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.

4

Equity and Evaluation in Informal STEM Education

Cecilia Garibay, Rebecca M. Teasdale

Abstract

Informal STEM education institutions seek to engage broader cross sections of their communities to address inequities in STEM participation and remain relevant in a multicultural society. In this chapter, we advance the role that evaluation can play in helping the field adopt more inclusive practices and achieve greater equity than at present through evaluation that addresses sociopolitical contexts and reflects the perspectives and values of non-dominant communities. To do this for specific projects, we argue that evaluation should privilege the voices and lived experiences of non-dominant communities, engage communities in identifying desired outcomes, and ensure multicultural validity of instruments, measures, and inferences. At the field-wide level, we urge evaluators to examine conceptualizations of "broadening participation in STEM," evaluate community-based partnerships, and address replication and scaling. Ultimately, these actions can lead to greater equity in evaluation and in informal STEM education.

> *. . . the reproduction of social disadvantage in [informal science education (ISE)] has less to do with purposefully exclusive practices on the parts of ISE institutions and their representatives, than with the ingrained values, systems and behaviours of ISE practitioners, their visitors and their "non-visitors", as well as society more broadly.* (Dawson, 2017, p. 221)

> *Those who attempt to study and advance knowledge of multicultural and culturally competent evaluation inevitably find that the central issue is to move beyond narrow culture-bound assumptions toward diverse sociocultural perspectives and experiences.* (Hopson, 2003, p. 1)

Evaluation of informal science, technology, engineering, and mathematics (STEM) education programs and exhibitions is fundamentally tied to issues of equity. As the above quotations illustrate, informal STEM education (ISE) and evaluation are rooted in values and assumptions that reflect cultural, social, and historical forces that shape professional and institutional practices. In ISE, these practices privilege the perspectives, concerns, and preferences of dominant groups and disadvantage non-dominant groups (Ash & Lombana, 2013; Dawson, 2014a; Feinstein, 2017; Philip & Azevedo, 2017). Likewise, evaluation practices reflect the larger power and social structures in society (House & Howe, 1999; Mertens, 1999); they are often grounded in the dominant culture worldview and values and marginalize non-dominant communities (Hopson, 2003; Stanfield, 1999; Thomas & Madison, 2010).

This chapter explores the relationship between equity and evaluation in the context of ISE in the United States, focusing specifically on this issue's theme of summative evaluation of ISE outcomes. We advance the role that evaluation can play in the future to help the field enact more inclusive practices and achieve greater equity by assessing outcomes in ways that explicitly consider sociopolitical contexts as well as the values, practices, and lived experiences of non-dominant communities. We situate our arguments in two bodies of literature: the emerging discussion about equity in ISE and the flourishing literature on foregrounding culture and context in evaluation.

Equity and Informal STEM Education

Though STEM is increasingly important for daily life and work, communities of color have been traditionally underrepresented in STEM careers and inadequately served in STEM education (Dawson, 2014a; Herrenkohl & Bevan, 2017; U.S. Bureau of Labor Statistics, 2017). Engagement in ISE, for example, is dominated by privileged groups (Ash & Lombana, 2013; Dawson, 2014a, 2017; Feinstein & Meshoulam, 2014), and ISE institutions actively seek to engage broader cross sections of communities to help reduce educational inequities and remain relevant in an increasingly multicultural society (Farrell & Medvedeva, 2010; Feinstein, 2017). This positions ISE institutions as key actors in national efforts to broaden participation in STEM (Bell, Lewenstein, Shouse, & Feder, 2009; Center for the Advancement of Informal Science Education [CAISE], 2018; Committee on Equal Opportunities in Science and Engineering [CEOSE], 2017). Most often, broadening participation activities are defined as those that encourage full participation

in STEM education and careers by communities of color, White women, and people with disabilities (CEOSE, 2017). To date, however, ISE efforts have largely focused on diversifying ISE participants without addressing the conditions that have created inequities (Dawson, 2014a; Feinstein, 2017; Philip & Azevedo, 2017).

Currently, the ISE field lacks shared terminology and definitions of notions such as *diversity* and *equity* (Feinstein & Meshoulam, 2014). We use the term *diversity* to refer to the ways in which people are similar and different, including but not limited to identities, social locations, lived experiences, and values (Garibay & Huerta Migus, 2014; Nightingale & Sandell, 2012). *Equity* implies fair access to resources (such as education) that advances social justice by allowing for self-determination and full participation in society (Dawson, 2014a; Feinstein, 2017; Philip & Azevedo, 2017). Efforts to achieve diversity in ISE may also advance equity, depending on the extent to which the underlying assumptions and norms that shape experiences are addressed. While ISE is often considered sufficiently self-directed for all individuals to orchestrate rewarding experiences (e.g., Falk, 2001), ISE institutions are fundamentally structured to align with the dominant culture practices of "expected" visitors (Dawson, 2014b; Garibay, Lannes, & González, 2017). Members of non-dominant communities, in contrast, often find their practices and norms are not reflected, leading them to experience ISE as unwelcoming, confusing, and/or irrelevant (Dawson, 2014b; Garibay, 2009; Garibay et al., 2017).

Efforts to diversify ISE participation that focus exclusively on providing greater access, inadvertently locate the "problem" within underserved communities. These efforts aim to help communities overcome limitations in their ability to access, navigate, and/or appreciate traditional ISE practices, reflecting deficit-based, assimilationist, and/or colonialist perspectives (Bell et al., 2009; CAISE, 2018; Dawson, 2014a; Garibay et al., 2017; Philip & Azevedo, 2017). In contrast, equity-focused efforts locate the "problem" within ISE institutions and aim to reconceptualize ISE to reflect a broader range of cultures and practices (CAISE, 2018; Dawson, 2014a; Garibay et al., 2017). Consistent with literature in the field of organizational change, we refer to these actions as *inclusive practices* (e.g., Sabharwal, 2014).

Foregrounding Culture and Context in Evaluation

The practice of evaluation has faced similar calls to address its grounding in dominant culture practices and assumptions. Like ISE, evaluation is shaped by values and beliefs that are culturally defined (Hood, Hopson, & Kirkhart, 2015). Though frequently portrayed as value-neutral, evaluation advances particular values through decisions around study purposes and audiences, key questions, methods, dissemination of results, and social relational aspects of the evaluator's role (Greene, 1997, 2012). These choices

are often guided by a form of "epistemological ethnocentrism" (Reagan, 1996, cited in Hopson, 2003) through which evaluation practice is shaped by norms and assumptions that privilege dominant worldviews and cultures and marginalize those of non-dominant communities (Hopson, 2003; Thomas & Madison, 2010). The very notion of value-free evaluation assumes dominant culture experiences are normative and universal (Hopson, 2009; Stanfield, 1999).

