM" was conceived by the National Science Foundation in the 1990s to develop skilled workers and boost the American economy, but the concept may confuse some educators who have not traditionally included engineering in their curriculum and may have varying definitions of technology (Blackley & Howell, 2015, p. 102). In addition, there may be an incentive for after-school programs to position their work under the STEM umbrella in order to become eligible for funding opportunities available to after-school STEM initiatives through the Department of Education, the Department of Defense, the National Science Foundation, and other avenues of federal, state, and local government (Afterschool Alliance, 2012). See Appendix B for our definitions of science, technology, engineering, and mathematics.

Along with the emphasis on science as a content area, the tools also emphasized learners' *interest* in science (e.g., views on science, attitude toward science). This seems to be related to the assumption that interest in science will lead to higher performance and ultimately academic and professional success in science. Interest development as it relates to emerging expertise has been theorized substantively in the extant literature (cf. Hidi & Renninger, 2006). However, there have also been notable critiques of the connection between interest and achievement. For example, there is a weaker correlation between interest and achievement among female students (Schiefele et al., 1992). Additionally, research in this area has often focused on correlation rather than causation (Kpolovie et al., 2014; Schiefele et al., 1992), leading educators to wonder whether promoting interest in science will lead to academic and professional success. Underrepresentation in STEM jobs is also the result of systemic factors, such as inequities in schools; pervasive racism and sexism; and limited access to material, relational, and ideational resources (Nasir, 2012). A few of the tools in this review asked about some related factors (see Home/School Environment), but by and large they were absent from measurement. These absences may perpetuate inequities in OST STEM learning.

Most of the program tools relied on participant self-reporting using Likert-type scales. This, paired with the individualistic focus of the constructs surveyed, suggests that the most

important data necessary to assess after-school program success are related to individual participants' self-perceptions (e.g., career interest) and performance (e.g., self-efficacy).[2] Success at the level of the individual has been a hallmark of sociopolitical ideals in the United States since the American Revolution and has continued to influence educational priorities through notions of meritocracy and achievement (Saa Meroe, 2014). In our analysis there was a paucity of measures related to perceptions and success at the level of the community, however they remain underdeveloped. For example, how are meaningful relationships being developed between participants, neighbors, after-school programs, and other community-based organizations? Are the voices, cultures, and ideals of underserved communities being amplified or suppressed (Garibay & Teasdale, 2019)? How is STEM learning connected to social change? It would be difficult to analyze these questions, among others related to community development, based on the available data.

Fortunately, there is evidence that after-school STEM programs that look to foster community, achieve positive results for underserved participants. For example, one study (Hughes, 2015) found that an after-school science program helped to build a collective identity among female participants, which in turn facilitated more participation in STEM practices and pathways to STEM careers. Evaluators may look at similar participant trajectories as they consider community-level measures.

Implications

A recurring theme in our findings was that social and political dimensions were largely absent from program measures. This is a particularly important consideration as we work towards a more equitable approach to STEM education. The STEM field has historically excluded people of color, women, people with disabilities and other nondominant groups. At the same time, after-school STEM programs are increasingly serving more underserved groups of learners. Pursuing justice in a landscape of unequal power relations and diverse learners requires attending to the multiple aspects of equity that Philip & Azevedo theorized (2017). This approach invites a pluralistic account of what equity means to diverse communities, and how different aspects are prioritized. For example, some communities may be focused on developing social capital by creating peer networks. Others may focus on fostering STEM commitment by ensuring learners have access to STEM courses.

---

[2] Self-reports may also simply be the easiest kind of data to collect.

In our survey we found measurements related to achievement, interest, and identity, but not much in the way of communities' everyday STEM practices or STEM as a tool for social justice and community development. Taking a justice-centered approach to this work requires moving beyond merely increasing underserved groups in a STEM-to-workforce pipeline and moving towards something Vakil (2018) describes as linking "learning to critical pedagogies of freedom and liberation by engaging the ethical and political implications as well as unrealized possibilities for technology to transform and empower communities" (p. 47). If after-school programs are committed to a justice-centered approach, they should incorporate tools that they can use to measure their success. This may include providing contextual information to supplement student self-reports, and/or rethinking the kinds of measures they are using altogether. Likewise, evaluators should recognize how the communities they serve are already working towards creating a more equitable STEM ecosystem, and what opportunities exist for centering these efforts in their evaluations.

