Gil G. Noam, Ed.D., Ph.D. (Habil.)
Ashima M. Shah, Ph.D.

Game-Changers and the Assessment Predicament in Afterschool Science

June 2013

A paper inspired by the Summit on Assessment of Informal and Afterschool Science Learning organized by the National Research Council and the Program in Education, Afterschool, and Resiliency (PEAR) at Harvard University

CONTENTS

Executive Summary 5
Context 8
Science Learning Beyond the School Day 10
Assessing Science Learning: The Current Approach 12
The Game-Changers 14
  Game-Changer 1: OST Takes on New Expectations 16
  Game-Changer 2: Science Teaching and Learning Look Different 18
  Game-Changer 3: Organizations Collaborate to Pool Resources and Expertise 21
  Game-Changer 4: Stakeholders Demand Outcomes Data 23
Game-Changers and Opportunity 25
Designing Assessments, Shaping the Field 27
Choosing Core Outcomes for Afterschool Science 30
  Outcomes for Student Engagement 32
  Outcomes for Learning Science Content and Practices 34
Need for Common Assessments 36
  The Case for Common Assessments 37
  Challenges to Designing Common Assessments 38
  Level 1: Surface Structure—Fast and Large Scale 41
  Level 2: Deep Structure —Quality and Processes 41
    Measuring Outcomes Over Time 42
Concerns, Questions, and Suggestions 44
Call to Action 46
References 49
Appendix 1 52
Appendix 2 53
White Papers 55
Summit Participant List 56

"Our nation needs an educated young citizenry with the capacity to contribute to and gain from the country's future productivity, understand policy choices, and participate in building a sustainable future. Knowledge and skills from science, technology, engineering, and mathematics—the so-called STEM fields—are crucial to virtually every endeavor of individual and community life. All young Americans should be educated to be 'STEM-capable,' no matter where they live, what educational path they pursue, or in which field they choose to work."

– Carnegie Corporation of New York
Institute for Advanced Study Commission on Mathematics & Science Education, p.vii

Executive Summary

Out-of-school time (OST), especially afterschool programming, is starting to create more science learning experiences for youth that go beyond typical classroom science. These programs often have youth-development oriented staff, flexible time and space, and less pressure for using specified curricula or covering particular standards. In order to effectively leverage these diverse learning environments, stakeholders must come to some consensus about the best ways to engage youth in science during these additional hours. If the afterschool field does not take the lead in establishing relevant indicators and assessment tools, it will be forced to measure itself by state and national standards used in schools, which are often misaligned with the philosophy and goals of OST learning.

This report explores what we call an "assessment predicament," by positing and exploring four educational game-changers that we believe make the afterschool science field a hotbed of innovation. These game-changers also inform and guide how the afterschool field thinks about and creates evidence that captures the rich, diverse, and unique opportunities and contexts of out-of-school science learning.

Our inspiration for this paper was the 2012 Summit on Assessment of Informal and Afterschool Science Learning, organized by the Board of Science Education at the National Research Council and the Program in Educational, Afterschool, and Resiliency (PEAR) at Harvard University and McLean Hospital. Our paper reflects not only the many significant discussions held during the summit, but many subsequent conversations with leaders in the field as well.

The four game-changers are:

1. OST Takes on New Expectations: Science is sometimes short-changed as school districts face pressures to meet performance targets for subjects such as English and mathematics. The afterschool environment will increasingly be tasked with filling in the gaps, leading to additional pressure to provide quality science experiences.

2. Science Teaching and Learning Look Different: School science is no longer about step-by-step lab experiments and passive learning. New standards are pushing for an integrated focus on content and practices in which students "learn by doing" and teachers support this hands-on exploration.

3. Organizations Collaborate to Pool Resources and Expertise: Afterschool programs are often at the nexus of different organizations and learning resources. For example, a science center or a botanical garden may bring parts of their exhibits into OST settings. Additionally, afterschool programs are increasingly finding themselves the beneficiaries of a trend toward collaboration, such as groups of funders or sponsors coming together as networks to improve afterschool programming on a larger scale.

4. Stakeholders Demand Outcomes Data: The relatively new culture of assessment is here to stay and the OST field feels increasing pressure similar to schools to demonstrate outcomes.

In an educational climate shaped by these game-changers, the afterschool science field must build on existing assessment work, consider the development of common tools that assess particular outcomes across sites, and capture evidence about science quality at different levels.

Build on existing assessments. Self-report tools such as student and teacher surveys can track participants' experiences with OST programming. For example, student surveys allow youth to report their level of interest, engagement, or motivation for science. Observation tools, on the other hand, allow observers to look closely at interactions among teachers and students as science activities take place. As new assessments are designed, developers must be careful not to create a variation of the existing top down models, where national assessments are imposed upon the afterschool world.

Developing Common Assessments. It is important to focus in on the areas where informal and afterschool science can make the greatest difference and assess those over time, rather than measure many outcomes poorly or focus primarily on individual programs. By focusing on a relatively small set of outcomes, such as science engagement or science practices, the field can create common assessments that are used across programs, regardless of structure, curriculum, and location. This will allow the field to aggregate findings and create a common language for quality in afterschool science. Also, by aggregating data, assessment tools can be refined and improved and policymakers can see trends at a large-scale.

Collect Evidence at Different Levels. Two levels of assessment development are proposed in this report. The first level includes tools that are quick to administer and easily operationalized across programs. The second level includes tools that help measure the quality of programs more deeply, looking at characteristics of their activities, staff, and resources that make them more or less successful at engaging students in science. In-depth research and evaluation is also needed because programs differ greatly and the field does not want common assessments to create a culture of "teaching to the test" which could result in a reduction of innovation and creativity in program practices and weakening of students' learning experiences. In-depth work can capture the nuances of particular interventions and settings.

Moving Forward

As a next step, we propose a further refinement of an assessment agenda and define funding needs, followed by focused research in areas that will inform the creation of effective assessments. We also recommend that the four funders who made the Summit possible continue to pursue a shared agenda and bring other funders who are interested in innovative science learning is schools, programs and cultural instiutions "to the table." The diverse and growing world of afterschool science can encourage youth to be intrigued by the world around them, to make discoveries, to collaborate, to meet mentors, and to feel empowered by science. As a field, we must embrace this great opportunity and create assessments that can best measure what makes afterschool science worth investing in for the benefit of today's youth.

CONTEXT

From June 10-12, 2012, experts and leaders in assessment and measurement, informal science, and afterschool programming gathered at the *Summit on Assessment of Informal and Afterschool Science Learning* organized by the Board on Science Education at the National Research Council (NRC) and the Program in Education, Afterschool, and Resiliency (PEAR) at Harvard University and McLean Hospital. The questions guiding the Summit built on PEAR's 2008 report, *Toward a Systematic Evidence-Base for Science in Out-of-School Time: The Role of Assessment* (Hussar, Schwartz, Boiselle, & Noam, 2008). That report critically reviewed the state of assessment in science, technology, engineering, and mathematics (STEM) and set the foundation for a searchable website that is widely used called Assessment Tools for Informal Science or ATIS. After those two stages of field building were complete, it became clear that an expert convening was needed to help the field move forward at a time when assessment and evaluation had become more of a national focus.

The Noyce, Moore, Bechtel, and Samueli Foundations generously supported the Summit and charged PEAR to author a paper that highlights key points of consensus, controversy, and dilemma when designing assessments for science learning in afterschool and summer settings.

> *"...tens of millions of Americans, young and old, choose to learn about science in informal ways—by visiting museums and aquariums, attending after-school programs, pursuing personal hobbies, and watching TV documentaries, for example. There is abundant evidence that these programs and settings, and even everyday experiences such as a walk in the park, contribute to people's knowledge and interest in science." (LSIE, 2009)*

We want to thank first and foremost Ron Ottinger who was part of this plan from the very beginning. Ann Bowers, Alan Friedman, Penny Noyce, Cary Sneider, and Uri Treisman, who are also from the Noyce Foundation, played very active roles in creating the vision and planning the successive steps towards an assessment paradigm for the informal science field. They understood that this exciting and productive learning space can only get traction in the policy world if we take assessment and research very seriously. Their colleagues, George Bo-Lin and Janet Coffey at the Moore Foundation, Soo Venkatesan and Arron Jiron at Bechtel, and Gerald Solomon at the Samueli Foundation were equally committed and decided to

support the summit as a collaborative. We received a great deal of intellectual support at the National Science Foundation, especially from Sue Allen, Leslie Goodyear, Larry Suter, and David Ucko. After a year of preparation with some of these stakeholders, we approached Martin Storksdieck at the NRC who was very open to bringing our idea to the NRC and supporting the endeavor with his colleagues, Stuart Elliott, Margaret Hilton, and Judy Koenig. We'd also like to thank Cathy McEver for carefully capturing the many rich interactions at the Summit through her detailed notes and summaries.

This paper reflects the many ideas that emerged during the Summit and the further work and discussions we have had since with leaders in the field. We are indebted to the Summit participants (see participant list starting on p.53) and their thoughtful contributions. While this report does not include a transcript nor attribute particular ideas to individuals, it does attempt to share ideas and perspectives that were presented at the Summit. We intended to create a paper that would be most useful for a diverse audience, including evaluators, researchers, practitioners, and policymakers. Whenever appropriate and possible, we incorporated comments from colleagues who reviewed and provided feedback on earlier drafts of the report (see p.53). We benefitted greatly from the input and tried to incorporate as many ideas and positions without losing the main thrust of this paper.