To counter this ethnocentrism, approaches have been developed that attend explicitly to culture and privilege the lived experiences of people of color, indigenous peoples, and others that have been historically marginalized (Chouinard & Cousins, 2009; Hopson, 2009; Samuels & Ryan, 2011). Among these, culturally responsive evaluation (CRE) has emerged as a prominent framework that recognizes the fundamental importance of the culture(s) of programs, individuals, and communities; engages issues of power, privilege, and marginalization; and advances equity and social justice (Frierson, Hood, & Hughes, 2010; Hood et al., 2015; Hopson, 2009; Samuels & Ryan, 2011). CRE and related approaches also attend closely to context, including the sociopolitical climate, geographic location, economic conditions, and timing associated with a program (Thomas, 2004). Evaluators who foreground culture argue that a program must be evaluated in light of its own unique context (LaFrance, 2004).

In this chapter, we draw on this literature to illuminate the key role evaluation can play in ISE as we look to the future. Although cultural responsiveness in ISE evaluation does not in itself guarantee equity, it can—when done with intention—help illuminate unexamined assumptions, perspectives, and practices that perpetuate inequities. Evaluation can foster more inclusive practices and can play an important role in advancing equity at two levels: specific ISE projects (such as programs or exhibitions) and the ISE field overall. For particular projects, we illustrate how evaluation can and should (a) privilege the voices and lived experiences of non-dominant communities, (b) engage communities in identifying desired project outcomes, and (c) ensure multicultural validity of instruments, measures, and inferences. Concerning the ISE field overall, we describe how evaluation can and should (a) examine conceptualizations of broadening participation, (b) evaluate community-based partnerships, and (c) address replication and scaling.

Evaluating Specific Projects

To strengthen ISE going forward, we argue that evaluations of specific projects should advance equity and inclusive practices in three ways: (a) privilege the voices and lived experiences of non-dominant communities, (b) involve community members in identifying desired project outcomes, and (c) strive for multicultural validity of instruments, measures, and inferences.

Privilege Community Voice

Attending to context and culture in ISE evaluation begins with privileging the voices and lived experiences of community members, particularly communities that have traditionally been excluded or marginalized (Hood et al., 2015; Hopson, 2009). This positions communities as experts about their own lives rather than relying on proxies (Madison, 2007) and ensures that findings reflect a range of perspectives beyond those currently engaged in ISE. Further, when seeking to understand the experience of non-dominant communities, evaluators must actively reject a deficit view (Hopson, 2009; Stanfield, 1999) and guard against deficit models embedded in programs that, in essence, blame individuals for social or educational problems (Thomas & Madison, 2010). This requires specific awareness and sensibilities of the evaluator (Kirkhart, 1995), often enhanced by shared lived experience between evaluators and community members (Hood et al., 2015; Thomas, 2004). Thus, there is a critical need for evaluators of color within ISE, a need shared with the field of evaluation more generally (Hood, 2001; Hopson, 1999).

Centering community perspectives also supports reflective practice, a type of "double-loop learning" in which ISE staff and evaluators examine their own assumptions, attitudes, and understandings about communities and practices (Argyris & Schön, 1995; Rogers & Williams, 2006; Samuels & Ryan, 2011). Through reflection, ISE staff and evaluators can gain insight about the diversity of communities, uncover and examine institutional attitudes, and identify practices that foster or preclude full participation. Ultimately, this builds the organizational capacity necessary to implement inclusive ISE practices (Garibay & Huerta Migus, 2014).

To illustrate, we consider a summative evaluation of an outdoor "parklet" developed to engage Latino families in STEM content (Garibay Group, 2015). Parklets are temporary public spaces that typically extend sidewalks to the width of parking lanes on which they are installed and provide communities with public space and amenities (City of San Francisco, 2015). Located in San Francisco's Mission District, a historically Latino neighborhood, the parklet featured exhibits about water and sustainable water use that were designed and installed by the Exploratorium, a San Francisco science center. The ISE team identified three primary intended outcomes for the parklet: families would (a) actively engage with the exhibits, (b) deepen their understandings about STEM content, and (c) connect to the experience personally and find it culturally relevant. Evaluation of the parklet was grounded in the lived experiences of Mission neighborhood residents, and the evaluation team included bilingual/bicultural evaluators, some of whom were familiar with the Mission's Latino community. The team worked with a neighborhood school to broker relationships with families and interviewed administrators to deepen understanding of contextual issues.

Several aspects of the evaluation findings provided a basis for reflection among ISE staff and evaluators. First, the evaluation illuminated aspects of the Mission's sociopolitical context that had not been considered during the parklet's conception and initial development. Findings revealed that gentrification in the Mission and potential displacement of neighborhood residents had led many community members to feel skeptical of organizations coming from outside the neighborhood. This had affected where the parklet could be placed, and its eventual location was in an area that was less frequented by Latino residents than other candidate locations.

During the design process, exhibit developers anticipated that families' preferences for using Spanish and English would vary (reflecting the complexity of linguistic practices in informal learning; see Garibay & Yalowitz, 2015; Yalowitz, Garibay, Renner, & Plaza, 2015), but the evaluation illuminated many other important dimensions of diversity within the community beyond language that needed to be considered. For example, the study revealed nuances in how families with varying levels of formal education engaged with the exhibits and made meaning of their experiences. Among families in which caregivers had more formal education, group interactions focused on knowledge acquisition, concepts were "taught" or facilitated by a more knowledgeable individual (usually a parent), and learning unfolded unidirectionally. This resembled interactions documented in studies of families visiting museums (e.g., Briseño-Garzón, Anderson, & Anderson, 2007; Moussouri, 2003) and the learning arrangements in schools.

Families in which caregivers had less formal schooling displayed more fluid, dynamic, and exploratory interactions. These groups engaged with the exhibit topics but did not focus exclusively on the science content or processes depicted. Instead, group members introduced a range of ideas and drew on past experiences to make sense of what they encountered. Conversations often involved storytelling and discussion of related experiences. Adults and children worked together to interpret content rather than one person facilitating the interaction.