Limitations

We believe this review is a useful starting point for an ongoing dialogue among practitioners, administrators, policymakers, evaluators, and researchers as we continue to better understand each other, our goals, and our strategies to achieve these goals. The constructs described here are some of the most prominent that are being measured by STEM programs in the United States. We aimed to create a comprehensive list of tools but realize there were some limitations along the way.

For example, we were unable to gain access to seven of the tools. This was either because we never heard back from the host organization or because the tool had not been fully developed at the time of our research. We were also limited to the tools listed in the OST STEM databases (see Conducting the Literature Review). It is likely that there are other relevant organizations measuring program success, but they did not show up in our database searches. Finally, it is possible that some of these tools may have been developed for in-school use but are being used by after-school programs. This context would affect what is being measured and, ultimately, how the organizations are thinking about learning.

For those who are interested in learning about some of the instruments that did not make it into our review, InformalScience.org is a great place to start. Some selection criteria that may be helpful in thinking through which instruments to examine include the number of settings in

which the instrument has been tested, the number of constructs measured, whether the instrument has been validated, and whether it defines each construct it aims to measure.

Acknowledgements

References

Afterschool Alliance. (2012). *Know your funders: A guide to STEM funding for afterschool.* https://www.afterschoolalliance.org/STEM-Funding-Brief-10182012.pdf

Allen, P. J., Chang, R., Gorrall, B. K., Waggenspack, L., Fukuda, E., Little, T. D., & Noam, G. G. (2019). From quality to outcomes: A national study of afterschool STEM programming. *International Journal of STEM Education*, *6*(1). https://doi.org/10.1186/s40594-019-0191-2

Azevedo, F. S. (2015). Sustaining interest-based participation in science. In K. A. Renninger, M. Nieswandt, & S. Hidi (Eds.) *Interest in mathematics and science learning* (pp. 281-296). American Educational Research Association. https://doi.org/10.3102/978-0-935302-42-4_16

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). Academic Press. (Reprinted in H. Friedman [Ed.], *Encyclopedia of mental health*. Academic Press, 1998). https://www.uky.edu/~eushe2/Bandura/BanEncy.html

Bang, M., & Vossoughi, S. (2016). Participatory design research and educational justice: Studying learning and relations within social change making. *Cognition and Instruction*, *34*(3), 173-193. http://dx.doi.org/10.1080/07370008.2016.1181879

Banks, J. A. (2006). *Race, culture, and education: The selected works of James A. Banks*. Routledge.

Barron, B., & Bell, P. (2015). Learning in informal and formal environments. In L. Corno & E. Anderman (Eds.), *Handbook of educational psychology* (3rd ed., ch. 24). Lawrence Erlbaum.

Barron, B. (2006). Interest and self-sustained learning as catalysts of development: A learning ecology perspective. *Human Development*, *49*(4), 193-224. http://dx.doi.org/10.1159/000094368

Bell, J., Besley, J., Cannady, M., Crowley, K., Grack Nelson, A., Philips, T., Riedinger, K., & Storksdieck, M. (2019). *The role of engagement in STEM learning and science communication: Reflections on interviews from the field*. Center for Advancement of Informal Science Education.

Bevan, B. (2016). STEM learning ecologies: Relevant, responsive, and connected. *Connected Science Learning*, *1*(1). NSTA Press. https://www.nsta.org/connected-science-learning/connected-science-learning-march-2016/stem-learning-ecologies

Bevan, B., Calabrese Barton A., & Garibay, C. (2018). *Broadening perspectives on broadening participation in STEM: critical perspectives on the role of science engagement*. Center for Advancement of Informal Science Education.