SCIENCE LEARNING BEYOND THE SCHOOL DAY

School and afterschool environments often have complementary and overlapping goals (Noam, 2002; NRC, 2009; Pittman, et al., 2004). While reform efforts call for science to become more hands-on, engaging and student-centered, classroom teachers are often constrained by limited instructional time and pressures to prepare students to perform well on content assessments. Afterschool programs, on the other hand, typically have more flexibility to be more student-centered and youth development oriented. They can take the time to provide hands-on, interactive experiences where students can take ownership of their learning (Noam & Shah, in press). Unlike public schools, however, afterschool programs often rely on non-tax resources and attendance is not mandated by law. Federal and local governments, businesses, and foundations can choose whether or not to fund afterschool science programming, and therefore, many programs struggle for an identity and financial resources in an education system that lacks a coherent way of organizing and funding them.

In conversations among researchers and practitioners in schools, museums, science centers, and afterschool youth sites, the terms “informal science,” “afterschool science,” “informal STEM,” and “science in out-of-school time,” are often grouped together, or even used interchangeably. To sharpen our message, in this report we will focus on out-of-school time (OST) environments that offer programming with hands-on, engaging, and enriching science activities. We recognize that youth often visit museums, botanical gardens, science centers, zoos, etc. with their families or schools and have free choice and exploratory experiences. These types of experiences can be extremely valuable and enjoyable, but are outside the scope of this paper. Our suggestions will concentrate on OST or afterschool science programming (see Appendix 1) that has pre-planned activities, a designated facilitator/teacher/leader, and some structure (e.g., a science club that meets in the cafeteria afterschool twice a week, a community center’s afterschool program that has a 30 minute science block each day or a museum or science center’s camp or weekend programming with pre-planned activities and curricula).

In this report, we consider assessments for these experiences that take the form of sustained programming versus individual visits, as the assessment issues with each are different. While we will focus on assessment needs for afterschool science settings, we recognize that science learning is most effective when students are engaged, motivated, and encouraged to think and make meaning. These opportunities can and should happen both in and out of school.

ASSESSING SCIENCE LEARNING: THE CURRENT APPROACH

Based on numerous measures at the state, national, and international levels, great numbers of students are falling behind in science, technology, engineering, and mathematics (STEM) subjects. As a result, school districts, state departments of education, and policy makers at all levels are calling for "all hands on deck" to address this issue.

For example, many principals and state departments of education want a longer school day to give students more "time on task" with curricula that are aligned with standards at the state and district level. The Next Generation Science Standards (NGSS), scheduled to be adopted by states in 2013, will also guide decisions about science curricular decisions and assessments.

Meanwhile, some out-of-school time learning is becoming more integrated into the school day within the expanded day learning model (Malone, 2011). Bringing informal science opportunities into all day learning is one potential way to address these standards and for some, to cover more content.

Unfortunately, the very characteristics that make afterschool science an alternative entry point for students who might not connect with school science could be lost as stakeholders disagree about how to use time outside of school. There is no consensus about the structure and shape of the 15 hours or so a week that could be used for additional science instruction through extended, expanded, or afterschool learning and the four to six weeks available in the summer.

We call this the "struggle for the soul of OST," and believe it needs to be taken into consideration when thinking about appropriate assessments. Adding to the complexity, the different players in the afterschool world often have perspectives on assessments that do not always align. For example, some funders want to see the impact of their contributions; others want all their funds to be used to improve materials and the quality of an experience. Still other funders want to understand the strengths and weaknesses of particular programs that they support, so they can make smarter future investments in the field (Friedman, 2012, WP, p.3). Practitioners, on the other hand, can have mixed views on assessments, as they can help them improve their activities, but they can also hold them to particular outcomes that they may not feel are appropriate for their particular program vision or goals.

Thus, the afterschool field faces an assessment predicament: if it does not establish some indicators and tools of its own, it will increasingly be forced to measure itself by state and national test outcomes, which are mismatched with the philosophy and goals of many informal science experiences in OST. The *Defining Youth Outcomes for STEM Learning in Afterschool* study surveyed groups of experts to define what could be possible, appropriate, and feasible outcomes and indicators for all afterschool STEM programs (Afterschool Alliance, 2013). We will review some of its findings later in this paper.

How do we make these decisions moving forward? What is the state of science in out-of-school time? In this report, we will share four key game-changers in education that we believe are pushing the afterschool science field to become a potential hotbed of innovation. These forces have led to new ways of thinking about time in and out of school for science learning, new approaches to teaching and learning, as well as various entry points for engaging different types of learners.

THE GAME-CHANGERS

In the field of education, there have been many major shifts over the past decades. We call these game-changers because they do not represent incremental developments; instead, they create new approaches and configurations that allow for creativity and innovation in the educational ecosystem. The game-changers create fundamental evolutions in how we understand education, the learner, and the use of time for science learning. They also provide new possibilities for theory and research. The four game-changers focus on the following trends:

- OST is taking on more responsibility for providing students' access to STEM learning experiences
- There is a push for more integrated learning that focuses on core concepts and practices
- Increasingly, organizations are partnering in order to pool resources and expertise
- There is a need for outcome data on the impact of students participating in science programming during OST

While schools can be confined to their existing structures and are often slower to change, afterschool programs have the advantage of being able to try new approaches, new curricula, and new ways of organizing their time and space. In other words, they can respond more rapidly and nimbly to these game-changers.

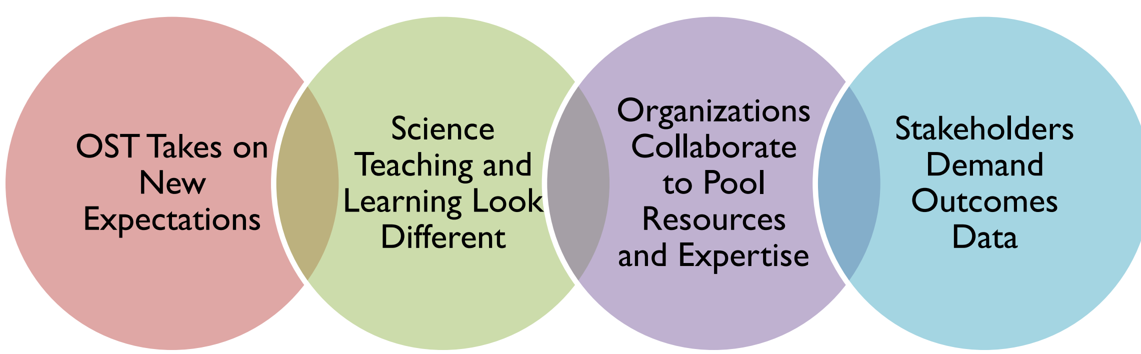


Figure 1 *Four game-changers in the educational landscape as identified by the authors to structure this report.*

Afterschool science programs take place in different settings and have varying goals and organizational structures, with no centralized authority overseeing all of their operations. Afterschool science allows the flexible time and space for new approaches to learning and offers a laboratory to experiment with ways to respond to and flourish amidst these game-changers. We anticipate that innovation from outside the established school day will influence changes to the status quo. In order to respond to the game-changers, the field will need to transform some of its traditional ways of doing things and provide new openings for experimentation and application. We will now dig deeper into the four game-changers and how they influence the direction of this report's greatest concern: assessments for science teaching and learning in OST. While there are many new developments in the educational and STEM space, we chose to focus on these four game-changers because we think they have the potential to shift and shape the future of STEM in OST. Understanding these changes is critical to shaping an agenda for assessment in OST because with these changes will come new learners, new learning environments, new expectations, and new assessments.

Game-Changer 1: OST Takes on New Expectations

The role of afterschool programming is shifting. This youth development environment can improve students' opportunities to engage with science and to become familiar with possible science careers.

While some afterschool programs are still seen as a sanctuary for children of working parents, where they can engage in enriching activities and stay away from the dangers of unsupervised time, the role of afterschool in the educational landscape has shifted significantly. Afterschool programs are now receiving funding to create learning environments that can complement and supplement school learning. Changes in school day priorities have also contributed to new afterschool expectations. For example, subjects such as math and reading are often tested more frequently than science, so science often gets shortchanged. In a statewide study in California, 40 percent of elementary teachers reported that they spend just 60 minutes or less teaching science each week and only one-third of elementary teachers reported feeling prepared to teach the subject (www.lawrencehallofscience.org/story/statewide_study).

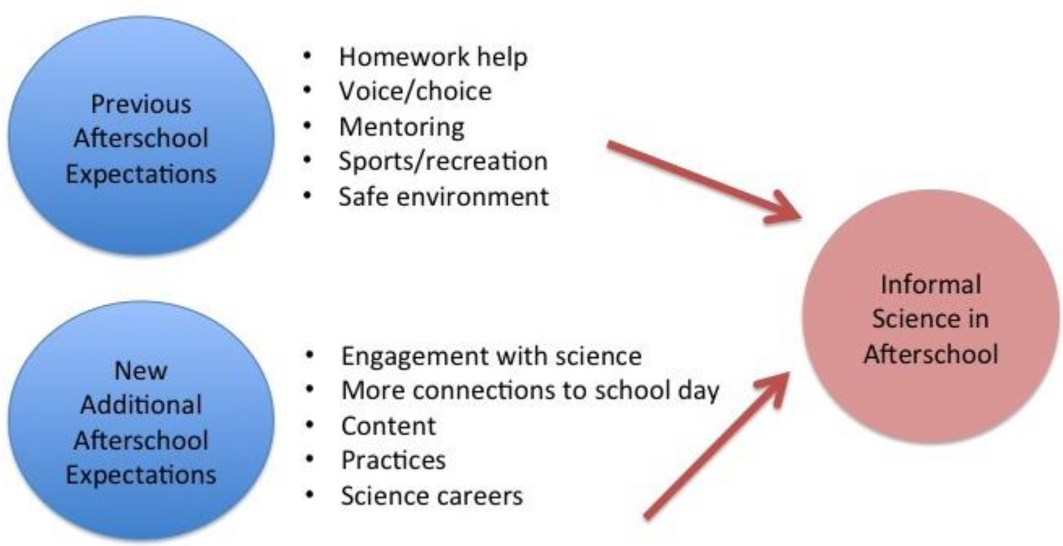


Figure 2: *Shifting role of afterschool to incorporate access to science learning opportunities.*

Given the scant amount of time devoted to science and the minimal preparation and confidence of some classroom science teachers, participating in a daily afterschool science club or joining a multi-week science summer camp can significantly contribute to the science education of youth. Afterschool and summer programs often have an advantage over school science activities: they can use time more creatively and leverage more flexible teaching approaches. While a teacher in a formal school classroom may have only a class period for students set up, conduct, and collect data from an experiment, an afterschool program may be able to have students collect data over days or even weeks, helping students learn more about a particular content area. All these changes have placed added pressure on afterschool programs to provide quality science experiences, whether they are prepared to or not.

