Measuring STEM Learning in Afterschool and Summer Programs:
A Review of the Literature
June 16, 2020

Erica Mallett & Ari Hock
with Bronwyn Bevan

1.0 Introduction

The most recent synthesis of research on learning confirms that learning is a complex cultural process that occurs across both time and setting (NASEM, 2018). The time spent outside of the school classroom – more than 80% of a school-aged person's waking hours – has been shown to be pivotal in developing interests, identities, and capacities to engage with ideas and subject matter they may also encounter in school, including STEM (Banks, 2006; National Research Council, 2006; OSTP, 2018).

Indeed, scores of studies have demonstrated the importance of family members, role models, and out-of-school time experiences for developing young people's awareness and commitment to pursuing particular STEM learning pathways (Bell et al., 2015; Falk et al., 2017; Halim et al., 2018; National Research Council, 2009, 2015). Out-of-school time (OST) experiences may include structured afterschool, weekend, and summer time 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., attitudes, 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 document is focused on two lines of argument in the literature that pertain to the value and potential of afterschool 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.

Afterschool and Summer STEM Programs and Equity

Particularly for young people from communities historically excluded from STEM fields, leveraging OST can be pivotal for advancing more equitable outcomes in who chooses to participate in and contribute to STEM learning activities and pathways (Calabrese Barton & Tan, 2012; Ito et al., 2016; Nasir, Rosebery, and Warren, 2015). 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 support student achievement in in-school STEM.
2. Informal STEM learning can deepen students' interest, identity, and excitement in STEM, which can then be capitalized on and developed by school STEM programs.
3. Informal STEM learning can expand young people's 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 young people concerned with broader issues of social justice and community development, thus supporting personal agency among broader communities of young people 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 some minoritized youth. Furthermore, engagement in informal settings may serve as an introduction to STEM for students in schools and districts where STEM opportunities are limited.

Some young people may be committed to STEM and seek and benefit from academic enrichment opportunities. These students are likely to already identify with STEM or academic STEM. Other young people who identify with social justice or community development struggles may come to appreciate and identify with STEM while using it for broader community purposes. For example, students concerned about environmental contamination near their school in the Bronx learn to code in order to program small robotic toy dogs to find danger spots (see Jeremijenko, 2019).

Programs may emphasize one or more of these characterizations of equity over the others. However, each one can make strong contributions to young people's relationship with STEM. They are not mutually exclusive, and can actually be complementary. The call from Philip & Azevedo (2017) is not to prioritize one approach over the other, but to articulate the approach, to understand why students might opt in or thrive with that approach, and to develop coherence across design, implementation, and evaluation.

Afterschool and Summer Programs in the STEM Learning Ecosystem

Because young people's interest in and commitment to STEM fields may fluctuate and change over the months and years, not to mention the lifespan (Azevedo, 2015; Barron, 2006), many have argued for the importance of providing multiple, diverse, and redundant opportunities for young people to engage in STEM, and for actively brokering awareness and uptake of these opportunities (Barron & Bell, 2016; 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 be rich in 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 young people in those communities who "belongs in STEM" and who doesn't (Bevan, 2019).

Enriching the STEM Learning Ecosystem means providing a range of inclusive and inspiring STEM learning options for young people 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 find programs that match their interests and needs at any moment in their development. This includes opportunities for young people committed to social justice to pursue academic-oriented STEM programs more deeply to refine their skills, as well as for young people committed to STEM academics to branch out and apply those skills in programs oriented towards community development or other goals (Collins & Bilge, 2016; Collins, 2018). Attention to the historical dimensions of STEM learning ecosystems, including the field's marginalization of people of color and women, is essential for ensuring full participation in such activities.

2.0 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. 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 afterschool and summer STEM programs aim to accomplish, how they do so, and how they document progress in meeting their goals for young people in STEM.

We conclude that the instruments reviewed assume two things: 1) that individuals will change based on appropriate intervention and 2) that improvement on these scales represents some type of success. However, the first assumption neglects systemic factors that contribute to inequitable outcomes, and the second assumption leaves us unclear as to whether the kind of success being measured is the kind of success we should be measuring. (See 4.0 Conclusion for more discussion.)