Blackley, S., & Howell, J. (2015). A STEM narrative: 15 years in the making. *Australian Journal of Teacher Education*, *40*(7). http://dx.doi.org/10.14221/ajte.2015v40n7.8

Blosser, P. E. (1984). *Attitude research in science education. Information Bulletin, No. 1.* ERIC Clearinghouse of Science, Mathematics and Environmental Education. https://files.eric.ed.gov/fulltext/ED259941.pdf

Briggs, L., Trautmann, N., & Phillips, T. (2019). Exploring challenges and lessons learned in cross-cultural environmental education research. *Evaluation and Program Planning*. *73*,156-162. https://doi.org/10.1016/j.evalprogplan.2019.01.001

Calabrese Barton, A., & Tan, E. (2010). We be burnin'! Agency, identity, and science learning. *The Journal of the Learning Sciences, 19*(2), 187-229. https://doi.org/10.1080/10508400903530044

Chalkiadaki, A. (2018). A systematic literature review of 21st century skills and competencies in primary education. *International Journal of Instruction*, *11*(3), 1-16. https://doi.org/10.12973/iji.2018.1131a

Chen, S. (2006). Views on science and education (VOSE) questionnaire. *Asia-Pacific Forum on Science Learning and Teaching, 7*(2). https://www.eduhk.hk/apfslt/v7_issue2/chensf/index.htm

Collins, P. H. (2018). Confronting color-blind STEM talent development: Toward a contextual model for black student STEM identity. *Journal of Advanced Academics*, *29*(2), 143-168. https://doi.org/10.1177/1932202X18757958

Collins, P. H., & Bilge, S. (2016). *Intersectionality*. Polity Press.

DiSalvo, C. (2012). *Adversarial design (Design thinking, design theory)*. MIT Press.

Engeström, Y. (1999). Expansive visibilization of work: An activity-theoretical perspective. *Computer Supported Cooperative Work (CSCW), 8*(1), 63-93. https://link.springer.com/content/pdf/10.1023/A:1008648532192.pdf

Falk, H. J., Dierking, L. D., Babendure, J., Pattison, S., Meyer, D., Verbeke, M., Coe, M., Palmquist, S., Storksdieck, M., & Canzoneri, N. (2017). *The learning SySTEM.* STEM Ready America. http://stemreadyamerica.org/the-learning-system/

Flaskerud, J. H. (1988). Is the Likert scale format culturally biased?. *Nursing Research*, *37*(3), 185–186. https://psycnet.apa.org/doi/10.1097/00006199-198805000-00013

Fu, A. C., Kannan, A., & Shavelson, R. J. (2019). Direct and unobtrusive measures of informal STEM education outcomes. *New Directions for Evaluation*, *2019*(161), 35-57. https://doi.org/10.1002/ev.20348

Garibay, C., & Teasdale, R. M. (2019). Equity and evaluation in informal STEM education. In A. C. Fu, A. Kannan, & R. J. Shavelson (Eds.), *Evaluation in informal science, technology, engineering, and mathematics education. New Directions for Evaluation*, 161, 87-106. https://doi.org/10.1002/ev.20352

Germann P. J. (1988). Development of the attitude toward science in school assessment and its use to investigate the relationship between science achievement and attitude toward science in school. *Journal of Research in Science Teaching, 25*(8), 689-703. https://doi.org/10.1002/tea.3660250807

Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, *48*(10), 1159-1176. https://doi.org/10.1002/tea.20442

Grack Nelson, A., Goeke, M., Auster, R., Peterman, K., Lussenhop, A. (2019). Shared measures for evaluating common outcomes of informal STEM education experiences. *New Directions for Evaluation*, 161, 59-86. https://doi.org/10.1002/ev.20353

Greenberg, M. T., Domitrovich, C. E., Weissberg, R. P., & Durlak, J. A. (2017). Social and emotional learning as a public health approach to education. *The Future of Children, 27*(1), 13-32. https://www.jstor.org/stable/44219019