Game-Changer 2: Science Teaching and Learning Look Different

The typical approaches of using step-by-step lab experiments or science textbooks filled with vocabulary words and diagrams have been shifting over the past decades. Science educators agree that students learn science by asking questions, building on prior knowledge, and refining explanations through exploration (National Research Council, 2000).

New standards are also re-defining how we think about science teaching. For example, the Common Core and the Framework for K-12 Science Education that were the foundation for the Next Generation Science Standards calls for more integrated and relevant learning that prepares youth for today's fast-paced, technologically rich world (see Figure 3).

Common Core Habits of Mind

- *Make sense of problems and persevere in solving them*
- *Use appropriate tools strategically*
- *Respond to varying demands of audience, task, purpose, and discipline*
- *Use technology and digital media strategically and capably*

The Forum, 2012 (www.forumfyi.org/files/ost_7.31.pdf)

Figure 3: *The Framework and Common Core Standards are strongly recommending an integrated focus on content and practices so students can actively build their understandings. These standards increase the expectation for learning outlined in general afterschool quality standards (e.g., California Afterschool Program Quality Self-Assessment Tool).*

Additionally, these standards push for teaching, learning, and assessing students not only on what they know but also on whether they can demonstrate their knowledge. This emphasis on active performance can really build engagement, motivation, and interest—areas that are very important to the afterschool world (Krajcik, 2012, WP). Therefore, afterschool science can be the perfect space to help propel these changes in science teaching, since it is often difficult to change the structure of science classroom learning. The "Habits of Mind" listed in the Common Core "…encompass a range of skills that are critical both to academics but also to success in work and life" (Forum, 2012). Several of these skills are integral components of STEM activities. The NGSS framework (NRC, 2012) encourages science learning that goes beyond listing isolated and disconnected facts, and instead focuses on core ideas by using cross-cutting concepts (that apply across science domains like physics, chemistry, biology, etc.) and practices (behaviors that scientists engage in) that build over years so learning opportunities have coherence and foster deeper understanding (http://www.nextgenscience.org/).

Sample of Scientific and Engineering Practices from the Framework for K-12 Education:

- Asking questions
- Developing and using models
- Planning and carrying out investigations
- Analyzing and interpreting data

(NRC, 2012)

Teaching science through inquiry has been one reform effort focused on students engaging in more scientific practices and building deeper content understanding (NRC, 2000). With the reforms in the last two decades, new approaches to teaching science have begun to make their way into many curricula and schools; however, moving beyond the textbook to engage students in inquiry is very challenging and requires a lot of preparation and training for teachers (e.g., Crawford, 2007; Davis, Petish, & Smithey, 2006). Therefore, there is still a great deal of work to be done so that these practices become prevalent in schools (e.g., Anderson & Helms, 2001).

On the other hand, afterschool programs already provide hands-on, exploratory types of activities. With new efforts to bring science into these environments, afterschool facilitators, who are already skilled at engaging students, now have to learn how to engage them in science specifically. There are many opportunities to increase access to science learning opportunities in afterschool. Programs differ dramatically in terms of available resources, science curricula, and training for staff. With federal and private funding available, the field must think about ways to assess what these experiences provide, knowing that the impact is purposefully and importantly different than the experiences in school.

> *Because the NGSS framework aligns particularly well to the student-centered, project-based methods of teaching and learning long embraced by informal educators, promising practices honed in informal settings have the potential not just to support the work of schools in implementing NGSS, but to help lead the way forward. This is particularly true in the area of building students' interest and motivation to engage in STEM learning – the foundation upon which proficiency, skills and eventual pursuit of STEM careers is built.*
> *(Traphagen, Sneider, & Morrison, 2012, p.1)*

Standards and new pedagogies act as game-changers because while they struggle to fit into the decades-old school science paradigm, they can have a different impact on afterschool science programming, which is still developing. Afterschool programs offer a new paradigm for learning—learning that is relevant, sustained, aspirational, and anxiety-free, with mentoring and social/emotional support. The strategies used in these settings can help bolster the experiences students are having more formally in school.

Game-Changer 3: Organizations Collaborate to Pool Resources and Expertise

Afterschool programs can be defined as "intermediary spaces," as they are positioned at the intersection of different organizations and resources for learning. These spaces can help students translate what they learn in school to afterschool, to the community, to their personal lives, and beyond (Noam, Biancarosa, & Dechausay, 2002; Noam, 2001). Furthermore, afterschool programs allow for connections between community-based settings and programs and educational opportunities for youth. For example, science centers can bring parts of their exhibits into afterschool settings, or they can run programs on-site. Afterschool programs may have partnerships with a local botanical garden, biotechnology research facility, or environmental group. Due to the flexibility of afterschool programs, some leaders and administrators are seeing the value in pooling resources (e.g., science materials, learning spaces, technology, funding) and expertise (e.g., local scientists and experts, STEM graduate students and researchers, community leaders, etc.) in creative ways to maximize the quantity of science programming as well as the quality. In addition to individual programs pairing up with other organizations or programs, there is also a trend of groups of funders or sponsors coming together as networks to coordinate and improve afterschool science programming on a larger scale. These connecting forces can help provide much needed access for students across the country. A recent report by *Change the Equation* points out that only 19% of households take advantage of STEM programming offered during out-of-school time. Furthermore, participation in STEM programming in OST is low among elementary and high school students, and participation is higher in lower-income urban areas but very low in rural areas (Change the Equation, 2012).

Schools cannot singlehandedly prepare students to engage and achieve in science… partnerships, connections, and links among different organizations and efforts are needed to pool resources, capitalize on a range of expertise, and increase both quantity and quality of learning opportunities.

There are several examples of these partnerships: local administrations (e.g., a mayor's office) are connecting different organizations (e.g., hospitals, universities, school districts, community youth organizations) to create learning opportunities for students, especially in inner cities.

The Collaborative for Building After-School Systems (CBASS) is supporting citywide afterschool systems in cities such as New York, Chicago, Baltimore, Boston, Providence, and Washington, D.C. These cities' intermediary organizations, as they are frequently called, connect afterschool programs with neighborhood groups, cultural institutions, policymakers and funding sources. Through these connections, creative, innovative organizations and their networks can increase capacity and programming quality.

Another significant development is the 21st Century Community Learning Centers initiative funded by the U.S. Department of Education. It supports afterschool programming that offers academic enrichment and youth development supports, focused on students who attend high-poverty and low-performing schools (http://www2.ed.gov/programs/21stcclc/index.html). Many centers have started incorporating science activities into their offerings.

Statewide afterschool networks have been organized as well, to coordinate efforts to improve science programming across states, in and out of school. For example, in California, the *Power of Discovery: STEM*^{2} initiative of the California Afterschool Network (CAN) and the California STEM Learning Network (CSLNet) are spearheading a large-scale initiative to provide technical assistance and resources for afterschool providers across the state to improve their practices and increase students' STEM interest, knowledge and skills (http://stem.afterschoolnetwork.org/rfq). The Mott Foundation has funded large afterschool networks and has recently begun an initiative with the Noyce Foundation to create a network of five STEM-focused state networks with plans to grow and reach more states over time. This trend toward coordination and network development is in part due to funders' interest in supporting initiatives that go beyond individual programs and can lead to greater impact. By combining efforts, the potential for afterschool opportunities in science can be significant.

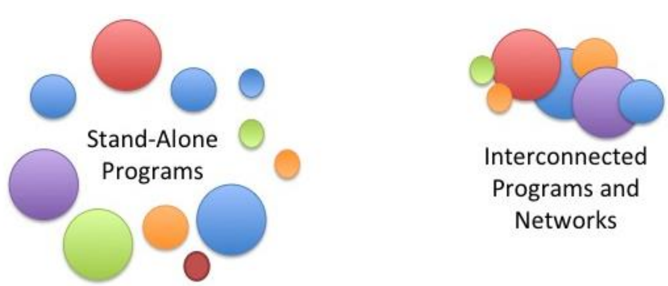


Figure 4: *Weaving together different configurations of resources, settings, and expertise will result in greater impact.*

Game-Changer 4: Stakeholders Demand Outcomes Data

While national and state assessments are often criticized for causing schools to place too much emphasis on test preparation and being a source of anxiety and pressure for teachers and students, the culture of assessment seems to be here to stay, and getting stronger. Assessments have been a transformational force, changing the very landscape of education. Teachers must teach and students must learn a set of standards and perform well on the assessments to assure their school receives funding and supports (e.g., No Child Left Behind Act of 2001).