Finally, the evaluation illuminated assumptions that broadening participation simply requires removing barriers to access. The science center aimed to remove barriers to ISE experiences by placing museum-style exhibits (with an assortment of interactive components and explanatory panels and labels) in a community setting. When these exhibits were encountered in a parklet, however, they seemed out of context and were unfamiliar to many families, prompting some adults to wonder: *What were these devices? Were they put there by the local school or the City? Were they demonstrating something?* Some families were unsure whether it was acceptable to touch or interact with the components and looked to the evaluation team for information about the content or what they should do. Additionally, families saw little in the exhibition that was culturally relevant, although the topic of water was seen as relevant to San Franciscans as a whole. This revealed the ways in which ISE needs to shift its focus beyond providing

access and reconsidering its practices in order to reflect a broader range of cultures and experiences.

By privileging neighborhood residents' perspectives, the evaluation revealed the meaning community members made of their experiences, provided insight about the diversity of the neighborhood the science center sought to serve, and uncovered assumptions about how families interact with museum exhibitions and, more generally, learn together.

Identify Desired Outcomes

Issues of equity and inclusive practices arise early in the evaluation process, often when a project is under development and intended outcomes are specified. Typically, desired outcomes reflect the aims of ISE leaders and staff and project funders. These intentions may not align, however, with the priorities of ISE participants (Allen et al., 2007; Heimlich & Horr, 2010; Perry, 2012) and/or the broader community. Such misalignment can stem, in part, from the considerable variability among ISE participants. Disconnects can also arise due to different perspectives about what counts as learning and what types of knowledge are important, as these understandings are informed by cultural values and priorities (Goodnow, 1990; Rogoff, 2003; Stein, Garibay, & Wilson, 2008). In the same way, conceptualizations of the "problem" to be solved or "need" to be met by ISE reflect culture, values, and assumptions and may differ between ISE staff and members of various communities (Hopson, 2003; Madison, 1992). Further, misalignments can stem from "evaluating down"—that is, adopting a paternalistic and/or deficit viewpoint when specifying intended outcomes for traditionally marginalized groups (Hopson, 2009).

As the field looks to the future, ISE projects should consider community perspectives when defining desired outcomes. A key strategy is front-end evaluation, which is conducted during the beginning phases of planning and has been frequently used to inform exhibition themes and content, improve communication of key ideas, and remove barriers to visitor engagement (Diamond, Luke, & Uttal, 2009; Dierking & Pollock, 1998; Silverman & Korn, 1994). Going forward, we encourage the use of front-end evaluation to understand community members' interests and needs and to enable evaluators and ISE staff to specify intended outcomes that reflect culturally appropriate norms and values (Allen et al., 2007).

A front-end evaluation of a summer environmental science program for youth illustrates these ideas (Garibay, 2006). In this study, a bicultural evaluator worked with the ISE team to examine the program's stated goals, specify intended outcomes, and develop measurable indicators. Youth and parents were interviewed to understand the extent to which proposed target outcomes and indicators aligned with the context and the intended audiences' perspectives and experiences. The evaluation revealed important differences between the ways in which "active parental engagement" was

conceptualized by the ISE staff and parents. On the one hand, the ISE staff identified regular attendance at parent meetings and evening events as *the* key measure of parental engagement. Parents, on the other hand, saw parental engagement as multifaceted. They described, for example, the care and responsibility they took in ensuring that their children were well-rested and ready for the program, getting their children to the program on time, checking in with their children to ask about the program, and being supportive and encouraging of the activities and topics in which their children showed interest. The front-end evaluation also found that parents' work schedules and responsibilities for younger children meant that attending meetings and events was not always possible.

The findings led the ISE team to recognize the hidden assumptions and dominant cultural norms embedded in the proposed indicators and, ultimately, to revise and broaden them to include more culturally relevant behaviors and activities. This process was consistent with calls to "define 'success' or 'effectiveness' in conjunction with a community, using measures grounded in that community's perspectives and values" (Stein et al., 2008, pp. 191–192).

Attend to Multicultural Validity

A third way in which ISE evaluation should attend to equity and inclusive practices is by ensuring that instruments and measures used in evaluation, and inferences drawn from resulting data, are valid for the culture and context in which the evaluation is situated. Validity refers to the soundness and appropriateness of the understandings and inferences drawn from an evaluation (Kirkhart, 1995; Messick, 1988; Thorndike & Thorndike-Christ, 2010). In multicultural settings, validity requires the use of rigorous, culturally appropriate methodology; attention to lived experience (including both the evaluator's awareness of their own cultural location and incorporation of community member's lived experiences into the evaluation); relationships characterized by trust and respect; congruence between theory and cultural context; and consideration of the evaluation's consequences for the community (Kirkhart, 2013b).

At a methodological level, evaluators must examine whether instruments, measures, and data collection procedures will generate accurate, trustworthy understandings (Kirkhart, 2010). This is done in several ways: considering the extent to which measures are relevant to community members' lived experiences; conducting validity studies using culturally appropriate criteria and norming assessments with appropriate reference groups; ensuring that constructs are equivalent across cultures and languages; and determining whether generalizations may or may not be extended across individuals, communities, and cultures (Kirkhart, 1995). Evaluators must actively recognize and incorporate culturally-specific contextual factors into the study, use culturally-relevant interpretive strategies to generate findings,

examine their own values and assumptions, and ensure that interpretations and evaluative conclusions are aligned with communities' lived experiences (Kirkhart, 1995; 2013a; Madison, 2007; SenGupta, Hopson, & Thompson-Robinson, 2004).

Issues of multicultural validity arose in the parklet evaluation described above. Rather than start with predefined, normative measures of what constituted STEM engagement among dominant group visitors, the parklet evaluation revealed and centered community members' authentic practices and styles of interactions. Instruments were developed simultaneously in English and Spanish to address construct equivalence, and data were recorded and analyzed in the language in which they were collected, allowing evaluators to capture nuances not always directly translatable between languages.

The relationship between measurement, equity, and inclusive practices was also evident in an evaluation of a research project concerning facilitation strategies used by science center staff at math exhibits (Garibay & Huerta Migus, 2017). Drawing on established practices for studying math engagement, ISE researchers created rubrics to code video recordings of interactions between families and ISE staff. The rubrics ranked varying types of family interactions as indicators of levels of math engagement (low to high). The evaluation uncovered a tension between the use of existing and accepted measures of math engagement and the need to consider the cultural norms and practices of families from non-dominant groups that participated in the study. Evaluators noted that using the "low to high" rubric could lead to unintentional norming of certain interactions by considering particular behaviors indicative of "more successful" engagement than others. This, in turn, could devalue other types of interaction.