Halim, L., Abd Rahman, N., Zamri, R., & Mohtar, L. (2018). The roles of parents in cultivating children's interest towards science learning and careers. *Kasetsart Journal of Social Sciences, 39*(2), 190-196. https://so04.tci-thaijo.org/index.php/kjss/article/view/242487

Harty, H., & Beall, D. (1984). Toward the development of a children's science curiosity measure. *Journal of Research in Science Teaching*, *21*(4), 425-436. https://doi.org/10.1002/tea.3660210410

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hughes, R. (2015). An investigation into the longitudinal identity trajectories of women in science, technology, engineering, and mathematics. *Journal of Women and Minorities in Science and Engineering, 21*(3), 181-213. https://ui.adsabs.harvard.edu/link_gateway/2015JWMSE..21..181H/doi:10.1615/JWomenMinorScienEng.2015013035

Jenkins, H., Itō, M., & boyd, d. (2016). *Participatory culture in a networked era: A conversation on youth, learning, commerce, and politics*. Polity Press.

Joshi, A., Kale, S., Chandel, S., & Pal, D. (2015). Likert scale: Explored and explained. *British Journal of Applied Science & Technology*, *7*(4), 396-403. http://dx.doi.org/10.9734/BJAST/2015/14975

Khishfe, R., & Abd-El-Khalick, F. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. *Journal of Research in Science Teaching*, *39*(7), 551-578. https://doi.org/10.1002/tea.10036

Kpolovie, P. J., Joe, A. I., & Okoto, T. (2014). Academic achievement prediction: Role of interest in learning and attitude towards school. *International Journal of Humanities Social Sciences and Education*, *1*(11), 73-100.

Lee, J. W., Jones, P. S., Mineyama, Y., & Zhang, X. E. (2002). Cultural differences in responses to a Likert scale. *Research in Nursing & Health*, *25*(4), 295-306. https://doi.org/10.1002/nur.10041

Merriam-Webster. (n.d.). Framework. In *Merriam-Webster.com dictionary*. Retrieved June 25, 2020, from https://www.merriam-webster.com/dictionary/framework

Merriam-Webster. (n.d.). Instrument. In *Merriam-Webster.com dictionary*. Retrieved June 25, 2020, from https://www.merriam-webster.com/dictionary/instrument

Merriam-Webster. (n.d.). Mathematics. In *Merriam-Webster.com dictionary*. Retrieved June 25, 2020, from https://www.merriam-webster.com/dictionary/mathematics

Merriam-Webster. (n.d.). Measure. In *Merriam-Webster.com dictionary*. Retrieved June 25, 2020, from https://www.merriam-webster.com/dictionary/measure

Nasir, N. S. (2012). *Racialized identities: Race and achievement among African American youth*. Stanford University Press. https://doi.org/10.1515/9780804779142

Nasir, N., Rosebery, A. S., Warren, B., & Lee, C. D. (2006). Learning as a cultural process: Achieving equity through diversity. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (489-504). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.041

National Academies of Sciences, Engineering, and Medicine. (2018). How people learn II: Learners, contexts, and cultures. *National Academies Press*. https://doi.org/10.17226/24783

National Research Council. (2009). *Learning science in informal environments: People, places, and pursuits*. The National Academies Press. https://doi.org/10.17226/12190

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. The National Academies Press. https://doi.org/10.17226/13165

National Research Council. (2015). *Identifying and supporting productive STEM programs in out-of-school settings*. Committee on successful out-of-school STEM learning. Board on Science Education, Division of Behavioral and Social Sciences and Education. The National Academies Press.

Noam, G. G., & Shah, A. (2014). Informal science and youth development: Creating convergence in out-of-school time. *Teachers College Record*, *116*(13), 199-218.