The afterschool world, including OST science programming, is pressured to show outcomes. For example, The National Science and Technology Council Committee on STEM Education (CoSTEM), mandated by the America COMPETES Reauthorization Act of 2010, calls for two outcomes: 1) increasing interest in and identification with STEM, and 2) increasing participation of learners from underserved groups (Feder & Weiman, 2012, WP, p.5). They are also interested in measuring how STEM programming can guide more youth to pursue STEM careers and how it can increase overall interest and proficiency in science.

> Afterschool science programs are unique and offer a range of opportunities to youth. Unfortunately, differences in where they take place, when they happen, how they happen, and for what purpose they are designed are what make them difficult to assess with existing instruments. Progress needs to be swift to avoid school-based *assessments*

With these ambitious goals, there is also federal pressure to develop evidence standards for determining that programming is moving towards high quality STEM outcomes. And private funders looking to invest in afterschool networks also want proof that their contributions are affecting the lives of youth and engaging them with science in ways beyond what the school day can provide. Federal funding streams for afterschool programming (e.g., 21^{st} Century Community Learning Centers), National Science Foundation-funded interventions and many private foundations that invest in afterschool programs are calling for evidence of science learning. In order to secure and maintain funding, afterschool programs need to produce positive outcomes at the student and program level. There are efforts to involve practitioners and evaluators in the conversations needed to define possible outcomes (e.g., Delphi study by Afterschool Alliance, 2013).

In this paper, we identified game-changers that we feel are working together to create a climate conducive to new approaches and perspectives on assessment and learning in STEM. Together, they have created the potential for significant innovation and change. Specifically, in the area of assessments, the afterschool field is in an exciting position to specify what it does for youths' learning trajectories and how its value can be defined and measured. A new suite of assessments must be developed that can:

- Address the integrated, practice-oriented ideals of the latest standards
- Be used across large networks of programs
- Align with new pedagogies well suited to afterschool science
- Measure outcomes that are relevant to these out-of-school experiences where students can be engaged in and motivated by science in ways that can propel their future educational, career, and life decisions.

GAME-CHANGERS AND OPPORTUNITY

The game-changers have provided the field with an opportunity—an exciting moment that encourages creativity, collaboration, and a chance to build the field thoughtfully and innovatively. By extending and expanding time, using both school settings and informal spaces, a whole new world of learning has opened up. Science learning is often hands-on and collaborative. It integrates content ideas while exposing students to practices used in the field. Thus, science activities that take place both during the school day and afterschool can build strong links between the settings and also highlight the strengths of both types of learning environments.

No longer should the discussion be about school versus afterschool, but instead, about the best ways to learn science and how school and afterschool settings can provide complementary and unique access to that learning. This perspective has important implications for how we think about assessments. In Figure 5, we illustrate how the game-changers give rise to a set of actions in afterschool science assessments. By designing a system of assessments to respond to and capitalize on the game-changing forces in education, the afterschool field can take ownership of what assessments look like, how they should be administered, and what they should tell us that is different from what is revealed by assessments in the formal educational world.

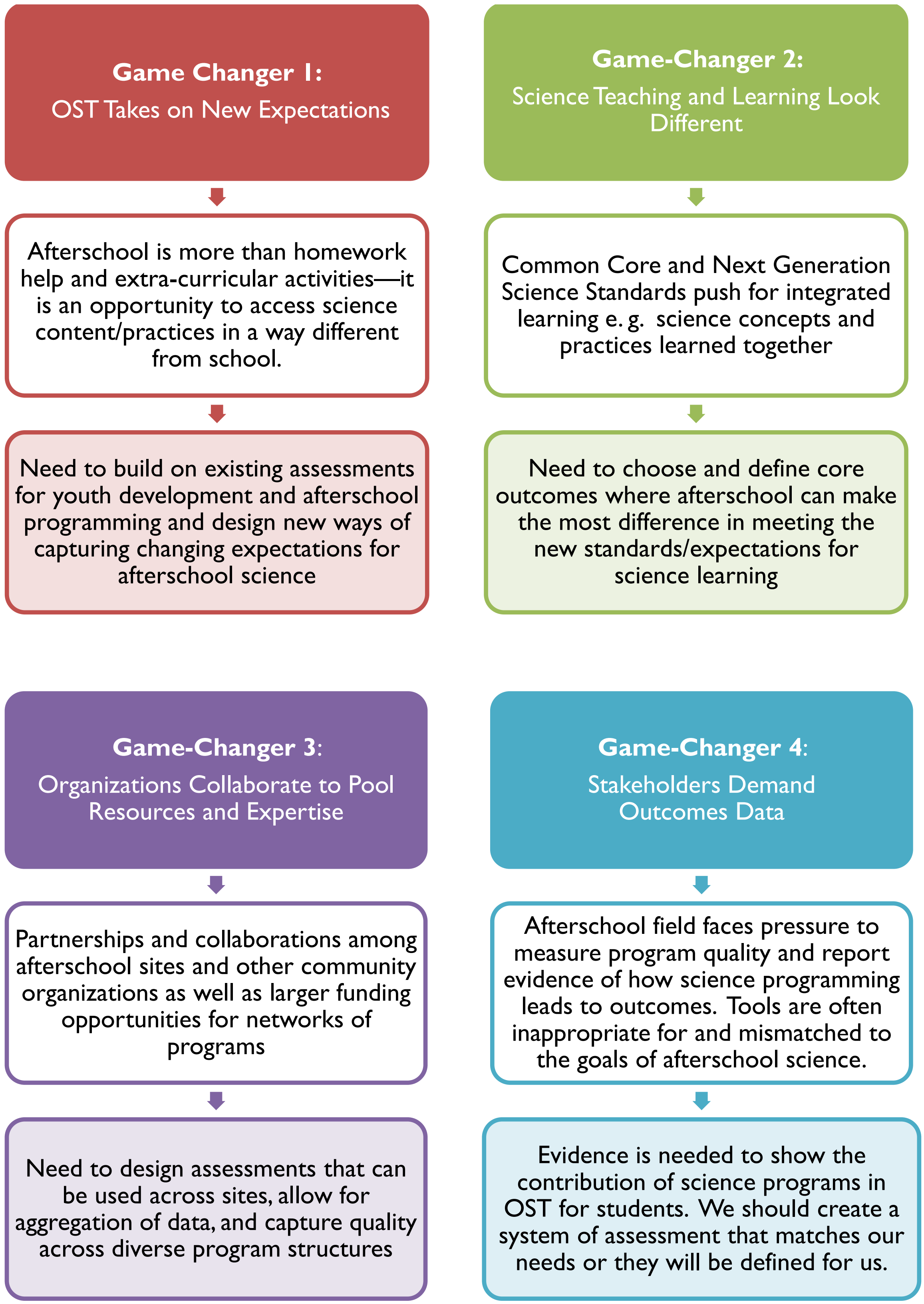
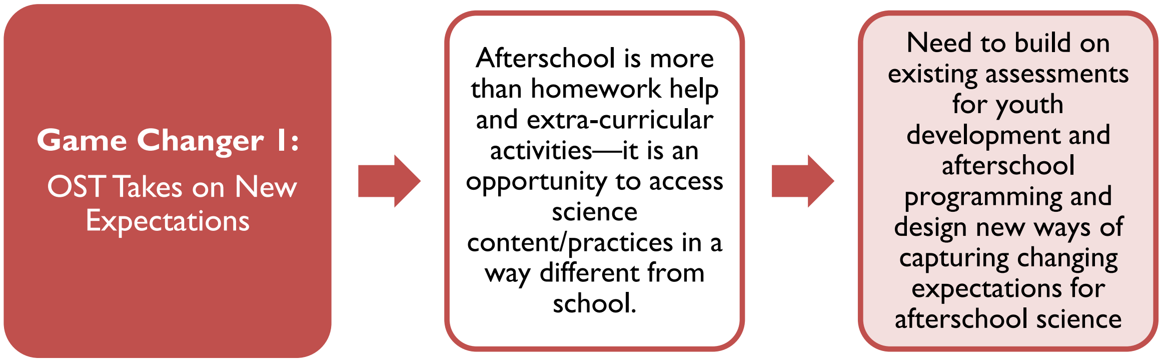


Figure 5: *How game-changers give rise to new assessment designs in afterschool science*

DESIGNING ASSESSMENTS, SHAPING THE FIELD



Given the game-changers, the afterschool science field must:

- Build on existing assessments
- Come to some consensus about what outcomes are manageable and appropriate
- Capture new standards about what students should know and be able to do
- Create assessment systems that allow reporting of outcomes across program types

Assessment presents complicated challenges in the world of OST, which is traditionally not constrained by the testing expectations in schools. We want to know the impact of STEM experiences in informal settings, and the details of program content, structure, and delivery that will lead to the outcomes we seek. Yet, we do not want to end up with informal STEM programs that look too much like school. Right now, the great advantage of OST venues is that they engage children and youth in enjoyable, hands-on settings that build participants' interest in STEM, free of the testing constraints of the formal classroom.