2.1 Methods

The primary procedures used to create a broad collection of OST STEM measurement tools, disaggregated by age and measures, were: (a) 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, and STEM Ready America), and (b) examining reference lists of subsequent tools, research papers, and websites.

Researching these tools took place over a three-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 do not posit that this list is complete, and welcome suggestions to add tools that we did not surface through our search. We began our search with the PEAR Institute search engine, 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.

2.2 Measuring with the Likert Scale

The majority (27) of the tools we identified used student self-reports on Likert scales. The Likert scale was developed to reliably and validly measure attitudes, which involve cognitive, affective, and psychomotor components (Joshi, Kale, Chandel, & Pal, 2015). These characteristics make the scale well-poised for measuring constructs in afterschool STEM programs. However, there are some limitations. A 1988 report on two studies with non-English-speaking refugees found several issues using the Likert scale cross-culturally. For example, several Spanish speakers preferred to use a dichotomous response (sí or no), rather than the continuum that was called for. It is possible that the amount of variation the scale intends to measure could be meaningless to some cultural groups (Flaskerud, 1988). Other indications of cross-cultural variation have been documented as well (cf. Briggs, L., N. Trautmann, & T. Phillips 2019; Lee, Jones, Mineyama, & Zhang, 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." We would be wise 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 afterschool 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 unobtrusive measuring techniques that provide more context on STEM programs while melding more seamlessly into the day-to-day learning activities of the youth involved (Fu, A. C., Kannan, A., & Shavelson, R. J., 2019).

In our review, the OST STEM programs used other data collection methods, which included observations (3) and interviews (3). Although these are time- and cost-intensive, there are benefits that justify the extra resources involved. For example, these data collection methods enable researchers to capture more nuanced responses than a Likert scale allows. Participants are not limited to a numeric response and may explain their beliefs and feelings in detail. This is particularly useful for the complicated nature of the constructs listed below.

3.0 Constructs Measured in Existing Out of School STEM Evaluation Instruments

In this section we describe the different constructs[1] we encountered in our review of the 36 tools surfaced in our search. We made a list of the measures within each tool, totaling 76 measures.

[1] We define a construct as "the measurable part of an outcome" (Grack Nelson, A., Goeke, M., Auster, R., Peterman, K., & Lussenhop, A, 2019).

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 different constructs, such as attitude, engagement, and motivation; or self-efficacy and learning. Each tool focuses on at least two constructs, though some primarily focus on one construct. 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.

We had access to different levels of detail on how the tools had 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. For example, many of the instruments contained items relevant to school learning. By and large, they 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.

Appendix A lists all of the tools and provides more detailed information.

We now present each construct, along with a brief description and some examples.

Attitudes Towards Science

Attitudes toward science was identified as a core component in 7 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, the 4-H Science Initiative tool, measured aspects of participant attitudes and opinions toward science and the 4-H science program itself. All of the instruments that measure attitude use student self reports using multiple choice and/or likert scale-like questions. Tools are reported to have been 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), its utility (CARS), and their desire to learn more (Modified Attitude Toward Science Inventory):

Example 1. ***Attitude Toward Science Survey***
(5 pt Likert Scale from strongly agree to strongly disagree)
(1) Science is fun.
(5) If I knew I would never go to science class again, I would feel sad.

Example 2. ***Changes in Attitude 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 Toward Science Inventory
(5 pt Likert Scale from strongly agree to strongly disagree)
24. 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. Most of those (6) have questions that are STEM-specific, while others (3) explore questions about the environment and obstacles one may face in relation to STEM. Almost all (8) of the measurement tools in this section use a self report Likert-type scale. One, the Exploring Youths' Interest-Related Pursuits tool, involves an interview protocol where participants are interviewed by an adult in the program. The majority (8) 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)
2. I would dislike being a scientist after I leave school
14. 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
24. 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 7 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, exhibiting the need to know more about oneself or one's environment, and examining/exploring stimuli in an effort 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 (Hart & 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 strongly disagree to strongly agree)
2. I like to watch television programs about science.
8. I want to know what causes wind.