Ogden, J., & Lo, J. (2012). How meaningful are data from Likert scales? An evaluation of how ratings are made and the role of the response shift in the socially disadvantaged. *Journal of Health Psychology*, *17*(3), 350-361. https://doi.org/10.1177/1359105311417192

Organisation for Economic Co-operation and Development. (2006). OECD program for international student assessment 2006. https://nces.ed.gov/surveys/pisa/pdf/quest_pisa_2006_student.pdf

Philip, T. M., & Azevedo, F. S. (2017). Everyday science learning and equity: Mapping the contested terrain. *Science Education, 101*(4), 526-532. https://doi.org/10.1002/sce.21286

Saa Meroe, A. (2014). Democracy, meritocracy and the uses of education. *The Journal of Negro Education*, *83*(4), 485. https://www.jstor.org/stable/10.7709/jnegroeducation.83.4.0485

Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.),*The role of interest in learning and development* (183-212). Erlbaum.

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26*(3-4), 207-231. https://doi.org/10.1080/00461520.1991.9653133

Science Learning Activation Lab. (2016a). *Scientific sensemaking*. (Measures Technical Brief version 3.2). http://activationlab.org/wp-content/uploads/2018/03/Sensemaking-Report-3.2-20160331.pdf

Science Learning Activation Lab. (2016b). *Engagement in science learning activities*. (Measures Technical Brief version 3.2). http://activationlab.org/wp-content/uploads/2018/03/Engagement-Report-3.2-20160803.pdf

Sneider, C., & Noam, G. G. (2019). The common instrument suite: A means for assessing student attitudes in STEM classrooms and out-of-school environments. *Connected Science Learning*, *1*(11). https://www.nsta.org/connected-science-learning/connected-science-learning-july-september-2019/common-instrument-suite

Subedi, B. P. (2016). Using Likert type data in social science research: Confusion, issues and challenges. *International Journal of Contemporary Applied Sciences, 3*(2), 36-49. http://www.ijcar.net/assets/pdf/Vol3-No2-February2016/02.pdf

Vakil, S. (2018). Ethics, identity, and political vision: Toward a justice-centered approach to equity in computer science education. *Harvard Educational Review*, *88*(1), 26-52. https://doi.org/10.17763/1943-5045-88.1.26

Yerdelen-Damar, S., & Peşman, H. (2013). Relations of gender and socioeconomic status to physics through metacognition and self-efficacy, *The Journal of Educational Research*,*106*(4), 280-289. https://doi.org/10.1080/00220671.2012.692729

Appendix A

Measurement Tool List

4-H Science Youth Survey

http://www.pearweb.org/atis/tools/62

Attitude Toward Science

http://www.pearweb.org/atis/tools/attitude-toward-science-in-school-assessment-atssa

Changes in Attitudes about the Relevance of Science (CARS)

http://www.pearweb.org/atis/tools/27

Children's Attitude Towards Technology Scale (CATS)

http://www.pearweb.org/atis/tools/children-s-attitudes-toward-technology-scale-cats

Children's Science Curiosity Scale (CSCS)

http://www.pearweb.org/atis/tools/3

Common Instrument Suite

https://www.thepearinstitute.org/common-instrument-suite

Competency Belief Survey

http://activationlab.org/wp-content/uploads/2018/03/CompetencyBeliefs_STEM-Report_20170403.pdf

Dimensions of Success

https://www.thepearinstitute.org/dimensions-of-success

Draw A Scientist Test

http://www.pearweb.org/atis/tools/5

Emerging STEM Learning Activation

http://activationlab.org/wp-content/uploads/2018/08/Emerging_STEM-Report_08.02.18.pdf

Engagement Observation Protocol

http://stelar.edc.org/instruments/engagement-observation-protocol

Engagement in Science Learning Activities

http://activationlab.org/wp-content/uploads/2018/03/Engagement-Report-3.2-20160803.pdf

Exploring Youths' Interest-Related Pursuits

https://connectedlearning.uci.edu/wp-content/uploads/2019/05/Experiences-of-Interest-Interview-Protocol-Longitudinal-Survey-of-Connected-Learning.pdf

Fascination in STEM Survey

http://activationlab.org/wp-content/uploads/2018/03/Fascination_STEM-Report_20170403.pdf