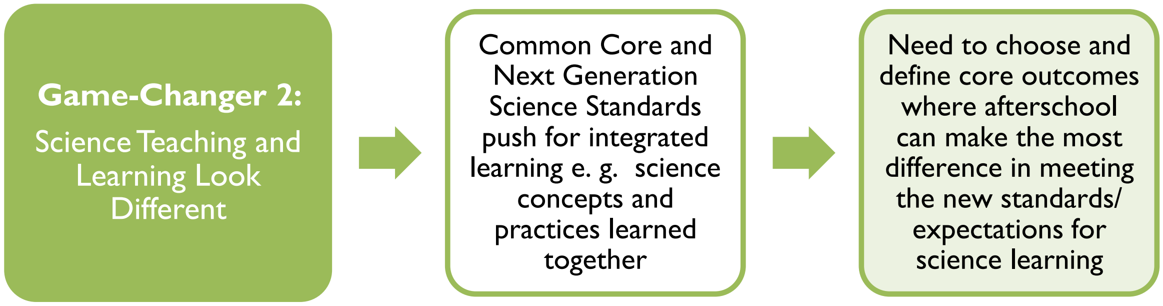
We can learn from the existing work in the field, capturing what occurs in unique afterschool settings (Piha & Newhouse, 2011). Self-reporting tools such as student and teacher surveys can describe participants' experiences with OST programming. For example, student surveys allow youth to report their level of interest, engagement, or motivation for science. Students rate how much they agree or disagree with statements on a 4-point or 5-point scale. Sample items include:

- "Science lessons are fun." (TOSRA)
- "I would prefer to find out why something happens by doing an experiment rather than being told." (TOSRA)
- "When I grow up and have kids, I will take them to a science museum." (Common Instrument)
- "Science is something I get excited about."(Common Instrument)

Observation tools, on the other hand, allow observers to look closely at interactions among teachers and students as activities take place. Tools include the *Youth Program Quality Assessment* (YPQA), *Program Quality Observation Scale* (PQO), and *Assessment of Afterschool Program Practices Tool* (APT). In his white paper, Drew Gitomer points out: "protocols can range from those that ask observers to identify fairly discrete actions by a teacher (e.g., The teacher asks open-ended questions during the lesson) to those that require relatively high levels of inference that take into account not just the teacher action, but also consideration of evidence from students (e.g., The teacher asks questions that promote student thinking and reasoning)," (2012, WP, p.4). The *Dimensions of Success (DoS)* observation tool, for example, allows observers to collect evidence about the unique features of science learning experiences in afterschool.

The website ATIS (http://www.pearweb.org/atis) offers descriptions of a range of tools available in the field along with their psychometric properties and references. We can also build on the insightful work of the *Science Learning Activation Lab (http://activationlab.org/)* and the *Citizen Science Toolkit (http://www.birds.cornell.edu/citscitoolkit)* to inform the development of tools. Even though several of the tools claim that they assess "engagement" or "student attitudes" or "student understanding of science," the field must continue to study and define these constructs as well as develop appropriate instruments to measure them. As new assessments are designed, developers must be extremely careful not to create a variation of the existing top down model, where national assessments made by those outside of the field are imposed upon the afterschool world. Instead, this should be an opportunity to work with programs and intermediaries to create assessments that are appropriate for afterschool settings, and to create a sustainable and effective system of technical assistance for data collection, management, and reporting.

CHOOSING CORE OUTCOMES FOR AFTERSCHOOL SCIENCE



While we have described differences in structure between schools and informal settings, it is important to note that at the core, they do "…share a common interest in enriching the scientific knowledge, interest, and capacity of students and the broader public," (NRC, 2009, p.13-14). By nature, school and afterschool science emphasize unique outcomes and define particular outcomes in their distinctive ways. It is important to focus in on the areas where informal and afterschool science can make the greatest difference and assess those over time, rather than measure many outcomes poorly. As Larry Suter pointed out, several international data collection efforts have items that look at cognitive outcomes but also include separate measures of student goals, values, and interests toward science (2012, WP, p.8). We can learn from those large-scale efforts to achieve agreement across different frameworks and definitions of learning across many theories. The recent Delphi study in the U.S. yielded consensus about three major outcomes for STEM in afterschool as well as indicators and sub-indicators of those outcomes (Afterschool Alliance, 2013, p.6). These three outcomes included:

- Youth develop interest in STEM and STEM learning activities
- Youth develop capacities to productively engage in STEM learning activities
- Youth come to value the goals of STEM and STEM learning activities

Many conversations at the Summit focused on the first two types of outcomes—those that focus on student engagement and science content knowledge and practices.

It is clear that not every program has the goal to meet all outcomes, and this was a significant concern at the Summit. A great number of informal afterschool science programs see their primary charge as increasing students' interest in, curiosity about, and engagement with science. Others argue that the purpose of afterschool science is to contribute to students' science learning and consequently, their performance in school and success along the STEM career trajectory. Both of these outcomes complement and are dependent on each other. For example, studies from the formal education world suggest that lack of engagement has an adverse effect on student performance (Lee & Shute, 2010). The informal field could potentially capitalize on these findings and show how informal experiences boost engagement in science across contexts (e.g., school vs. afterschool).

In the end, we want today's youth to be excited, interested, and motivated to learn science for life—as STEM professionals or informed citizens. To reach that goal, they must have access to purposeful, enriching, and meaningful opportunities to learn important science content and engage in practices.

If the afterschool field can appropriately measure and show positive results on these outcomes, policy-makers will have more reason to support afterschool science programming. Also, increasing opportunities for engagement with science is critical in afterschool settings before students have to make coursework and major decisions in college or higher education. For example, adolescents' experiences with science and expectations about the type of career they would like to pursue have been found to predict the likelihood of taking more science courses and pursuing a science-related career later in their lives (Tai, Liu, Maltese, & Fan, 2006). Without being engaged, students are unable to learn the content and perform science-related activities.

Outcomes for Student Engagement

There is a large body of literature about the construct of student engagement, with a range of definitions and types. There is evidence that academic performance is related to student engagement in school, which can include behavioral, cognitive, and affective subcomponents (Fredericks, Blumenfeld & Paris, 2004). Ellen Skinner and Michael Belmont described student engagement as "…sustained behavioral involvement in learning activities accompanied by positive emotional tone" (1993, p.572). They go on to describe the opposite, as students who "…are passive, do not try hard, and give up easily in the face of challenge." (Skinner & Belmont, 1993, p.572). In her white paper, Ann Renninger argues that while the terms engagement, interest, curiosity, and motivation all have their own literature bases, you cannot define them without referring to the others (2012, WP, p.1-2). She offers the following simple definitions:

- **Engagement** refers to connecting for some period of time to any of a variety of tasks or activities.
- **Interest** refers to both the state of being engaged with and also the predisposition to return to engagement with particular content (e.g., science).
- **Curiosity** describes a disposition to explore and question.
- **Motivation** in its most general usage refers to the will to engage.

For the sake of this report, when we discuss "engagement," we are referring to a combination of these terms. Because afterschool science learning experiences can be approached more creatively than those during the more constrained school day, they are primed to provide students with ample opportunities to engage with science intellectually and practically. This is critical, as several longitudinal studies have linked student engagement with academic achievement (i.e., Alexander et al., 1993; Fincham et al., 1989). As engagement in science decreases in schools that emphasize memorization and test taking over hands-on exploration, afterschool science programs have an important role to play in remedying this.

In summary, engagement with science represents a natural fit for afterschool settings (Noam & Shah, in press). Engagement can last a lifetime, leading to a deep interest in high school and college science courses and perhaps pursuit of a science career. It is the job of the afterschool science field to make sure policy makers understand the great significance of engagement in science and to find measurement strategies that show real results in this domain. Longitudinal work will have to further demonstrate that engagement in children and youth can have a deep impact on later career choices and success in STEM professions.

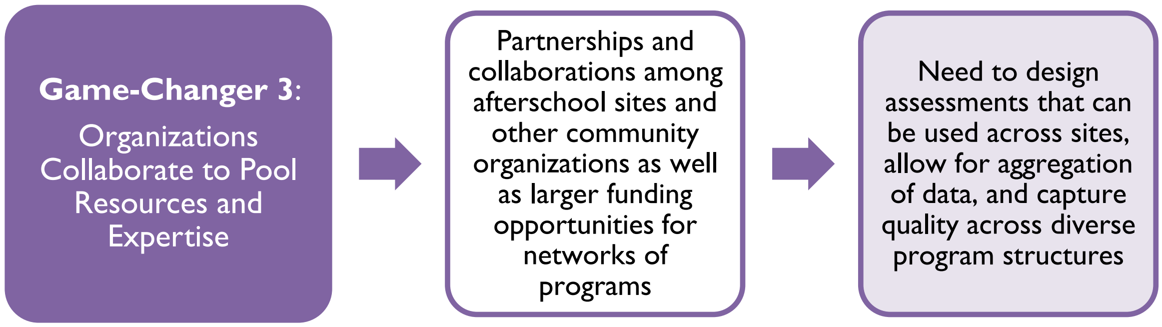
Outcomes for Learning Science Content and Practices

While content knowledge is often the focus of teaching and assessment in schools, content learning also happens in OST. Similarly, engaging students in scientific practices occurs in both environments, however OST settings often provide students with the space to explore these practices more fully. For example, conducting scientific investigations has features of an informal activity: students are doing a hands-on activity—often collaborating with others— communicating their findings, and making decisions about what they investigate and how. Afterschool and summer programs can tackle a wide range of science content areas—drawing on physics, aerospace, earth science, chemistry, and biology as well as engineering design tasks and technology challenges. Therefore, it is difficult to create assessments that focus on particular content outcomes. Instead, it might be useful to look at particular scientific practices and how students use them in different content contexts—skills such as asking scientific questions, developing and using models, planning and carrying out investigations, identifying patterns in evidence, and building explanations. Science activities also require skills listed in the Common Core such as problem solving, communication, and collaboration.

As Common Core is adopted by states, it will set the stage for collecting data on these skills across settings. Practices in NGSS such as learning to collect data to answer a question or mastering how to build an argument with evidence can be transferred across content areas. For example, demonstrating the ability to examine data and propose scientific explanations can occur whether a student is collecting data about plant growth under different conditions or tracking the rate of a chemical reaction in various beakers. In this way, content and science practices are learned in an intertwined way. At this level it is possible to assess skills, knowledge, and practices across content areas and across settings.