Example 2. Fascination in STEM
(4-point Likert scale from No! to Yes!)
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 from No! to Yes!)
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. Engagement is defined as one’s focus, participation, and persistence on a task (Science Learning Activation Lab, 2016). There are cognitive, affective, and behavioral dimensions of engagement, which may involve movement and gestures (Bell, J., Besley, J., Cannady, M., Crowley, K., Grack Nelson, A., Philips, T., Riedinger, K., & Storksdieck, M., 2019a).

There were 2 tools that focused exclusively on engagement (the Engagement Survey and the Engagement Observation Protocol) and 9 that measured some form of engagement. The remaining 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 (7) use 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, 2016), to choices to continue to practice skill development (REF) to documenting the nature of student engagement in the science activities (REF).

Example 1. Engagement in Science Learning Activities
(4-point Likert Scale from No! To Yes)
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	W/whom	What done	What with	Participate	Affect
Different engagement type Points of child's transition Science content changes Activity structure	Adult Facilitator Peer Self	Ask, Answer, Connect, Describe, Discuss, etc.	Metacognition, Ideas, Procedure, Challenge/ Problems, Artifacts, etc.	Active: takes initiative, Passive+: listening, attentive, alert, etc.	+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 tech (1), books/resources at home (2), bullying (1), class climate (1), collaboration (2), education level (0), environmental issues/topics (1), general habits (1), language (0), 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 climate (0), school habits (1), science courses (8), support (6), teacher/adult perceptions (2), and tech 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, included multiple questions that got at 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, 2 incorporated interviews, and 1 used an interview protocol.

Questions ranged from where learners enjoy pursuing enjoyable activities (Survey of Principles of Connected Learning) to 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)

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)
3. 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 metre 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).[2] 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

[2] For more on defining interest, see Bell, J., Besley, J., Cannady, M., Crowley, K., Grack Nelson, A., Philips, T., Riedinger, K., & Storksdieck, M. (2019b).

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.

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)
9. I would like to learn more about technology at school.
10. I am NOT interested in technology.

Motivation

Motivation was measured in just three tools, one of which was 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, Brickman, Armstrong, and Taasoobshirazi (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 10 to 18. Questions addressed issues such as science relevance to daily life 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-Pt Likert Scale from NO! To YES!
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." Four tools measure Views of Science specifically while the remaining tools (9) have one or more questions about students' views of science. Views of science is defined as one's views about the nature of science and attitudes toward teaching and issues related to the nature of science (Chen, 2006). The nature of science as evidence-based reasoning, and of the scientific enterprise as the building of tentative, evidence-based understanding, is seen as a primary goal of current science improvement efforts (e.g., NRC, 2006, 2009). These NOS/VOS tools described here have been used with a range of learners between 5 and 18. Views of science are thought to be linked to learners' engagement and thus 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. Examples of questions measuring Views of Science include:

Example 1. Views of Scientific Inquiry - Primary School Version (VOSI-P)
Open-Ended Interview questions
1. What kinds of work do scientists DO?
2. 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 strongly disagree to strongly agree)
2. 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
10111 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 corpus explicitly examined self-efficacy. However, 17 of the 36 tools measured some aspect of it. Self-efficacy is commonly defined in the literature as "...people's beliefs about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives" (Bandura, 1994). Many scholars posit that these beliefs determine how people motivate themselves and behave in life (c.f. Schunk, 1991). Some (3) of the 17 tools measure ability, some (7) measure confidence, and some (11) measure metacognition.[3] The majority (14) of tools used a self-report Likert-type scale. One 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 9-18. Questions addressed issues such as reflection on learning approaches (REF), to thinking (REF), and academic self-assessment (REF).

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
53. I Am unwilling to change my ideas when evidence shows that the ideas are poor.
38. 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
2. I am a persistent student.
17. I do not feel I am bright enough for science.

---

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

Example 4. 4-H Science Youth Survey
4-Pt Likert Scale from Never to Always
10. When I have decisions to make...
b. I think before making a choice

STEM Practices

While most of the constructs described above are generically 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, etc. A majority of tools (25) have questions related to STEM practices. Questions range from 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 5 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.