Resisting dominant-group norming can be challenging; therefore, evaluators must scrutinize the contextual and cultural relevance of constructs and instruments and carefully consider the extent to which they can make valid, meaningful inferences across communities and cultures (Chouinard & Cousins, 2009; Kirkhart, 1995). This requires ISE evaluators and staff to identify and document the assumptions embedded in instruments; to mitigate those assumptions when possible; and to clearly articulate instruments' limitations so that evaluative claims can be described and critiqued. In the math facilitation study, evaluation findings helped the research team reconsider how they had defined "successful" math engagement and move away from measures that privileged certain styles of interaction. Instead, researchers focused on the alignment between families' behaviors and facilitators' actions. The team also disaggregated data by demographic categories to check for bias.

Discussion of multicultural validity is particularly critical given the current conversation about shared or common measures in ISE (see Allen & Peterman, this issue; Grack Nelson, Goeke, Auster, Peterman, & Lussenhop, this issue). We recognize that field-wide measures have the

potential to bring much-needed rigor to ISE evaluation, strengthen the evidence base that underpins evaluative claims, enable cross-project and cross-institution conversation about findings, and reduce the burden of instrument development for individual evaluators. As Grack Nelson et al. (this issue) emphasize, however, validity arguments are built on evidence gathered in particular contexts with particular audiences. Evaluators must, therefore, carefully consider the extent to which a shared measure is valid for a given context, culture, and population.

Further, we strongly caution that shared measures and constructs can become de facto representations of normative, dominant culture practices that are then accepted as defining the intended outcomes of ISE. In efforts to broaden STEM participation, however, not all programs can be held to the same standards (Chubin, Harkavy, & Martin-Vega, 2017), since context and culture can intertwine deeply with programs and, in fact, help *constitute* programs rather than merely serve as settings or backdrops (Chouinard & Cousins, 2009; Greene, 2005; Rog, 2012).

Relatedly, we recognize that many summative evaluations of ISE outcomes seek to understand "what works" and to draw comparisons of relative effectiveness across interventions and settings. Often, the methods that support causal inferences require evaluators to strive to reduce complexity and strip away contextual and cultural elements to maximize internal validity (Allen et al., 2007; Campbell & Stanley, 1963). We concur with Cronbach and Associates (1980), however, that these efforts reduce the relevance of evaluation findings. As ISE evaluation continues to develop, we urge the field to embrace an expanded perspective on effectiveness by seeking to understand "what works, for whom, under what circumstances." To do this, we encourage evaluators to continually ground their findings (including data from shared measures) in context and to look to emerging views of causality and causal methods that consider—rather than strive to strip away—context, culture, and complexity (see Gates & Dyson, 2017 for an overview).

Contributions to the Larger ISE Field

Addressing inequities in STEM participation and remaining relevant in an increasingly multicultural society requires the ISE field to look beyond individual projects and enact field-wide efforts to broaden participation in STEM. Evaluation can play a key role in advancing equity and inclusive practices in field-wide efforts if evaluators (a) critically examine conceptualizations of broadening participation, (b) evaluate community-based partnerships, and (c) address replication and scaling.

Examine Conceptualizations of Broadening Participation

Evaluators have a key role to play in critically examining how "broadening participation" is defined and operationalized in field-wide conversations

about STEM learning and engagement. Therefore, we encourage evaluators to pose and investigate questions such as: *Participation by whom? In which aspects of STEM? To what end?* This requires evaluators to explicitly address issues of power and privilege and to situate their work in terms of social justice (Greene, Millett, & Hopson, 2004; House & Howe, 1999), as they direct evaluative efforts toward the distribution of social "benefits and burdens" (Madison, 2007, p. 108).

As a first step, we encourage evaluators to examine how "broadening participation" is defined in the specific projects they evaluate, and within the field more broadly, as well as the extent to which different communities play—or should play—roles in establishing that definition. Conversations about broadening STEM participation often focus on how to attract and support students in completing degrees and seeking careers in STEM, practices associated with the STEM career "pipeline" (e.g., Chubin et al., 2017). Communities, however, may have a range of priorities beyond career preparation, such as addressing health disparities (e.g., Logan et al., 2015), confronting inequities associated with climate change (e.g., Morello-Frosch, Pastor, Sadd, & Shonkoff, 2009), or responding to local environmental health concerns (e.g., Corburn, 2005). Broadening participation could, therefore, focus on equity of opportunities to access the existing STEM knowledge base, contribute to the generation of STEM knowledge, and/or realize the benefits of STEM (Thomas, as quoted in Clewell & Fortenberry, 2009, p. 12)—that is, to live "lives empowered by STEM literacy, knowledge, and identity" (CAISE, 2018, p. 2).

Considering communities' perspectives, concerns, and values is necessary to ensure that broadening participation is not "done to" or "done for" communities but is genuinely aligned with community values, priorities, and needs. This is consistent with calls to develop ISE in cooperation with communities, rather than seeking to serve them as clients (Feinstein, 2017; Feinstein & Meshoulam, 2014), and calls to include communities in defining their own needs and solutions (Madison, 1992).

To illustrate how "broadening participation" can be conceptualized differently even within the same project, we turn to a citizen science program which seeks to foster understanding of the value of green spaces for birds and "reach diverse urban audiences who do not already participate in science or scientific investigation" (Cornell Lab of Ornithology, 2016, p. 1). Implementation research in five U.S. cities revealed two contrasting approaches adopted in the project (Garibay Group, 2018b). The first approach focused on having ISE staff provide special, short-term programs and events for underserved groups. For example, ISE staff developed and led hands-on activities at an afterschool program in an underserved urban neighborhood. The second approach sought to build capacity in communities to develop and lead their own programs on an on-going basis. For example, ISE staff facilitated activities jointly with afterschool program staff drawn from the underserved community. Afterschool staff increased their level of involvement

over time, with the explicit long-term goal of equipping the afterschool staff to implement the program independently.

The two approaches had different implications for program design, desired outcomes, and project timeframe. When "broadening participation" was conceptualized as reaching out to underserved audiences, the personnel and skill requirements centered on developing and implementing programs, project ownership resided with programming staff, and timelines were short. The second approach aimed to build capacity within the community, which required different personnel and skills. Activities centered on strategy development and relationship building and included staff in senior positions to envision high-level aims and focus explicitly on carrying out a long-term strategy.

