caise | center for advancement of informal science education

Principal Investigator's Guide:

Managing Evaluation in Informal STEM Education Projects

caise | center for advancement of informal science education

VSA Visitor Studies Association

Principal Investigator's Guide: Managing Evaluation in Informal STEM Education Projects

Rick Bonney, Kirsten Ellenbogen, Leslie Goodyear, Rachel Hellenga
Editors

Rick Bonney, Rachel Hellenga, Jessica Luke, Mary Marcussen, Sasha Palmquist, Tina Phillips, Lauren Russell, Saskia Trail, Steven Yalowitz
Authors

Center for Advancement of Informal Science Education
Association of Science-Technology Centers
818 Connecticut Avenue NW. 7th Floor
Washington, DC 20006

InformalScience.org

Center for Advancement of Informal Science Education. (2011). Principal Investigator's Guide: Managing Evaluation in Informal STEM Education Projects. Washington, DC: Author. Retrieved from http://informalscience.org/evaluation/evaluation-resources/pi-guide

This material is based upon work supported by the National Science Foundation (Award Nos. DRL-0638981 / DRL-1212803). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Introduction

An initiative of the Visitor Studies Association, produced in partnership with the Center for Advancement of Informal Science Education.

This Guide is designed to help principal investigators and other leaders of informal STEM (Science, Technology, Engineering, and Math) education projects integrate evaluation into all phases of project design, development, and implementation. Such projects include exhibits, media projects, websites, community science projects, afterschool programs, festivals, family activities, online games, citizen science projects, and other efforts to help people learn about science in the course of their everyday lives.

Project evaluation, carefully designed and conducted in collaboration with project developers, can help any informal STEM project improve its deliverables, achieve its goals, and measure the degree to which its targeted objectives have been met. Also, when results are disseminated widely, evaluations can provide critical information for helping projects throughout the STEM education field improve their overall practice.

By design, most of the authors of this guide are not professional evaluators. Rather, they are informal STEM education project leaders who have, through many years of experience, developed effective means of collaborating with evaluators. Also by design, this guide is not a how-to manual on developing and implementing project evaluations. Instead, it explores the use and application of evaluation in informal STEM education projects with the goal of improving partnerships between project teams and their evaluators. References and links to the many valuable resources that do provide information on how to conduct project evaluations are included along the way.

This Guide should prove particularly useful for Principal Investigators who have received funding from any directorate of the National Science Foundation (NSF) that addresses informal STEM education and which requires external project evaluation. The Guide will also inform prospective Principal Investigators who are preparing proposals for submission to NSF. However, the guide is not an NSF policy document. While the authors and editors have done their best to reflect current thinking of NSF staff, all questions about NSF policy should be directed toward an appropriate NSF program officer.

While the Guide is rooted in experiences and stories that are drawn mainly from NSF-funded projects, the process of developing a proposal, implementing a project, and working with an evaluator described in its pages should be relevant to practitioners working in most informal education environments regardless of the funding source.

We welcome feedback on the Guide and encourage you to comment on its overall quality, value to you, missing information, or other ways in which we can improve future editions. Please send feedback to caise@InformalScience.org.

This page intentionally left blank.

Table of Contents

Chapter 6: Reporting and Dissemination: Building in Dissemination from the Start

Chapter 1
True Stories from Your Peers:

The Interplay Between Evaluation and Project Implementation

RACHEL HELLENGA

Introduction

I am not a professional evaluator, nor are the other contributors to this chapter of the Guide, but we all have traveled to the land of evaluation and are here to report on some of the opportunities and adventures that lie therein. Welcome, fellow travelers. The true stories that follow are intended to engage you in the equivalent of a headlong rush through a foreign market full of colorful textiles, delicious-smelling spices, and tasty treats. I've interviewed many project leads—and added a few stories drawn from my own experience—to paint a vivid picture of the interplay between evaluation and project implementation.

In the following chapters, friendly colleagues who have been in your shoes will address key topics and practical considerations to help you successfully integrate evaluation into your informal STEM education projects. You'll learn about the principles of evaluation and how different types of evaluation can be used (chapter 2). You'll learn how to find an evaluator whose expertise matches your project needs (chapter 3) and how to work with that individual throughout all phases of project development and implementation (chapter 4). You'll delve into the details of designing an evaluation plan (chapter 5). And, you'll learn how evaluation findings can be used to improve your project, inform the work of your peers, and even influence the overall success of your institution and the entire field of informal STEM education (chapter 6). Throughout these chapters you'll find stories told by PIs and other project leads who have integrated evaluation into the implementation of exhibitions, television shows, science café, after-school programs, science festivals, and other informal science education projects.

Who should read this chapter?

Read this chapter if you like to see concrete examples before diving into underlying principles. You may also find this chapter helpful if you are new to the informal STEM education field and want to understand the operating context for evaluation activities. If you prefer to get straight to the details, feel free to jump right to chapter 2.

If you haven't jumped ahead, you're about to read several vignettes organized by the phases of typical project development.