4.0 Conclusion

From surveying 36 OST 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: attitudes towards 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, social and emotional learning (SEL) is another construct that is gaining popularity as a measure of success in afterschool STEM programs (Noam & Shah, 2014). However, in our survey just one tool measured the connection between STEM and SEL (i.e. the Common Instrument Suite, Allen et al., 2019; Sneider & Noam, 2019). There were also five tools that include 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 7 and 18 years old. One barrier to more rapid SEL measurement adoption is cost; the SEL tools were among the only ones in our survey that were protected behind a paywall. 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).[4] It is unclear why this is the case. Perhaps this finding could highlight the disconnect between policymakers and educators in the United States; “STEM learning” was conceived by the National Science Foundation in the 1990s to develop high-skilled workers and boost American economic productivity, but the concept has confused some educators who have not traditionally included engineering in curricula and may have “differing interpretations of the meaning of ‘technology’” as a subject (Blackley & Howell, 2015). In addition, there may be an incentive for afterschool programs to highlight their work under the STEM umbrella in order to become eligible for the myriad funding opportunities available to afterschool 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, 2019).

In our research and evaluation work, it’s important to understand the distinctions between the disciplines within STEM. For this, we rely on the National Research Council’s Framework for K-12 Science Education (2012, p.11-12):

> 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 . . . . We use the term “engineering” in a very broad sense to mean any engagement in a systematic practice of design to achieve solutions to particular human problems. Likewise, we broadly use the term “technology” to include all types of human-made systems and processes—not in the limited sense often used in schools that equates technology with modern computational and communications devices. Technologies result when engineers apply their understanding of the natural world and of human behavior to design ways to satisfy human needs and wants.

Along with the emphasis on science as a content area, the tools also emphasized learners’ *interest* in science (e.g. views on science, attitudes towards science). This seems to be couched in 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 among female students (Schiefele, Krapp, & Winteler, 1992). Furthermore, the research in this area has focused on correlation rather causation (Schiefele, Krapp, & Winteler, 1992; Kpolovie, Joe, & Okoto, 2014), leading educators to wonder whether promoting interest in science will lead to academic and

[4] There are some exceptions, like PEAR.

professional success. It may be useful to think about underrepresentation in STEM jobs as the result of systemic factors, such as inequities in schools, omnipresent racism and sexism, and access to material, relational, and ideational resources (Nasir, 2012). A few of the tools asked about related factors (see Home/School Environment), but by and large they were absent from measurement.

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 afterschool program success are related to individual participants' self-perceptions (e.g. career interest) and performance (e.g. self-efficacy).[5] 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 academic achievement (Aundra Saa Meroe, 2014). On the other hand, in our analysis there was a notable paucity of measures related to perceptions and success at the level of the community. For example, how are meaningful relationships being developed between participants, neighbors, afterschool programs, and other community-based organizations? Are the voices, cultures, and ideals of minoritized groups being amplified or suppressed (Garibay, C., & Teasdale, R. M., 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 has been some research that looks at the long-term impact of STEM programs on participants. For example, one study (Hughes, R., 2015) found that an afterschool 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.

4.1 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 minoritized groups. At the same time, afterschool STEM programs are increasingly serving multicultural groups of learners. Pursuing justice in a landscape of unequal power relations and diverse learner needs requires attending to the four aspects of equity that Philip & Azevedo theorized (2017; see Afterschool and Summer STEM and Equity). 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 underrepresented 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 afterschool programs are committed to a justice-centered approach,

---

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

they should incorporate tools that they may 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.

4.2 Limitations

This review sought to gain a better understanding of STEM learning ecosystems by synthesizing information on the tools and constructs being used to measure afterschool STEM program success. We aimed toward creating a comprehensive list of tools, but realize there were some limitations along the way. For example, we were unable to gain access to some (7) 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. In total, these tools include: Boys and Girls Club - Big Think STEM, Build IT, Dimensions of Success, Girls Inc., Project Exploration, Techbridge Girls, and Test of Mathematics Related Attitudes (TOMRA). We were also limited to the tools listed in the OST STEM databases (see Section 2.0). 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 afterschool programs. This context would affect what is being measured and, ultimately, how the organizations are thinking about learning.