Within and across projects, evaluators should carefully examine the different approaches to broadening participation in ISE and the implications of those approaches. Rather than simply measuring a project's success in achieving its intended outcomes, evaluators should draw explicit lines between the project's intended outcomes and larger initiatives in the field—at the community, regional, state, or national levels. We encourage evaluators to examine the rationale underpinning a project's intended outcomes and the extent to which the desired outcomes align with larger conceptualizations of broadening participation.

By making connections and spotlighting alignment, evaluators can prompt critical reflection at the program level and also advance field-level efforts to refine our understanding of what it means to broaden participation. Evaluators are uniquely positioned and equipped to speak across projects, identify and integrate empirical evidence, and develop strategies to assess progress. For example, the first author of our chapter serves on the leadership team of the CAISE Broadening Participation Task Force that is currently engaging stakeholders across the country to develop a field-wide strategy for broadening participation in ISE (CAISE, 2017). In a similar fashion, we urge evaluators to be at the table with researchers, educators, policy makers, funders, and communities to develop an expansive and critical conceptualization of broadening participation as we move forward.

Evaluate Community-Based Partnerships

Partnering with local communities is increasingly viewed as critical to broadening STEM participation over the long term and on a large scale. ISE organizations, therefore, are encouraged to partner with schools, afterschool programs, and community-based organizations to extend and amplify what they could achieve independently (National Research Council, 2014; Ucko, 2016). To advance these efforts, funders are beginning to focus on partnership-level outcomes such as the development of collaborative infrastructures and the potential for long-term partnership as a critical aspect of broadening participation (see, e.g., National Science Foundation,

2018). This represents a shift for the ISE field in that partnerships have often been considered a means to accomplish participant-level program outcomes rather than seen as an outcome in and of themselves.

As we move forward, ISE evaluators must be prepared to evaluate entire partnerships, in addition to individual programs or activities, and to disseminate lessons learned across the ISE field. This will require attention to partnership features, such as roles, responsibilities, communication, and relationship dynamics (Garibay Group, 2018b); partners' varying values, understandings, and viewpoints (Stein & Valdez, 2016); and partnership infrastructure such as networks that connect organizations across different sectors and geographies, shared metrics and measurement strategies, and systems for communicating promising practices and lessons learned (National Science Foundation, 2018).

To illustrate, we share an evaluation of a project through which Native American and non-Native American youth and their families engaged with traditional ecological knowledge and western ecological science via traveling exhibitions, a website, and activity kits (Stein & Valdez, 2016). The project also provided opportunities and resources for collaboration between the Oregon Museum of Science and Industry and six Native American partner organizations and groups: the Indigenous Education Institute, The National Museum of the American Indian, the Confederated Tribes of the Umatilla Indian Reservation, Tulalip Tribes, Pacific America Foundation and Waikalua Loko Fishpond Preservation Society, and Eastern Band of Cherokee Indians (Stein & Valdez, 2016). The evaluation examined not only the outcomes associated with the exhibitions, website, and kits but also the experience and outcomes of the larger partnership.

The evaluation team enacted a collaborative, participatory process to examine the intended partnership outcomes, including increasing participants' skills and confidence related to collaboration between science museums and tribal communities. The findings revealed differences in how science museum participants and tribal participants conceptualized "increased confidence" related to collaboration. "For several tribal partners in particular, confidence was defined as increased trust toward a non-Native science institution to represent their story, their voice and their community in an inclusive and respectful way" (Stein & Valdez, 2016, p. 13). The science museum team expressed "increased confidence" through familiarity and comfort with cultural protocols, reduced fear of making mistakes, and sharing knowledge with mutual respect. The study also identified similarities across participants' understandings of collaboration, reflecting themes of respect, honoring one another's voices and knowledge, working toward shared goals, and mutual benefit.

The study illuminated the importance of building relationships over time and in-person, and allowing sufficient time for collaborative decision-making. This was illustrated both when partners came to understand one another's goals and perspectives as the project unfolded and when

communication broke down and all voices did not feel heard. The evaluation also noted the importance of building on prior relationships and involving all partners in the earliest stages of project conceptualization and development. By examining the partnership as well as its public-facing "products," this evaluation provided critical feedback to inform long-term change by highlighting key elements of a successful partnership; identifying strategies the ISE institution should enact and/or continue to enact in the future; and, perhaps most important, illuminating similarities and differences in how partnerships and collaboration were understood and enacted.

Findings such as these have critical implications for the field going forward. We urge evaluators to leverage evaluation findings to support reflection and conversation within and across programs, and within and across institutions, to foster field-level dialog about partnership and ISE. The robust infrastructure within ISE offers a number of interconnected platforms (such as CAISE, InformalScience.org, Association of Science-Technology Centers, Visitor Studies Association, etc.) that evaluators can use to elevate reflection and discourse from the project level to the national level.

Address Replication and Scaling

Finally, as efforts to broaden STEM participation take on greater urgency, STEM educators, funders, and leaders seek to replicate successful programs and scale them up to reach more individuals and communities. Context and culture, however, are deeply intertwined with the conceptualization of needs and desired interventions and with the implementation and outcomes of programs (Chouinard & Cousins, 2009; Greene, 2005; Rog, 2012). Replication and scaling can, therefore, actually constrain the ability of programs, partnerships, and ISE institutions to create meaningful change if expectations for replication and scaling do not account for cultural and contextual factors.

To help advance equity and inclusive practices at the field level, evaluators should attend to the ways that context and culture constitute programs, examine how replication may or may not be appropriate for a given program, and identify what adaptations may be needed if program models are transferred across contexts. When reporting findings, evaluators should explicitly address these themes and offer evidence-based recommendations for scaling.

To illustrate, we provide an example of how evaluation can both illuminate the intertwining of program, culture, and context and speak to policy-level aims for replication and scaling. Children Investigating Science with Parents and Afterschool (CHISPA) was an initiative that sought to foster science engagement of Latino children through local-level partnerships (Garibay Group, 2018a). In eleven metropolitan areas with growing Latino populations, science museums led by the Phillip and Patricia Frost Museum of Science teamed with UnidosUS (formerly, National Council of

La Raza) and ASPIRA Association affiliate organizations that served these communities to offer science lessons for children attending afterschool programs, workshops for their parents, and a science center event for the whole family. Due to CHISPA's innovative model and initial successes, there was considerable interest in expanding the program to serve more families and communities.