Innovation Stance in STEM

http://activationlab.org/wp-content/uploads/2018/04/InnovationStance_STEM-Report_0425_2018.pdf

Modified Attitudes Towards Science Inventory

http://www.pearweb.org/atis/tools/7

Program for International Student Assessment (PISA)

https://nces.ed.gov/surveys/pisa/pdf/quest_pisa_2006_student.pdf

Relevance of Science Education Questionnaire (ROSE)

http://stelar.edc.org/instruments/relevance-science-education-rose-questionnaire

Scientific Attitude Inventory (SAI II)

http://www.pearweb.org/atis/tools/12

Science Motivation Questionnaire II (SMQ-II)

http://stelar.edc.org/instruments/science-motivation-questionnaire-ii-smq-ii

Science Opinion Survey (SOS)

http://www.pearweb.org/atis/tools/11

Science Process Skills Inventory (SPSI)

http://www.pearweb.org/atis/data/documents/000/000/010/Science_Process_Skills_Inventory_1.pdf

Science Students' Metacognition, Self-Efficacy and Learning Processes (SEMLI-S)

http://www.pearweb.org/atis/tools/32

Scientific Sensemaking Survey

http://activationlab.org/wp-content/uploads/2018/03/Sensemaking-Report-3.2-20160331.pdf

Simpson–Troost Attitude Questionnaire Revised (STAQ-R)

http://www.pearweb.org/atis/tools/21

STEM-Related Scales

https://conference.iste.org/uploads/ISTE2018/HANDOUTS/KEY_110836005/ISTE_RP.Impactsof
AfterSchoolRoboticsProgramsonSTEMInterests.5_30_18.pdf

Survey of Principles of Connected Learning

https://connectedlearning.uci.edu/wp-content/uploads/2019/05/Principles-Survey-Longitudinal-
Survey-of-Connected-Learning.pdf

Test of Science-Related Attitudes (TOSRA)

http://www.pearweb.org/atis/data/documents/000/000/004/TOSRA_BJF_.pdf

Valuing STEM Survey

http://activationlab.org/wp-content/uploads/2018/04/Values_STEM-
Report_20170403_online.pdf

Views of Nature of Science Questionnaire (VNOS-D)

http://www.pearweb.org/atis/tools/17

Views of Science and Education (VOSE)

http://www.eduhk.hk/apfslt/v7_issue2/chensf/chensf6.htm#six

Views about Science Survey (VASS)

http://www.pearweb.org/atis/tools/14

Views of Science Technology Society (VOSTS)

http://www.pearweb.org/atis/data/documents/000/000/002/vosts_2_.pdf

Views of Scientific Inquiry, Primary School Version (VOSI-P)

http://www.pearweb.org/atis/data/documents/000/000/016/VOSI-P_questionnaire.pdf

Wareing Attitudes Toward Science Protocol
http://www.pearweb.org/atis/tools/19

Women in Science Scale - Revised (WiSS-R)
http://www.pearweb.org/atis/tools/20

Appendix B

Glossary

Construct: the measurable part of an outcome (Grack Nelson, A., Goeke, M., Auster, R., Peterman, K., & Lussenhop, A, 2019)

Engineering: any engagement in a systematic practice of design to achieve solutions to particular human problems (National Research Council, 2012, p.11-12)

Framework: a basic conceptual structure (Merriam-Webster, n.d.)

Instrument: a measuring device for determining the present value of a quantity under observation (Merriam-Webster, n.d.)

Mathematics: the science of numbers and their operations (Merriam-Webster, n.d.)

Measure: refers to both a measured quantity and an instrument used for measuring (Merriam-Webster, n.d.)

Science: In the K-12 context, "science" is generally taken to mean the traditional natural sciences: physics, chemistry, biology, and (more recently) Earth, space, and environmental sciences (National Research Council, 2012, p.11-12).

Technology: all types of human-made systems and processes (National Research Council, 2012, p.11-12)