Many policymakers and funders do not support a singular focus on engagement in science (Ron Ottinger, personal communication, February 2013). There are two paths to pursue in response. One is to convince the skeptics that this is a legitimate and important outcome and that programs should focus on it. The other, which many experts at the summit proposed and the authors of this report agree with, is to help programs strengthen their offerings through professional development, technical support, online resources, training/coaching and other quality supports. This is the major frontier at present—defining the learning goals for informal science in afterschool (e.g., Afterschool Alliance, 2013), and then finding ways to demonstrate that these goals are being reached. Clearly, the first two game-changers (new expectations and standards) suggest that new assessments will be needed to move beyond measuring static knowledge, and instead to capture the use of knowledge through practices and real-world applications.

NEED FOR COMMON ASSESSMENTS



Currently, many afterschool and summer science programs use homegrown surveys, observation tools, or written assessments to monitor the quality of their programs or student progress on particular outcomes (Dahlgren, Noam, and Larson, 2008). In contrast schools, often use common tools, like state assessments or national exams, to assess student achievement in particular subject areas. We encourage the field to do the same, for example, by creating a survey that measures student engagement with science across afterschool settings. Another strategy would be to provide searchable databases or websites (e.g., ATIS) where programs access a set of assessment tools that have an established psychometric history. Finally, some programs may choose to develop their own instruments. In these cases, we recommend that guidelines be created for sharing how they've developed their tools and the constructs measured by them, so others can understand what they are assessing.

The Case for Common Assessments

An important part of the Summit involved discussions about the feasibility of common assessments that offer multiple methods (self-reports, observation protocols, multiple-choice questions) of assessing inputs and outcomes. Arguments **for** common instruments included:

- ***Encouraging a common language.*** Terms such as "engagement," or "science identity" are often used in the informal and afterschool worlds but are defined in different ways by experts.
- ***Creating high-quality assessments.*** With multiple settings collecting data about the use of particular assessments to measure specific outcomes, a common assessment instrument will produce more evidence for the strengths and weaknesses of the assessment so that continuing improvements can be made. Thus, instead of having many weak assessments that are created independently by a range of sites, programs can pool resources and data and evaluators and researchers can create stronger assessments.
- ***Pinpointing funding and policy avenues.*** If several afterschool programs can show that students' participation in their informal science activities increased a particular outcome, then there may be more support for that type of programming. If just a single afterschool program focuses on measuring a very specific outcome only for its site, it is harder to make a general case for informal science funding leading to support that particular outcome.

Challenges to Designing Common Assessments

As mentioned above, OST programs offered in a range of settings such as museums, science centers, community organizations, and afterschool environments differ so much that creating common assessments offer key challenges:

- ***Watering down outcomes.*** Because the outcomes have to be general enough for data to be measured across settings, we have to simplify and generalize to a point that we can lose the core of what we are trying to measure in any given program.
- ***Preserving the spirit of informal science experiences.*** Even assessments that focus on interest and engagement but are delivered in conventional ways— surveys that kids need to sit down and answer, question by question— risk dampening enthusiasm. Kids do not like to take these types of tests in out of school time. There is much work to do in the field to develop assessments that are appropriate for informal STEM venues, find the best delivery methods, and develop ways to aggregate results so they can be used as powerful evidence in policy and public awareness efforts. (Traphagen, Sneider, & Morrison, 2012, p.5)
- ***Embracing unique program features.*** Whether it is Lego Robotics, a summer camp on the ocean, or a series of Saturday workshops at the science center, informal science providers take pride in their unique resources, partnerships, and offerings. Unfortunately, using common instruments may lead to programs becoming homogenous — a variant on the idea of "teaching to the test." What will be lost is the rich variety of educational environments that make afterschool experiences so appealing to such a wide range of learners.
- ***The reality of introducing assessments.*** As we consider common assessments and implementation, we should keep in mind that the staffs in afterschool programs often have limited science training and little exposure to assessments. The more tailored the assessments are to their practices, the greater their motivation to implement assessments will likely be.

While individual programs may choose to also study specific outcomes that are unique to them (e.g., measuring students' attitudes towards conserving aquatic wildlife in a marine biology afterschool program), having some common outcomes that all informal and afterschool science programs can measure can help build the evidence base for the field.

EVIDENCE AT DIFFERENT LEVELS

Game-Changer 4: Stakeholders Demand Outcomes Data

The afterschool field faces pressure to measure program quality and report evidence of how science programming leads to outcomes using tools developed for the school world. These tools are often inappropriate and mismatched to the goals of afterschool science.

Evidence is needed to show afterschool science is worth it. We should create assessments that match our needs or they will be defined for us.

In order to create a suite of assessments that fully capture the inner workings of afterschool science programming, it is important to first ask what we need these assessments for? While all afterschool networks or organizations may not have the exact same needs, many are working to respond to the following common pressures:

- Convincing policymakers, funders, and other stakeholders that these experiences add value to the overall science education of our youth
- Assuring program quality as diverse science afterschool and summer programs are offered across the country
- Providing feedback to program leaders and staff on how to improve science instruction
- Tracking contributions of and links between in-school and out-of-school time experiences
- Evaluating programs and student outcomes for those who participate short-term and long-term
- Collecting evidence to make informed decisions about which programs are worth scaling up and which are not

With these varied goals, assessments need to be designed at different levels to address different purposes. At the Summit, participants discussed assessments that quickly measure outcomes at a glance, and then more in-depth ones that capture the nuances of particular interventions (see Appendix 2 for examples discussed at the Summit). Based on these discussions, we propose two levels of assessment development. The first level includes tools that are quick to administer and easily operationalized across programs. The second level includes tools that help measure the quality of programs, looking at characteristics of their activities, staff, and resources that make them more or less successful at engaging students in science.

Level 1: Surface Structure—Fast and Large Scale

At this level, tools need to be developed to assess a core set of outcomes across many programs. Data could be aggregated across settings and the evidence base for that outcome would be large enough to show the impact of afterschool science experiences. Self-report surveys are an efficient way to gather data. For example, for student engagement and interest, the self-report tool under development by PEAR (sponsored by the Noyce Foundation), called the Common Instrument, could be given to students before and after science learning experiences in a range of settings. If participation in particular types of activities were linked to increased engagement/interest in science, the evidence would be valuable for funders and other stakeholders.

Level 2: Deep Structure —Quality and Processes

At this deeper level, tools need to be developed to link program quality indicators with student outcomes. Level 2 assessments would go beyond general baseline data about trends across programs, and instead, capture interactions among students, facilitators, and resources in informal and afterschool settings. Detailed observation tools as well as new technology that capture interactions could be helpful. While validating instruments can be a long, resource-intensive process, by pooling resources and tackling the problem collaboratively, stronger validity arguments can be built in the field.

Since programs are required by funders and organizations like the National Science Foundation (NSF) to document their processes, it would be useful to strengthen the suite of tools they could choose from when planning their evaluation efforts. Different teams could work on a range of existing and new approaches: performance-based, stealth, embedded, online, and real-time assessments, among others. There are tools that already exist (see www.pearweb.org/atis and informalscience.org) and can be refined, as well as many others to be developed or validated. At this level, there is more opportunity for in-depth observations of students completing embedded tasks, interviews with students and teachers, and analyses of portfolios that can track progress over time.

We suggest starting with these two levels, however, over time, there may be other levels, such as in-depth case studies as well as longitudinal research to trace how particular informal and afterschool science experiences influenced students' learning trajectories and career pathways. There might also be opportunities for studies using different assessment designs. Focusing on two levels initially is manageable and a good way to make progress and contribute to the field in responsible, thoughtful ways.

Measuring Outcomes Over Time

Measuring student outcomes after participating in a science program is informative, but building an identity as a science learner and deciding to move through the science pipeline are things that develop over time. How do we track students' experiences with science over time? Summit participants highlighted the importance of following students along their learning trajectories — as they are introduced to ideas in school, through afterschool experiences, summer camps, and/or career guidance in high school. Going to an afterschool program once a week, for example, may not add up to the hours spent in school. However, tracking the cumulative impact of various afterschool science experiences over time may reveal that students are being affected more profoundly and in different ways than by school science. Students progress through their understandings and growth developmentally, and assessments of their science learning experiences should also track the evolution of their science interest, knowledge, identity, and career pursuits.

Summit participants agreed that longitudinal studies are very difficult, time-consuming, and costly, as students move from school to school and in some cases jump from one type of afterschool program to another. However, they also emphasized that they are important and that the field must think creatively about innovative technologies to collect longitudinal data. For example, students could have an online portal for logging their science education experiences. They could track what they did and when (in school/out of school, science club, textbook learning, etc.) and then that information could be analyzed over time while looking at particular outcomes. Also, new technologies could be developed to help monitor and track students' participation in science activities—for example, the nature of their contributions, whether or not they are engaged the entire time, etc. Finally, programs could upload information to a central database so that data could be aggregated across student experiences in different programs over time.

The question of " change attribution" also arose in Summit discussions. That is, when you are looking at a student's development and experiences over time, how do you know what weight to attribute to what experience? Do we attribute it to the Robotics Club that meets afterschool once a week, or to the amazing fifth grade science teacher? How do we know that improvements in certain science learning outcomes can be attributed to particular school or afterschool experiences? Careful attention must be paid to how data is used to build an evidence base for particular claims.

Afterschool environments play an interesting role in the educational landscape: on one hand, they offer a fun, exciting, safe space for students to have more choices and opportunities for enrichment. On the other hand, they offer extended time to reach students who may not connect with subjects like science during the very short time allotted during the school day. Some settings offer science in chaotic, superficial ways. Other programs introduce science in rigorous, content-heavy ways to align with school goals. Still others find a balance, where they engage and excite students in science while exposing them to important content and practices. Creating assessments that match this diverse afterschool world is itself, a game-changer. It can alter the way afterschool science is designed, experienced, funded, and connected to the school. As a game-changer, assessment in afterschool science allows for sophisticated questions and answers over the course of a student's educational experience.