The CHISPA summative evaluation found that participating afterschool programs varied widely in their structures and their relationships with parents and caregivers. In addition, the families served by the afterschool programs differed considerably from one community to the next in terms of countries of origin, adults' and children's language skills and preferences, adults' levels of education, and local context. Evaluation reporting, therefore, foregrounded the ways in which culture and context were intertwined with the program, highlighted the importance of customization in program success, and identified the program's flexible resources and structures that could be adapted to particular communities and organizations while achieving overarching aims. Findings were incorporated into a field-wide dissemination strategy that included posting the evaluation report to InformalScience.org, presentations at ISE conferences, webinars for science museums and afterschool programs, discussions with the funder, and conversation during a National Science Foundation video showcase.

Evaluators are uniquely situated to identify evidence-based strategies for replication, scaling, and adaptation, given their vantage across projects and their role in examining project outcomes in light of contextual and cultural factors. Going forward, their insights can support meaningful and successful scaling of projects to broaden participation in ISE engagement and learning.

Conclusion

This chapter has explored the relationship between equity and evaluation in the context of ISE in the United States, highlighting the ways in which evaluation can be leveraged to enact more inclusive practices and achieve greater equity going forward. We have argued that evaluations of specific projects should privilege the voices and lived experiences of non-dominant communities; engage communities in identifying desired project outcomes; and strive for multicultural validity of instruments, measures, and inferences. Further, we have called on ISE evaluators to actively engage in field-wide efforts to broaden STEM participation by critically examining conceptualizations of broadening participation, evaluating community-based partnerships, and addressing aims for replication and scaling.

ISE can play a key role in facilitating meaningful learning opportunities for non-dominant communities (Bell et al., 2009) and supporting community members in becoming empowered by STEM (CAISE, 2018). We look forward to ongoing conversation with other ISE evaluators about how we

can foster greater equity in our field and, more broadly, we invite continued discussion with evaluators in other fields of practice about the role evaluation can and should play in advancing equity in our communities and beyond.

Acknowledgements

The authors thank the ISE projects and evaluators cited here for generously sharing their expertise; Alice Fu, Archana Kannan, and Richard Shavelson for their guidance in framing this chapter and comments on earlier drafts; and Katrina Bledsoe and Stuart Shea for their feedback on previous versions of this chapter.

References

Allen, S., Gutwill, J., Perry, D. L., Garibay, C., Ellenbogen, K. M., Heimlich, J. E., … Klein, C. (2007). Research in museums: Coping with complexity. In J. H. Falk, L. D. Dierking, & S. Foutz (Eds.), *In principle, in practice: Museums as learning institutions* (pp. 229–246). Lanham, MD: AltaMira Press.

Argyris, C., & Schön, D. A. (1995). *Organizational learning II: Theory, method and practice*. Reading, MA: Addison-Wesley.

Ash, D., & Lombana, J. (2013). Reculturing museums: Working toward diversity in informal settings. *Journal of Museum Education*, *38*(1), 69–80.

Bell, P., Lewenstein, B., Shouse, A., & Feder, M. (Eds.). (2009). *Learning science in informal environments: People, places, and pursuits*. Washington, DC: National Academies Press. Retrieved from https://www.nap.edu/catalog/12190/learning-science-in-informal-environments-people-places-and-pursuits

Briseño-Garzón, A., Anderson, D., & Anderson, A. (2007). Adult learning experiences from an aquarium visit: The role of social interactions in family groups. *Curator: The Museum Journal*, *50*(3), 299–318.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin.

Center for the Advancement of Informal Science Education. (2017). *Introducing the CAISE task force on broadening participation*. Retrieved from http://informalscience.org/news-views/introducing-caise-task-force-broadening-participation

Center for the Advancement of Informal Science Education. (2018). *Broadening participation task force: February 2018 update*. Retrieved from http://www.informalscience.org/news-views/broadening-participation-task-force-february-2018-update

Chouinard, J. A., & Cousins, J. B. (2009). A review and synthesis of current research on cross-cultural evaluation. *American Journal of Evaluation*, *30*(4), 457–494.

Chubin, D., Harkavy, I., & Martin-Vega, L. (2017). *Better STEM outcomes: Developing an accountability system for broadening participation* (Report on an NSF-funded Workshop). Retrieved from https://www.nettercenter.upenn.edu/sites/default/files/Better_STEM_Outcomes_Developing_an_Accountability_System_for_Broadening_Participation_Final2.pdf

City of San Francisco. (2015). *San Francisco parklet manual*. Retrieved from http://pavementtoparks.org/parklets

Clewell, B., & Fortenberry, N. (Eds.). (2009). *Framework for evaluating impacts of broadening participation projects*. Retrieved from https://www.nsf.gov/od/broadeningparticipation/framework-evaluating-impacts-broadening-participation-projects_1101.pdf

Committee on Equal Opportunities in Science and Engineering. (2017). *Accountability for broadening participation in STEM: CEOSE 2015–2016 report to Congress*. Arlington, VA: National Science Foundation. Retrieved from https://www.nsf.gov/od/oia/activities/ceose/

Corburn, J. (2005). *Street science: Community knowledge and environmental health justice*. Cambridge, MA: MIT Press.

Cornell Lab of Ornithology. (2016). *What is CUBs?* Retrieved from https://celebrateurbanbirds.org/about/what-is-celebrate-urban-birds/

Cronbach, L. J. & Associates. (1980). *Toward reform of program evaluation*. San Francisco, CA: Jossey-Bass.

Dawson, E. (2014a). Equity in informal science education: Developing an access and equity framework for science museums and science centres. *Studies in Science Education*, *50*(2), 209–247.

Dawson, E. (2014b). "Not designed for us": How science museums and science centers socially exclude low-income, minority ethnic groups. *Science Education*, *98*(6), 981–1008.

Dawson, E. (2017). Social justice and out-of-school science learning: Exploring equity in science television, science clubs and maker spaces. *Science Education*, *101*(4), 539–547.

Diamond, J., Luke, J. J., & Uttal, D. H. (2009). *Practical evaluation guide: Tools for museums and other informal educational settings* (2nd ed.). Lanham, MD: AltaMira Press.

Dierking, L. D., & Pollock, W. (1998). *Questioning assumptions: An introduction to front-end studies in museums*. Washington, DC: Association of Science and Technology Centers.