4.3 Feedback from convening participants

Participants in the From Common Measures to Measures in Common national convening offered invaluable feedback to a previous draft of this review. We incorporated much of it into the current version. Due to the scope of the project, we realize we are unable to include everything. Here we list some of the document's strengths and weaknesses, as per the feedback we received.

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 OST STEM programs in the United States. The review highlights the importance of considering opportunities for youth to learn STEM by examining the various STEM ecosystems within which they live. It also foregrounds considerations for achieving more equitable learning outcomes.

However, it became clear that Philip & Azevedo's (2017) conception of equity in STEM is insufficient to define equity for the purposes of our work. Developing a shared definition of equity is an important cornerstone upon which our community may build. This may include more deeply considering the interplay between equity, identity, and intersectionality. It may also be useful to group all of the constructs reviewed according to the four categories that Philip & Asevedo (2017) have devised (see Introduction).

We received substantial feedback on the instruments we included – and excluded – in this review. Regrettably, we could not include all of the instruments that were suggested. For those who are interested in learning about more of them, 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.
There were also several comments related to the individual measures articulated. For example, some felt that Home/School Environment is more of a predictor of learning than a learning outcome. Others noted that identity was undertheorized in this review (and science identity, in particular).

One comment from a convening participant asked whether our goal is to share measures across different sectors, so as to track cumulative change over time; or whether it's to share measures across regions and communities. This is an important question to consider as we continue to think about what we are measuring and how we measure it.

References

Afterschool Alliance. Know Your Funders: A Guide to STEM Funding for Afterschool. Retrieved October 9, 2019, from https://www.afterschoolalliance.org/STEM-Funding-Brief-10182012.pdf.

Aikenhead, G. S. & Ryan A. G. (1992). The development of a new instrument: Views of science technology-society (VOSTS). *Science Education,* 76(5), 477-491. Retrieved from https://drive.google.com/file/d/1f4NxfKcUDbbS0myXuUw3AafwbAevHW1n/view

Arnold, M. E., Bourdeau, V. D., & Nott, B. D. (2013). Measuring science inquiry skills in youth development programs: The science process skills inventory. *Journal of Youth Development,* 8(1). Retrieved from https://drive.google.com/file/d/1Y8LafuKVmvx6WlfuAmgBHLyNkVy3Upvb/view

Aundra Saa Meroe. (2014). Democracy, Meritocracy and the Uses of Education. *The Journal of Negro Education*, *83*(4), 485.

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

Bell, J., Besley, J., Cannady, M., Crowley, K., Grack Nelson, A., Philips, T., Riedinger, K., & Storksdieck, M. (2019a). The Role of Engagement in STEM Learning and Science Communication: Reflections on Interviews from the Field. Washington, DC: Center for Advancement of Informal Science Education.

Bell, J., Besley, J., Cannady, M., Crowley, K., Grack Nelson, A., Philips, T., Riedinger, K., & Storksdieck, M. (2019b). The Role of Interest in STEM Learning and Science Communication: Reflections on Interviews from the Field. Washington, DC: 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).

Briggs, L., N. Trautmann, and T. Phillips (2019). Exploring challenges and lessons learned in cross-cultural environmental education research. Evaluation and Program Planning 73:156–162.

Chambers, D. W. (1983). Stereotypic Images of the scientist: The Draw-A-Scientist Test. *Science Education,* 67(2), 255-265. Retrieved from https://onlinelibrary.wiley.com/doi/epdf/10.1002/sce.3730670213

Chen, S. (2006). Views on science and education (VOSE) questionnaire. *Asia-Pacific Forum on Science Learning and Teaching,* 7(2). Retrieved from 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.

Collins, P.H. & Bilge, S. (2016). Intersectionality. Malden, MA: Polity Press.