CONCERNS, QUESTIONS, AND SUGGESTIONS

We anticipate several logistical and practical roadblocks in the journey ahead. Below, we raise our top concerns:

- ***Funding:*** How will we fund efforts to design and study new assessments? How can we pool resources across public and private sectors and gain the interest and sustained focus of various stakeholders?
- ***Expertise at the Program Level:*** At the program level, implementing new assessments and collecting data that can be aggregated beyond each individual program is time-consuming and detail-oriented. Given the many responsibilities of program staff already, how can we increase buy-in through training and technical assistance so that programs can implement tools/assessments and provide critical feedback?
- ***Data Management Infrastructure:*** With so many different programs across the country, how will data be managed and analyzed? What type of organizations will be needed to monitor this process and help relay the data to the appropriate researchers and assessment design teams? How will data be reported back to programs in user-friendly and efficient ways?
- ***Differing Motivations and Needs:*** While many would agree with the notion that the afterschool science field needs quality assessments that match its unique approach beyond school measures, the motivations and needs of stakeholders differ. Therefore, finding ways to come to consensus and make rapid progress forward will be critical.

These concerns and others were raised at the Summit and since. Interestingly, despite the range of expertise, roles in the formal and afterschool fields, and motivations, attendees felt that these big issues can be solved through strategic investments and a focused action plan. The following suggestions were made regarding the work ahead:

1. ***Leverage the expertise in the field.*** Researchers should collaborate on the creation of assessment tools, focusing on their areas of expertise — whether that is creating a theory of action, defining outcomes or tools, or pilot testing tools. This may involve pairing engagement experts with specialists in observation tool design, or STEM school content experts with afterschool researchers.
2. ***Build the evidence base carefully, not hastily.*** Participants agreed that we need to prioritize the key outcomes where we feel informal and afterschool science can make the most difference. By focusing, we can spend time making strong tools and going through enough trials in the field to refine the tool. In her white paper, Kirsten Ellenbogen suggested that the needs of both the field and individual programs or projects must be balanced by "selecting a small number of assessment questions to be used across the field" in order to "support the development of more coordinated data collection across projects, while still accounting for project specific goals and outcomes" (2012, WP, p.9).
3. ***Remember the audience.*** It is critical that we engage the informal science learning community in critiquing the process, developing the assessment tools, and measuring the results. It is important to make sure the assessments are valid and reliable and also that there is buy-in from those who will be using them. They need to be seen as useful, accurate, and worth the effort of administering them.
4. ***Make sure assessment design fits well in the climate created by the game-changers.*** It is important to recognize the forces surrounding the shifts in expectations for afterschool and how evidence and standards will play a role moving forward.

In the last section of this paper, we will introduce a set of next steps, in order to gain momentum and show success.

CALL TO ACTION

The Summit, and many discussions since, prompted diverse participants representing different communities of research, practice, and policy to work together towards measurement strategies in informal and afterschool science learning. Even with different frameworks, theories, and goals, the Summit attendees were able to offer a set of solutions and suggested short-term and long-term steps. As we have emphasized throughout this report, school-based assessments (e.g., NAEP, state science content tests) are inappropriate given the informal nature of the afterschool field. However, we need to consider how we can make immediate progress in developing measurement strategies that will convince policymakers and funders to provide the necessary supports for afterschool science programming.

The Summit showed that there is great potential for moving an assessment agenda forward for afterschool science. As a next step, we propose a convening to define and plan a detailed working agenda and funding process. This planning meeting should be focused on the areas of immediate work and should lead to the evolution of work plans that are measured in six-month intervals. Specifically, we recommend that the planning meeting address the following questions:

a) What assessment tools exist to address the desired outcomes and (sub-)indicators of learning the field can deliver? Where are the gaps?
b) What would comprise a convincing evidence base for policymakers and funders?
c) How can feedback from practitioners be incorporated in an assessment development plan?

Focused research should follow expeditiously and should include, at the least, four domains:

1) Refinement of existing measurements of student engagement and interest
2) Assessment of science content and practices, including embedded assessments and in-depth analysis of learning
3) Assessment of quality in science programming

4) Development of analytic strategies to assess multi-context learning of science (e.g., school, afterschool, summer)

We recommend that funders work closely together as they have in developing the research agenda and the Summit. The four funders were able to forge a common vision for the Summit and it will be important to include new funders and networks, as the investments will have to be significant and sustained. To do so, we must build relationships between school-based research and the afterschool science research base. The linking of common interests has already begun and will hold a great deal of promise, especially with cognitive and affective outcomes being highlighted by the NGSS. While both school and afterschool settings work to address both sets of outcomes, each setting is designed and equipped to prioritize particular outcomes over others (Sneider and Coffey, 2013). Thus, we can expect opportunities for strong links between both worlds as they capitalize on each other's strengths to improve the overall science learning opportunities for youth.

Participants introduced the concept of creating a "center" several times during the Summit and beyond. A center offers a structure to coordinate the sharing of resources and ideas in a way that will avoid unnecessary competition, repetition, and fragmentation. There are, of course, many different models for centers. Some are loose networks of existing research groups, while others are university-based, providing core coordination of a team of researchers. A specific plan for what kind of center is needed was not discussed at the Summit, but needs to be addressed in the near future to increase the likelihood of success. At the least, a data and reporting center is needed to integrate and coordinate data across programs, cities, and states.

In summary, when pursuing the assessment agenda, we do not want assessments to control the destiny of the afterschool science world. We would risk compromising the many exciting opportunities that exist for students to be intrigued by the world around them, to make discoveries, to collaborate, to meet mentors, and to feel empowered by science in the diverse and rich world of afterschool science. The possibilities are endless, but the challenges and roadblocks are real as we face an assessment predicament in afterschool science. Just because something sounds good, does not mean people will believe it or fund it. As a field, we must embrace this great opportunity and create assessments that can best measure what makes afterschool science important, special, and worth investing in for the benefit of today's youth. We must clarify for all stakeholders what we intend to do, how we plan to do it, and how we will know we have done it. The game has changed… and we have must decide if we are ready to play—collaboratively—in a way that bridges the research, policy, and practice communities.

REFERENCES

Afterschool Alliance (2013). Defining youth outcomes for STEM learning in afterschool. Retrieved from http://www.afterschoolalliance.org/STEM_Outcomes_2013.pdf

Alexander, K. L., Dauber, S. L. & Entwisle. (1993). First-grade classroom behavior:

Its short- and long-term consequences for school performance. *Child Development*, 64, 801–814.

Anderson, R. D., & Helms, J. V. (2001). The ideal of standards and the reality of schools: needed research. *Journal of Research in Science Teaching, 38*, 3-16.

Boiselle, E. Hussar, K. Noam, G. & Schwartz, S. (2008). *Toward a Systematic Evidence-Base for Science in Out-of-School Time: The Role of Assessment*. Cambridge, MA: Harvard University & PEAR.

Change the Equation. *Lost Opportunity: Few U.S. Students Participate in STEM Out-of-School Programs.* Vital Signs: Reports on the condition of STEM learning in the U.S. Retrieved August 2013 at http://changetheequation.org/sites/default/fines:CTEg%20Vital%20Signs%20Lost%20Opportunity.pdf

Crawford, B. A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of Research in Science Teaching, 44*(4), 613-642.

Dahlgren, C. T., Noam, G. G. & Larson, J. D. (2008). Findings for year one data for the Informal Learning and Science Afterschool Study. Paper presented at the annual meeting of the American Educational Research Association, New York: NY.

Davis, E. A., Petish, D., & Smithey, J. (2006). Challenges new science teachers face. *Review of Educational Research, 76*(4), 607-651.

Davis, K. S. (2002). "Change is hard": What science teachers are telling us about reform and teacher learning of innovative practices. *Science Education, 87*(1), 3-30.

Fincham, F. D., Hokoda, A., & Sanders, R. (1989). Learned helplessness, test anxiety, and academic achievement: A longitudinal analysis. *Child Development*, 60, 138–145.

The Forum for Youth Investment. (2012). The common core standards: What do they mean for out-of-school time?

Fredericks, J. A., Blumenfeld, P. C. & Paris, A. H. (2004). School Engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74, 59-109.

King, G., Murray, C., Salomon, J. & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review,* 94 [February]: 191–205

Lee, J. & Shute, V. (2010). Personal and Social-Contextual Factors in K-12 Academic Performance: An Integrative Perspective on Student Learning. *Educational Psychologist.* 45(3), 185-202.

Malone, H.J. (2011). Expanded Learning Time and Opportunities. *New Directions for Youth Development: Theory, Practice, Research.*

National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, D.C.: National Academy Press.

National Research Council. (2000). *How People Learn: Brain, Mind, Experience, and School: Expanded Edition*. Washington, DC: The National Academies Press.

National Research Council (NRC) (2009). *Learning Science in Informal Environments: People, Places, and Pursuits.*Committee on Learning Science in Informal Environments. Philip Bell, Bruce Lewenstein, Andrew W. Shouse, and Michael A. Feder, Editors. Board on Science Education, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council (2012). *A Framework for K-12 Science Education Practices, Crosscutting Concepts, and Core Ideas.* Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, D.C.: The National Academies Press.

Noam, G.G., Biancarosa, G., & Dechausay, N. (2002). *Afterschool Education: Approaches to an Emerging Field.* Cambridge, Harvard Educational Publishing Group.