Falk, J. H. (2001). Free-choice science learning: Framing the discussion. In J. H. Falk (Ed.), *Free-choice science education: How we learn science outside of school* (pp. 3–20). New York: Teachers College Press.

Farrell, B., & Medvedeva, M. (2010). *Demographic transformation and the future of museums*. Retrieved from https://www.aam-us.org/wp-content/uploads/2017/12/Demographic-Change-and-the-Future-of-Museums.pdf

Feinstein, N. W. (2017). Equity and the meaning of science learning: A defining challenge for science museums. *Science Education*, *101*(4), 533–538.

Feinstein, N. W., & Meshoulam, D. (2014). Science for what public? Addressing equity in American science museums and science centers. *Journal of Research in Science Teaching*, *51*(3), 368–394.

Frierson, H. T., Hood, S., & Hughes, G. B. (2010). A guide to conducting culturally-responsive evaluations. In J. Frechtling (Ed.), *The 2010 user-friendly handbook for project evaluation* (pp. 75–96). Arlington, VA: National Science Foundation.

Garibay, C. (2006). *Primero La Ciencia remedial evaluation*. Unpublished manuscript. Chicago, IL: Chicago Botanic Garden.

Garibay, C. (2009). Latinos, leisure values, and decisions: Implications for informal science learning and engagement. *The Informal Learning Review*, *94*, 10–13.

Garibay, C., & Huerta Migus, L. (2014). *The inclusive museum: A framework for sustainable and authentic institutional change*. Retrieved from http://community.astc.org/ccli/philosophy-approach/history-and-definitions

Garibay, C., & Huerta Migus, L. (2017). *REVEAL: Process evaluation report on culturally responsive research*. Retrieved from http://informalscience.org/reveal-process-evaluation-report-culturally-responsive-research

Garibay, C., Lannes, P., & González, J. (2017). *Latino audiences: Embracing complexity*. Retrieved from https://www.exploratorium.edu/genial-articles

Garibay, C., & Yalowitz, S. (2015). Redefining multilingualism in museums: A case for broadening our thinking. *Museums & Social Issues*, *10*(1), 2–7.

Garibay Group. (2015). *Ciencia Pública summative report*. Retrieved from http://informalscience.org/ciencia-publica-summative-report

Garibay Group. (2018a). *CHISPA final evaluation report*. Manuscript in preparation.

Garibay Group. (2018b). *Examining contextual factors that influence the implementation of projects designed to improve cultural diversity in informal STEM programming: Research report*. Manuscript in preparation.

Gates, E., & Dyson, L. (2017). Implications of the changing conversation about causality for evaluators. *American Journal of Evaluation*, *38*(1), 29–46.

Goodnow, J. J. (1990). The socialization of cognition: What's involved? In J. W. Stigler, R. A. Schweder, & G. Herdt (Eds.), *Cultural psychology: Essays on comparative human development* (pp. 259–286). Cambridge, UK: Cambridge University Press.

Greene, J. C. (1997). Evaluation as advocacy. *Evaluation Practice*, *18*(1), 25–35.

Greene, J. C. (2005). Context. In S. Mathison (Ed.), *Encyclopedia of Evaluation* (pp. 255–256). Thousand Oaks, CA: SAGE.

Greene, J. C. (2012). Values-engaged evaluations. In M. Segone (Ed.), *Evaluation for equitable development results* (pp. 192–206). New York: UNICEF.

Greene, J. C., Millett, R. A., & Hopson, R. K. (2004). Evaluation as a democratizing practice. In M. T. Braverman, N. A. Constantine, & J. K. Slater (Eds.), *Foundations and evaluation: Contexts and practices for effective philanthropy* (pp. 96–118). San Francisco, CA: Jossey-Bass.

Heimlich, J. E., & Horr, E. E. T. (2010). Adult learning in free-choice, environmental settings: What makes it different? In E. W. Taylor & M. M. Parrish (Eds.), *New Directions for Adult & Continuing Education*, *127*, 57–66.

Herrenkohl, L. R., & Bevan, B. (2017). What science and for whom?: An introduction to our focus on equity and out-of-school learning. *Science Education*, *101*(4), 517–519.

Hood, S. (2001). Nobody knows my name: In praise of African American evaluators who were responsive. In J. C. Greene & T. A. Abma (Eds.), *Responsive Evaluation. New Directions for Evaluation*, *92*, 31–44.

Hood, S., Hopson, R. K., & Kirkhart, K. E. (2015). Culturally responsive evaluation: Theory, practice, and future implications. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), *Handbook of practical program evaluation* (4th ed., pp. 281–317). Hoboken, NJ: Jossey-Bass.

Hopson, R. K. (1999). Minority issues in evaluation revisited: Re-conceptualizing and creating opportunities for institutional change. *American Journal of Evaluation*, *20*(3), 445–451.

Hopson, R. K. (2003). *Overview of multicultural and culturally competent program evaluation*. Woodland Hills, CA: California Endowment.

Hopson, R. K. (2009). Reclaiming knowledge at the margins: Culturally responsive evaluation in the current evaluation moment. In K. E. Ryan & J. B. Cousins (Eds.), *The SAGE international handbook of educational evaluation*. Thousand Oaks, CA: SAGE.

House, E. R., & Howe, K. R. (1999). *Values in evaluation and social research*. Thousand Oaks, CA: SAGE.

Kirkhart, K. E. (1995). Seeking multicultural validity: A postcard from the road. *American Journal of Evaluation*, *16*(1), 1–12.

Kirkhart, K. E. (2010). Eyes on the prize: Multicultural validity and evaluation theory. *American Journal of Evaluation*, *31*(3), 400–413.

Kirkhart, K. E. (2013a). Advancing considerations of culture and validity. In S. I. Donaldson (Ed.), *The future of evaluation in society: A tribute to Michael Scriven*. Charlotte, NC: Information Age Publishing.

Kirkhart, K. E. (2013b, April 21–23). *Repositioning validity*. Paper presented at the Center for Culturally Responsive Evaluation and Assessment conference. Chicago, IL.

Retrieved from https://crea.education.illinois.edu/docs/librariesprovider13/default-document-library/2013-crea-paper-karen-kirkhart-repositioning-validity.pdf?sfvrsn=2

LaFrance, J. (2004). Culturally competent evaluation in Indian Country. In M. Thompson-Robinson, R. Hopson, & S. SenGupta (Eds.), *In Search of Cultural Competence in Evaluation: Toward Principles and Practices. New Directions for Evaluation*, *102*, 39–50.