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

Flaskerud, J. H. (1988). Is the Likert Scale Format Culturally Biased?. [Editorial]. *Nursing Research*, *37*(3), 185–186.

Frantom, C. G., Green, K. E., & Hoffman, E. R. (2002). Measure development: The children's attitude toward technology scale (CATS). Journal of Educational Computing Research, 36(3), 249-263. Retrieved from https://drive.google.com/file/d/12Akp5F6s3utgVIKZyD_p_DcgURG1Rjg4/view?usp=sharing

Fraser, B. J. (1982). Test of science-related attitudes. Australian Council for Educational Research. Retrieved from https://drive.google.com/file/d/1OYyYnkPPW8Qp5u5YIl_BoiyM9Scc5Xfe/view

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.

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.

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. Retrieved from https://drive.google.com/file/d/1iu5-SkPRaOWPYLVG47vMio1dcxwDIIE9/view?usp=sharing

Gibson, H. L. & Chase, C. (2002). Longitudinal impact of an inquiry-based science program on middle school students' attitudes toward science. *Sci. Ed.,* 86: 693-705. Retrieved from https://drive.google.com/file/d/1y5PYYCFg7XO3gK88onEQuPtZOIR4f-8I/view

Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. J*ournal of Research in Science Teaching,* 48(10), 1159-1176. Retrieved from https://drive.google.com/file/d/1IdANf_KiIKGIJiGueyDLadsiO-Grwnzo/view

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.

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. Retrieved from https://drive.google.com/file/d/1Qtgq5u66uOy9NXCek4r02XLBEVO-3MEz/view?usp=sharing

Hidi, S., & Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist*, *41*(2), 111–127.

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

Joshi, A., Kale, S., Chandel, S., & Pal, D. (2015). Likert Scale: Explored and Explained. *British Journal of Applied Science & Technology*, *7*(4), 396–403.

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. Retrieved from https://drive.google.com/file/d/1wDr-atqS8rjkKIasjtoB9AbReEGkiC4f/view

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.

Maul, A., Penuel, W. R., Dadey, N., Gallagher, L. P., Podkul, T., & Price, E. (2016). Measuring experiences of interest-related pursuits in connected learning. *Educational Technology Research and Development,* 65(1), 1-28. Retrieved from https://drive.google.com/file/d/1bYin77-tqmajG-gdqYX9X9yWAI-dR-FB/view

Melchoir, A. L., Burack, C., & Hoover, M. (2018, April). Impacts of after-school robotics programming on STEM interests and attitudes. Presented at the 2018 annual meeting of the American Educational Research Association. Retrieved from https://drive.google.com/file/d/1JXUYuzf1eliEhjsmDi0uYMJ1C-wP8UV9/view

Mielke, M., LaFleur, J., Butler, A., & Sanzone, J. (2012). 4-H science initiative: Youth engagement, attitudes, and knowledge study. Policy Studies Associates, Inc. Retrieved from https://drive.google.com/file/d/1DCBYl2dNzYsQRtr6KdUM4yTy81bxN_Ar/view

Modified Attitudes Toward Science Inventory.

Moore, R. W. & Hill Foy, R. L. (1997). The scientific attitude inventory: A revision (SAI II). *Journal of Research in Science Teaching,* 34(4), 327-336. Retrieved from https://drive.google.com/file/d/1MjgS_GypQIXjvSnBgIx540TcE_aCdL63/view

Nasir, N. S. (2012). *Racialized identities: race and achievement among African American youth*. Stanford, CA: Stanford University Press.

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press.

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.

Owen, S. V., Toepperwein, M. A., Marshall, C. E., Lichtenstein, M. J., Blalock, C. L., Liu, Y., Pruski, L. A., & Grimes, K. (2008). Finding pearls: Psychometric reevaluation of the simpson-troost attitude questionnaire (STAQ). *Science Education,* 92, 1076-1095. Retrieved from https://drive.google.com/file/d/1gX7Gz9VfC-aP6-fo1ADdU3PrSOkGOvwg/view

Schreiner, C. & Sjøberg, S. (2004). Sowing the seeds of ROSE. Department of Teacher Education and School Development University of Oslo. Retrieved from http://stelar.edc.org/sites/stelar.edc.org/files/ad0404-sowing-rose.pdf