Noam, G. G. (2001). *Afterschool Time: Towards a Theory of Collaborations.* Paper written for Urban Seminar Series on Children's Mental Health and Safety: After-School Time. Retrieved April 2013 at http://naesp.org//resources/1/A_New_Day_for_Learning_Resources/Making_the_Case/Afterschool_Time_Toward_a_Theory_of_Collaborations.pdf

Noam, G. G., & Shah, A. (In press). Informal Science and Youth Development: Creating Convergence in Out-of-School Time. Chapter to appear in the National Society for the Study of Education (NSSE) Yearbook 2013.

Piha, S. & Newhouse, C. (2011). A Crosswalk Between the Learning in Afterschool Learning Principles and Afterschool Quality Measurement Tools. Retrieved August 2012 at

http://www.learninginafterschool.org/documents/A%20Crosswalk%20Between%2

Pittman, K. J., Irby, M., Yohalem, N., & Wilson – Ahlstrom, A. (2004). Blurring the lines for learning: The role of out-of-school programs as complements to formal learning. *New Directions for Youth Development.* 101 (Spring), 19 – 4.

Sneider & Coffey. (2013). Draft Title: Realizing the Vision of the Framework By Linking Formal and Informal STEM Education.

Skinner, E.A., & Belmont, M.J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85(4): 571-581.

Tai, R.H., Liu, Q.C., Maltese, A.V., & Fan, X. (2006). Planning early for careers in science. *Science Magazine*. 312(5777), 1143 – 1144.

Traphagen, K., Sneider, C., & Morrison, J. (2012). Draft Title: The Intersection of Informal STEM Education and NGSS.

Online links to websites cited in paper:

http://www2.ed.gov/programs/21stcclc/index.html

http://www.nextgenscience.org/

http://stem.afterschoolnetwork.org/rfq

APPENDIX I

Characteristics of Learning Environments for Science Learning

Settings	Time	Structure	Pedagogy	Goals or Intended Outcomes of this type of Programming
Schools (formal)	School Day	• Mandatory • K-12 Education • Graduation • Standards/Assessments • Structured and intentional time	Didactic or Hands-on, inquiry	• Content learning • Passing grade level • Graduation
Afterschool and Summer program settings	Out-of-school time	• Structured and intentional learning opportunities • Flexibility with time (can spend more time on science than in school day) • Could take place in schools, but with a non-formal feel; could take place in designed informal settings (museums, etc.) but have structured program	Didactic or Hands-on, inquiry	• Engagement • Motivation • Interest • Excitement • Attitude towards science • Exposure to content/practices that can support school learning **Some afterschool/summer programs are geared towards school performance, academic advancement (college prep courses), etc. They are beyond the scope of our focus.*
Museums Science centers, Zoos Botanical gardens, etc.	Out-of-School Time	• Visit with parents/friends as you wish • Free Choice • May not be intentional	Didactic (tour/ lecture/ presentation) or Hands-on, inquiry	• Interest • Motivation • Engagement • Excitement • Attitude towards science
Everyday life	Out-of-school time	• Helping parent fix car • Cooking with friends • Observing birds in backyard, etc. • May or may not be intentional	Self-directed Spontaneous	• Appreciating science • Value of learning/exploring

Appendix 2

Assessment Design Possibilities

Assessments for measuring science content in afterschool settings should be different than those used in traditional academic environments. First, they should be embedded in the afterschool activity and not feel like an add-on that is disconnected from the learning experience. Traditional paper-pencil tests do not fit the exploratory, student-centered nature of afterschool science and can be detrimental to programs' youth development focus. One solution is to use embedded assessments that do not take away from the hands-on, exploratory nature of the experience, but in fact, contribute to the learning. Embedded assessments can maintain activity flow and provide in-the-moment opportunities for reflection and feedback. In this way, they are not seen as something separate, additional, or evaluative, but rather as part of the learning process.

For example, if students are playing a video game or doing an online simulation, stealth assessments could be a part of that experience. In his white paper, Zapata-Riviera stated: "Stealth assessments are unobtrusive embedded assessments that are woven directly and invisibly into the fabric of the learning or gaming environment. During video game play, students naturally produce rich sequences of actions while performing complex tasks, drawing on the very skills or competencies that we want to assess," (2012, WP, p.2). Another advantage of embedded assessments is that they can provide feedback to students while they are doing the activity, so they can monitor their progress and refine their strategies.

Finally, when designing common assessments and attempting to collect data from many students at various afterschool science sites, it is often useful to give surveys that use simple multiple choice or Likert Scale items, where respondents pick a response on a scale and circling it. However, often it is hard to tell what a response to a survey question really tells us:

Question: Do you enjoy learning about science in school?

☐ 1 Strongly Disagree

☑ 2 Disagree

☐ 3 Neutral

☐ 4 Agree

☐ 5 Strongly Agree

Does the response of "2" mean the student does not enjoy science as a subject, does not like the particular way she learns science in school, or does not like her teacher? It is unclear what is motivating that response, and it can be difficult to use liking or not liking something as an indicator of overall excitement, interest, and engagement in a large subject area like science. On the other hand, if data were collected from students who attend a community-based science club once a week and students who are exposed to science only in school, it would be meaningful if those in the afterschool experience consistently rated their attitude towards science as higher on different items capturing that construct. Therefore, the data may be less powerful for individual students than for the aggregate. Summit attendees discussed some possible new approaches for improving assessments as development moves forward. These discussions were sparked by a presentation by Patrick C. Kyllonen, who shared examples of promising techniques including *forced-choice* design, *anchoring vignettes,* (King, et al., 2004), and *behaviorally anchored rating scales*. As we try to create the best assessments to match the afterschool science world, there will be exciting and meaningful possibilities for collaboration and innovation.

WHITE PAPERS

The following white papers were developed to spark thinking before the Summit and may be accessed online at:
http://www7.nationalacademies.org/bose/Assessment_Informal_Ed_Summit_Commissioned_Papers.html

In the report, the white paper will be cited in the text with this format:
(Author Last Name, 2012, WP, p.X)

Informal Science Learning and Education: Definition and Goals
Krishnamurthi, A. & Rennie, L.

The NRC and NSF Frameworks for Characterizing Learning in Informal Settings: Comparisons and Possibilities for Integration
Allen, S. & Bonney, R.

Informal Science Education Assessment in the Context of the 5-Year Federal STEM Education Strategic Plan
Feder, M. & Weiman, C.

Attitudes of Stakeholders towards Assessment in the Informal Science Education Realm
Friedman, A.

Evaluation Under Pressure: Balancing the Needs of the ISE Field with the Needs of Individual Projects
Ellenbogen, K.

Observational Methods for Assessment of Informal Science Learning and Education
Gitomer, D.

Using the NRC Framework to Engage Students in Learning Science in Informal Environments
Krajcik, J.

On Defining and Assessing Engagement, Interest, Curiosity, and Motivation in Informal Science Learning
Renninger, K.A.

Creating Assessment Frameworks: Experience from International Studies
Suter, L.

Embedded Assessment of Informal and Afterschool Science Learning
Zapata-Rivera, D.

SUMMIT PARTICIPANT LIST

Those who reviewed early drafts of the report and provided feedback/comments for us to consider are indicated with an asterisk next to the name.

Sue Allen
Allen & Associates

Nicole Ardoin
Stanford University

John Baek
National Oceanic And Atmospheric Administration NOAA

Eva Baker
UCLA/CRESST

Bronwyn Bevan
Exploratorium

George Bo-Linn
Betty and Gordon Moore Foundation

Rick Bonney*
Cornell Lab of Ornithology

Doris Chin
Standford University

Janet Coffey
Betty and Gordon Moore Foundation

Lynn Dierking
Institute for Learning Innovation

Rena Dorph
Lawrence Hall of Science;
University of California, Berkeley

Kirsten Ellenbogen
Science Museum of Minnesota

Stuart Elliott
National Research Council

(Evelyn) Margaret Evans
University of Michigan

John Falk
Oregon State University

Michael Feder
White House

Alan Friedman
FriedmanConsults.com

Drew Gitomer*
Rutgers University

Leslie Goodyear*
EDC

Kris Gutierrez
University of Colorado, Boulder

Edward Haertel
Stanford University

Joseph E Heimlich
Ohio State University; Institute for Learning Innovation

Joan Herman
UCLA/CRESST

Margaret Hilton
National Research Council

Margaret Honey*
New York Hall of Science

Sherry Hsi
Lawrence Hall of Science; University of California, Berkeley

Arron Jiron
S.D. Bechtel, Jr. Foundation

Judy Koenig
National Research Council

Joe Krajcik
Inst. For Research on Math & Science Education, Michigan State University

Anita Krishnamurthi*
Afterschool Alliance

Amy Kurpius
SK Partners, LLC

Patrick Kyllonen
Educational Testing Service

Christine Massey
Univ. of Pennsylvania, Institute for Research in Cognitive Science

Catherine McEver
Bureau of Common Sense

Barbara Means
SRI International

Gil Noam
Harvard University

Ron Ottinger
Noyce Foundation

K Ann Renninger
Swarthmore College

Sherman Rosenfeld
Weizmann Institute of Science, Rohovot, Israel

Marsha Semmel
Institute of Museum and Library Services

Ashima Shah
Program in Educational, Afterschool, and Resiliency

Richard Shavelson
SK Partners, LLC

Patrick Shields
SRI International

Cary Sneider
Portland State University

Gerald Solomon
Samueli Foundation

Brian Stecher
RAND

Martin Storksdieck*
National Research Council

Larry Suter*
University of Michigan

Robert Tai*
University of Virginia

Carol Tang
Coalition for Science After Achool

Uri Treisman
Dana Center UT Austin

David Ucko
Museums + more

David Uttal
Northwestern University

Deborah Vandell
University of California, Irvine

Soo Venkatesan
S.D. Bechtel, Jr. Foundation

Diego Zapata-Rivera
Educational Testing Service