Logan, R. A., Wong, W. F., Villaire, M., Daus, G., Parnell, T. A., Willis, E., & Paasche-Orlow, M. K. (2015). *Health literacy: A necessary element for achieving health equity*. Washington, DC: National Academy of Medicine. Retrieved from https://nam.edu/wp-content/uploads/2015/07/NecessaryElement.pdf

Madison, A. (1992). Primary inclusion of culturally diverse minority program participants in the evaluation process. In A. Madison (Ed.), *Minority Issues in Program Evaluation. New Directions for Evaluation*, *53*, 35–43.

Madison, A. (2007). New directions for evaluation coverage of cultural issues and issues of significance to underrepresented groups. In S. Mathison (Ed.), *Enduring Issues in Evaluation: The 20th Anniversary of the Collaboration between NDE and AEA. New Directions for Evaluation*, *114*, 107–114.

Mertens, D. M. (1999). Inclusive evaluation: Implications of transformative theory for evaluation. *American Journal of Evaluation*, *20*(1), 1–14.

Messick, S. (1988). The once and future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 33–45). Hillsdale, NJ: Lawrence Erlbaum.

Morello-Frosch, R., Pastor, M., Sadd, J., & Shonkoff, S. B. (2009). *The climate gap: Inequalities in how climate change hurts Americans & how to close the gap*. Los Angeles, CA: University of Southern California Program for Environmental and Regional Equity. Retrieved from https://dornsife.usc.edu/pere/climategap/

Moussouri, T. (2003). Negotiated agendas: Families in science and technology museums. *International Journal of Technology Management*, *25*(5), 477–489.

National Research Council. (2014). *STEM Learning is everywhere: Summary of a convocation on building learning systems*. Retrieved from https://www.nap.edu/catalog/18818/stem-learning-is-everywhere-summary-of-a-convocation-on-building

National Science Foundation. (2018). *NSF INCLUDES: Report to the nation*. Washington, DC: National Science Foundation. Retrieved from https://www.nsf.gov/news/special_reports/nsfincludes/index.jsp

Nightingale, E., & Sandell, R. (2012). Introduction. In E. Nightingale & R. Sandell (Eds.), *Museums, equality, and social justice* (pp. 1–9). Abingdon, UK: Routledge.

Perry, D. L. (2012). *What makes learning fun?: Principles for the design of intrinsically motivating museum exhibits*. Lanham, MD: AltaMira Press.

Philip, T. M., & Azevedo, F. S. (2017). Everyday science learning and equity: Mapping the contested terrain. *Science Education*, *101*(4), 526–532.

Reagan, T. (1996). *Non-western educational traditions: Alternative approaches to educational thought and practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

Rog, D. J. (2012). When background becomes foreground: Toward context-sensitive evaluation practice. In D. J. Rog, J. L. Fitzpatrick, & R. F. Conner (Eds.), *Context: A Framework for Its Influence on Evaluation Practice. New Directions for Evaluation*, *135*, 25–40.

Rogers, P. J., & Williams, B. (2006). Evaluation for practice improvement and organizational learning. In I. Shaw, J. C. Greene, & M. M. Mark (Eds.), *The SAGE handbook of evaluation* (pp. 76–97). London, UK: SAGE.

Rogoff, B. (2003). *The cultural nature of human development*. New York: Oxford University Press.

Sabharwal, M. (2014). Is diversity management sufficient? Organizational inclusion to further performance. *Public Personnel Management*, *43*(2), 197–217.
Samuels, M., & Ryan, K. (2011). Grounding evaluations in culture. *American Journal of Evaluation*, *32*(2), 183–198.
SenGupta, S., Hopson, R. K., & Thompson-Robinson, M. (2004). Cultural competence in evaluation: An overview. In M. Thompson-Robinson, R. Hopson, & S. SenGupta (Eds.), *In Search of Cultural Competence in Evaluation: Toward Principles and Practices*. *New Directions for Evaluation*, *102*, 5–19.
Silverman, L. H., & Korn, R. (1994). Enemies to allies: Using front-end evaluation to inform interpretive planning in an historical exhibition. *Current Trends in Audience Research and Evaluation*, *8*, 4–8.
Stanfield, J. H. (1999). Slipping through the front door: Relevant social scientific evaluation in the people of color century. *American Journal of Evaluation*, *20*(3), 415–431.
Stein, J. K., Garibay, C., & Wilson, K. E. (2008). Engaging immigrant audiences in museums. *Museums & Social Issues*, *3*(2), 179–196.
Stein, J., & Valdez, S. (2016). *Roots of wisdom summative evaluation: Professional audience impacts*. Retrieved from http://informalscience.org/roots-wisdom-summative-evaluation-professional-audience-impacts
Thomas, V. G. (2004). Building a contextually responsive evaluation framework: Lessons from working with urban school interventions. In V. G. Thomas & F. I. Stevens (Eds.), *Co-constructing a Contextually Responsive Evaluation Framework: The Talent Development Model of School Reform*. *New Directions for Evaluation*, *101*, 3–23.
Thomas, V. G., & Madison, A. (2010). Integration of social justice into the teaching of evaluation. *American Journal of Evaluation*, *31*(4), 570–583.
Thorndike, R. M., & Thorndike-Christ, T. (2010). *Measurement and evaluation in psychology and education*. Boston, MA: Pearson.
Ucko, D. (2016). *CAISE guide to resources for broadening participation in STEM*. Retrieved from http://informalscience.org/caise-guide-resources-broadening-participation-stem
U.S. Bureau of Labor Statistics. (2017). *Household data, annual averages*. Retrieved from https://www.bls.gov/cps/cpsaat11.htm
Yalowitz, S., Garibay, C., Renner, N., & Plaza, C. (2015). Bilingual Spanish–English intergenerational groups' experiences in bilingual exhibitions. *Museums & Social Issues*, *10*(1), 35–51.

CECILIA GARIBAY, PhD, is principal of Garibay Group, where she brings a bilingual/bicultural perspective to research and evaluation of informal learning projects and specializes in culturally responsive and contextually relevant approaches.

REBECCA M. TEASDALE, MA, is senior evaluation and research associate at Garibay Group and a doctoral candidate in educational psychology at the University of Illinois at Urbana-Champaign, focusing on evaluation methodology.