Dale H. Schunk (1991) Self-Efficacy and Academic Motivation, Educational Psychologist, 26:3-4, 207-231

Schwartz, R. S., Lederman, N.G., & Lederman, J. S. (2008, March). An instrument to assess views of scientific inquiry: The VOSI questionnaire. Presented at the 2008 annual meeting of the National Association for Research in Science Teaching. Retrieved from https://drive.google.com/file/d/1AXjj5bt1RhUWICbtWB0shfUfdjXFMGiB/view

Science Learning Activation Lab. (2016). Scientific sensemaking version 3.2. Measures

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

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

Science Learning Activation Lab. (2017). Innovation stance in STEM version 1. Measures Technical Brief. Retrieved from http://activationlab.org/wp-content/uploads/2018/04/InnovationStance_STEM-Report_0425_2018.pdf

Science Learning Activation Lab. (2017). Competency beliefs in STEM version 1.0. Measures Technical Brief. Retrieved from http://activationlab.org/wp-content/uploads/2018/03/CompetencyBeliefs_STEM-Report_20170403.pdf

Science Learning Activation Lab. (2017). Fascination in STEM version 1. Measures Technical Brief. Retrieved from http://activationlab.org/wp-content/uploads/2018/03/Fascination_STEM-Report_20170403.pdf

Science Learning Activation Lab. (2017). Valuing STEM (Values). Measures Technical Brief. Retrieved from http://activationlab.org/wp-content/uploads/2018/04/Values_STEM-Report_20170403_online.pdf

Science Learning Activation Lab. (2018). Emerging STEM learning activation survey version 1.1. Measures Technical Brief. Retrieved from http://activationlab.org/wp-content/uploads/2018/08/Emerging_STEM-Report_08.02.18.pdf

Science Learning Activation Lab. (?). Engagement observation protocol. Retrieved from https://drive.google.com/file/d/18A-__K2HUWLfXfTFpPYrFTRyqCoRrAW9/view

Siegel, M. A. & Ranney, M. A. (2003). Developing the changes in attitude about the relevance of science (CARS) questionnaire and assessing two high school science classes. *Journal of Research in Science Teaching,* 40(8), 157-775. Retrieved from https://drive.google.com/file/d/1QP1VOL-MKAaGyCPrkclzGfulgfUrcjet/view?usp=sharing

Suter, L. E. (2016). Outside school time: an examination of science achievement and non-cognitive characteristics of 15-year olds in several countries. International Journal of Science Education, 38(4), 663-687. Retrieved from https://drive.google.com/file/d/1rWYIxHPfYoWbkqj6qQM3BMHNChXzbu_S/view

The PEAR Institute. (2019). A Guide to PEAR's STEM Tools: Dimensions of Success &

Common Instrument Suite. The PEAR Institute: Partnerships in Education and Resilience. Retrieved from https://docs.wixstatic.com/ugd/e45463_1079ed979ef54b47b9da405ab760b7b3.pdf

Thomas, G., Anderson, D., & Nashon, S. (2008). Development of an instrument designed to investigate elements of science students' metacognition, self-efficacy and learning processes: The SEMLI-S. *International Journal of Science Education,* 30(13), 1701-1724. Retrieved from https://drive.google.com/file/d/1JxPIUGN_EtQBJA6761bRk8vgAjwmAHgP/view

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.

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.

Appendix A

Measurement Tool List

4-H Science Youth Survey
https://drive.google.com/open?id=1DCBYI2dNzYsQRtr6KdUM4yTy81bxN_Ar

Attitude Toward Science
https://drive.google.com/file/d/142ezyuuv2_Y-_P_ZsnVmy0Ta0zru8lSu/view?usp=sharing

Changes in Attitudes about the Relevance of Science (CARS)
https://drive.google.com/file/d/16XnTdA2sXTj53PXHV_hiqEad3ReV6yGO/view?usp=sharing

