

Attitudes of Stakeholders towards Assessment in the Informal Science Education Realm

Alan J. Friedman
Consultant for Museum Development and Science Communication
Alan@FriedmanConsults.com
May 4, 2012

INTRODUCTION

The practice and use of assessments in the informal science education (ISE) realm is highly diverse and inconsistent, with differing stakeholders having dramatically different attitudes towards which assessments (if any) they value. There are several reasons for this variety:

- There is no formal organization or hierarchy in the ISE realm, in contrast to the structure of formal education.
- Even within an ISE organization, there are different interests in assessment among the administrators, marketers, fundraisers, educators, exhibition developers, and boards.
- No single funding source or agency dominates the ISE realm, again unlike most public schools or universities. Different ISE funding sources come with different evaluation goals and hence different assessment interests.
- Much of ISE is simply not institutionally organized. It is difficult to set an assessment agenda for millions of individuals who make daily choices to participate in STEM-related hobbies, reading, television watching, and surfing the web.

This essay reviews the landscape of attitudes and uses of assessment on the part of informal science education stakeholders beyond the research community.¹

ISE assessment today largely follows one of two sets of rubrics, that of the NRC's

¹ This essay uses the terms "assessment" and "evaluation" as defined in *LSIE (2009)*. Among practitioners in ISE, the term "assessment" is rarely used, and instead "evaluation" is used to refer to assessments as well as evaluations.

Learning Science in Informal Environments (2009) and the NSF's *Evaluating Impacts of Informal Science Education Projects* (2008). ISE practitioners most often employ the NSF rubric because it is required for the NSF's Project Monitoring System. As suggested in the chart below, these two sets are similar.

Strands from NRC's <i>Learning Science in Informal Environments</i>	Impact categories from NSF's <i>Framework for Evaluating Impacts of Informal Science Education Projects</i>
Develop interest in STEM	Engagement or interest
Understand STEM knowledge	Awareness, knowledge, or understanding
Engage in STEM reasoning	Skills
Reflect on STEM	Behavior
Engage in the practice of STEM	Attitude
Identify with the STEM enterprise	Other

STAKEHOLDERS

FUNDERS: Certain funders, in particular the **NSF**, have largely driven the practice of assessment in ISE in the United States. NSF has long looked for assessment of learning of the particular cognitive objectives of a given ISE intervention, but more recently NSF has encouraged assessment of any combination of its impact categories, so that a measure of impact on interest alone is now potentially an acceptable goal in a proposal. Separately, NSF also encourages assessment for the purposes of research into the impacts of ISE in general, for the purpose of finding more effective models for ISE activities, and for the purposes of front-end, formative, and remedial evaluations.

Other funders have a less consistent interest in assessment. Some **corporate funders** decline to pay for assessment. One corporate funder told the author "We want to see all of our money on the exhibit floor," and another said the only assessment that would matter would be whether sales of their products increased as a result of their sponsorship of a major exhibition.

These funders do value user demographic information, however, and they have no objection to someone else paying for and performing assessment of LSIE's strands of learning in a program, although they have no urgent use for this data. **High-tech industry funders**, in contrast, are very interest in workforce development. So they value assessments of long-term impacts for the strands of interest and identity, especially if research can demonstrate that these strands lead to careers in STEM.

Recently some **private foundation funders** have become much more interested in impact assessments. The Noyce Foundation and the Gordon and Betty Moore Foundation, for example, have made major investments in developing assessments, as well as encouraging more rigorous assessment as part of project funding. From the perspective of these funders, the goal is not only to help them evaluate each investment the foundation has made, but also to make the funder "a smarter investor" for future grants. These stakeholders are primary audiences for and proponents of generalizable knowledge about ISE. Some other foundation funders tend to invest in assessments of STEM knowledge, with less attention to other outcomes listed in the two rubrics above, such as positive attitudes towards STEM or identifying with the STEM enterprise. At the same time, however, many foundation funders are also interested in youth development, and would welcome research connecting the LSIE strands with established measures of youth development such as resiliency.