Children's Attitude Towards Technology Scale (CATS)
https://drive.google.com/file/d/1n8zK_ZyIk_wCL8Kdo3VLnbenx-B-n_RT/view?usp=sharing

Children's Science Curiosity Scale (CSCS)
https://drive.google.com/file/d/1tepGCZyQ8S7XduELeOIy5UrgoXE7nsqn/view?usp=sharing

Common Instrument Suite

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

Dimensions of Success

Draw A Scientist Test
https://drive.google.com/file/d/1m5GuDZ9Z7MZlbzXVLzM5jQHv5OrP8q5K/view?usp=sharing

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

Engagement Observation Protocol
https://drive.google.com/file/d/18A-__K2HUWLfXfTFpPYrFTRyqCoRrAW9/view?usp=sharing

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

Exploring Youths' Interest-Related Pursuits
https://drive.google.com/file/d/1inUwwUgH8Lzi7uCEBjnYx8mH7nqxTK-i/view?usp=sharing

Fascination in STEM Survey
http://activationlab.org/wp-content/uploads/2018/03/Fascination_STEM-Report_20170403.pdf

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

Modified Attitudes Towards Science Inventory
https://drive.google.com/file/d/1UH5FCji4JoEfvYRXXuCArPvYr8sVhden/view?usp=sharing

Program for International Student Assessment (PISA)
https://nces.ed.gov/surveys/pisa/pdf/quest_pisa_2006_student.pdf

Relevance of Science Education Questionnaire (ROSE)
https://drive.google.com/file/d/13x8mrEj8QanVJRazGpM0ugX9InnGiXdR/view?usp=sharing

Scientific Attitude Inventory (SAI II)
https://drive.google.com/file/d/1MjgS_GypQIXjvSnBgIx540TcE_aCdL63/view?usp=sharing

Science Motivation Questionnaire II (SMQ-II)
https://drive.google.com/file/d/1F8nv4XmU2VBE1yLItaZ9DxbMoxxhXCU8/view?usp=sharing

Science Opinion Survey (SOS)
https://drive.google.com/file/d/15dNXuW7owMm8zE3WynpEnHO4s_f1vAEf/view?usp=sharing

Science Process Skills Inventory (SPSI)
https://drive.google.com/file/d/1YfkkOJcCV8k7MzAOx77qsAH6WU6KhNqW/view?usp=sharing

Science Students' Metacognition, Self-Efficacy and Learning Processes (SEMLI-S)
https://drive.google.com/file/d/1JxPIUGN_EtQBJA6761bRk8vgAjwmAHgP/view?usp=sharing

Scientific Attitude Inventory (SAI II)
https://drive.google.com/file/d/1MjgS_GypQIXjvSnBgIx540TcE_aCdL63/view?usp=sharing

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

Simpson Troost Attitude Questionnaire Revised (STAQ-R)
https://drive.google.com/file/d/1uRGpbH8rBprS0J3pLQvaXnsIe4tMC9Qs/view?usp=sharing

STEM Related Scales
https://drive.google.com/file/d/1JXUYuzf1eliEhjsmDi0uYMJ1C-wP8UV9/view?usp=sharing

Survey of Principles of Connected Learning
https://drive.google.com/file/d/1LyWg8D045YqhDgS2WCpurfPh-Wm2vl05/view?usp=sharing

Test of Science Related Attitudes (TOSRA)
https://drive.google.com/file/d/1OYyYnkPPW8Qp5u5YIl_BoiyM9Scc5Xfe/view?usp=sharing

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)
https://drive.google.com/file/d/1F_7pr_rAG3mR-_VmMg5R7ElrL5zZKiMb/view?usp=sharing

Views of Science and Education (VOSE)
http://www.eduhk.hk/apfslt/v7_issue2/chensf/chensf6.htm#six

Views about Science Survey (VASS)

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
https://drive.google.com/file/d/1C2pRsF1IzPIYRJyyO-iKJEWN_JRSp3nB/view?usp=sharing

Women in Science Scale - Revised (WiSS-R)
https://drive.google.com/file/d/1yFfuwsr7PWicoz6UK-gaPd2i0S-zzKfI/view?usp=sharing