ADMINISTRATORS: CEO's, Directors, Board Members, and other senior administrators occasionally are strong supporters of outcome assessments, particularly if those administrators have academic backgrounds or have used formative evaluation to improve their projects as educators, filmmakers, programmers, or exhibition developers. These administrators would find personal gratification and practical value in strong evidence that their institutions are

“making a difference” for the public good, and are likely to be equally pleased if the evidence is for any of the NRC or NSF strands. But a large number of administrative leaders, such as vice presidents for marketing or fundraising, typically see assessment of cognitive impacts as a nice but unnecessary expense unless funders have demanded it. These administrators are often looking for functions of their ISE activities beyond traditional cognitive education, which they perceive as owned by the schools. Assessments clarifying audience identities with STEM or with particular aspects of STEM (e.g., dinosaurs, astronomy), and how an institution and its individual offerings relate to those identities, have been increasingly recognized as valuable by these administrators because these factors relate directly to the organization’s ability to attract audiences and funders.

ISE PRACTITIONERS: staff such as **exhibit designers or educators** also have mixed views on assessment. Front end and formative assessments have the most appeal to them, because they can see the immediate uses of those results to help select and refine elements of their ISE projects. What visitors are already interested in, or could become interested in by visiting an exhibit or attending a program, has immediate utility in selecting and refining a project’s objectives. But assessments designed to measure gains in science knowledge may pose threats. If all or part of a project has no observed impact on knowledge, will those who designed and delivered the project be penalized? If an evaluation takes 15% of the project’s resources, and the resulting project is 15% smaller or shorter, will they get grief for not having produced more program? Research connecting the strands of learning to popularity of the institution and repeat visitorship would increase the acceptability of assessment for many non-academically based practitioners (and administrators as noted earlier).

FORMAL EDUCATORS: Whether a school class visits an aquarium, views a video, or

uses a STEM learning game depends on a **teacher, principal, or district office** approving that use of time and money. That approval usually depends on a claim that the ISE experience is aligned with state science standards and thus will contribute toward the school's performance on high-stakes tests. An exhibition or DVD on rare birds may increase empathy for endangered species, curiosity about the behavior of these unusual-looking animals, and evidence of evolutionary pressures, but if these outcomes are not on state-mandated tests, these findings are not compelling for a teacher or principal who is under pressure to meet annual progress requirements. What these stakeholders would like to see are assessments of the near-term impact of individual ISE experiences on their own tests. But because ISE activities and institutions may not be targeted at the same standards as those tests (few of which include interest, attitude, or engagement in their main scores), these stakeholders could also find a use for research showing that accomplishment of the LSIE strands through ISE activity contributes to improvements of their test-scores over the long run.

THE PUBLIC AUDIENCES: The reasons **individuals and families** give for deciding to use a particular ISE opportunity almost never include a statement about the strong data supporting the accomplishment of NRC or NSF learning strands. When surveyed about why they decided to participate in an ISE experience, consumers usually describe their choices in terms of entertainment value, opportunity for family interaction, or reinforcement of an individual's self-identification as a connoisseur of airplanes, birds, mountains, or stars. Frank Oppenheimer, the founder of the Exploratorium, was fond of noting that "nobody ever flunked museum." Separating ISE experiences from the experience of tests and grades is a positive factor for ISE consumers. Nevertheless, consumers who are parents bringing their children to an ISE experience might find assessments of conceptual learning compelling or useful if they

directly spoke to the consumers' additional preferences for experiences that help children advance in school, and might eventually lead to high-value careers like medicine. In addition, consumers visiting an ISE experience without children could potentially be influenced to participate by assessment results showing that an exhibit or activity is interesting and engaging (there is no research to date showing that ISE audiences are influenced by such assessment results). Positive reviews in newspapers, on the Web, or on television often talk about the reviewer finding a particular experience interesting and engaging, and the reviewer having learned a lot. Positive media reviews are frequently cited by audiences in describing why they chose to participate in an ISE experience. While individual reviews are not rigorous assessments, these impacts suggest that well publicized assessments might have similarly strong influence as well.

THE CHALLENGE FOR RESEARCH

Meeting the needs of the stakeholders described above is not solely a matter of creating better assessments for individual ISE experiences. The challenge is also to produce generalizable research on the ecology of STEM learning that will illuminate the connections among the varieties of formal and informal learning. Practitioners in ISE would love to be able to say that research demonstrates that high quality ISE helps formal education meet its goals, including expanding populations in the STEM workforce. But to make such a claim, the practitioners need broad research evidence that this is generally true at some size effect and under some plausible circumstances, based on assessments that are sufficiently robust and reliable to support such broad conclusions. Finally, practitioners need the assessment instruments and training to evaluate individual programs and identify when "quality ISE" is present; such assessment instruments need not be as robust and reliable as those used in the

broad research, but must meet appropriate standards for their individual uses.