1. First we'll look at setting project goals and objectives so you'll know what success looks like when it's time to measure your project outcomes.
2. Next we'll explore how front-end evaluation can help you better understand the needs and interests of your audience.
3. Following that, a few stories about formative evaluation will demonstrate how participant feedback can surprise you with valuable insights and can inspire course corrections.
4. Next we'll examine how summative evaluation can help you determine whether your project objectives have been achieved.
5. Finally, we'll talk about using your evaluation findings at the reporting and dissemination stages of project development.

Addressing Some Concerns about Evaluation: Time, Money, and Control

Some project leads fear that embracing evaluation means putting visitors or project participants in charge of the design process, a bit like Aesop's fable about a man and his donkey.

TRUE STORIES

A centuries-old dilemma: Getting to your destination while trying to please everyone.

[Aesop's Fable] A man is travelling to market with his son and donkey. The first passerby suggests that the man is foolish because "donkeys are to ride upon." The man puts his son on the donkey, only to hear from the next passerby that "the lazy youngster rides while his father walks." Father and son switch places only to hear "what a bad father to let his son trudge along." When both then hop on the donkey, someone observes that they are overloading it. Eventually the man and his son carry the donkey until they trip on a bridge and the donkey goes overboard. The moral of Aesop's tale: Please all and you will please none.

SPOTLIGHT ON PRACTICE

Wolfquest

This player's blog features screenshots of the immersive enhancements to WolfQuest 2.5. (Figure 1)

Evaluation shouldn't substitute participants' judgment for yours.

Here's a cautionary tale that acknowledges the fine line between inviting participant feedback... [David Schaller] Eighteen months after launch, WolfQuest had attracted a substantial audience: 450,000 game downloads and 150,000 registered users. So, for the next stage of this 3D wildlife simulation game developed by EduWeb and the Minnesota Zoo, we gave players a larger role in the design process. In fact, we put them in charge. In a slight departure from the methods of more traditional formal evaluation, we solicited ideas on new game features with a contest rather than through focus groups or surveys. After receiving nearly 1000 proposals from players we selected the best, then put them up for a vote on the WolfQuest forum.

To our surprise, proposals to improve gameplay and learning (e.g., adding bison as prey animals, creating better multiplayer missions) attracted fewer votes than proposals to add environmental enhancements (rain, snow, and changing times of day). Because we had made the vote public, we had to abide by the results. In retrospect, we realized our two big mistakes:

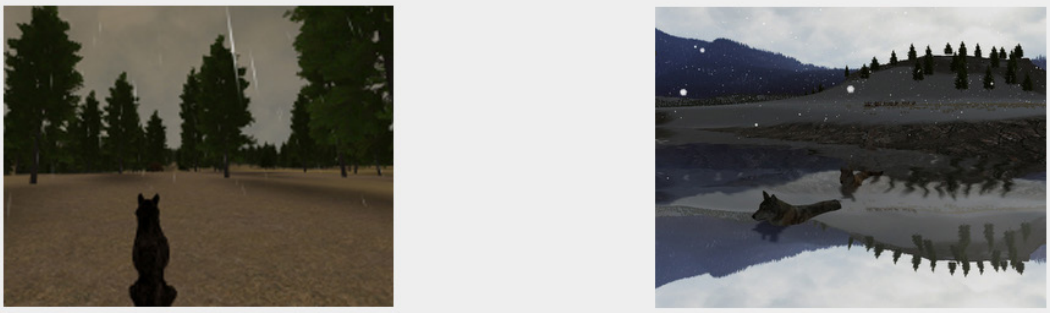


Figure 1

- We had no sampling strategy, so the vote was skewed by a core group of players who used the game mainly as a chat room and who preferred improvements that would enhance their sense of immersion in our wolf-world.
- Instead of using player input to inform project development, we had allowed it to enforce a particular direction.

Democracy has many merits, but benevolent dictatorship is generally a better model for formative evaluation.

In fact, the intent of evaluation is to inform project design rather than to generate or dictate it. This Guide will outline best practices in the integration of evaluation and project implementation and share examples of evaluation in action.

Let's acknowledge another concern: Project leads sometimes view evaluation as a process to be contained so that it doesn't derail the "real work" of project completion. It's true that evaluation can add to project cost and may extend your implementation schedule. However, when evaluation keeps your project on track and enables you to maximize your project's impact, it's time and money well spent. We'll look at a few situations in which evaluation saved money by preventing expensive mistakes; helped attract new funding; and eliminated lengthy debates in the conference room in favor of data collection and reality-testing.

Getting Started: Defining Project Outcomes

Determining goals and objectives requires describing specific outcomes that you hope your project will achieve for a visitor or participant. This process is described in detail in chapter 5, but here's a brief explanation. Typically outcomes are first stated as a series of goals that describe the broadest possible impacts that a project will have within its community, and then secondly as a series of measurable objectives that indicate specific and observable effects on the intended audience. Many evaluators coach project teams by asking some variant of the question: "What will your participants think, do, or feel as a result of their experience?" This exercise requires the team to define and articulate project outcomes that represent the desired effects on the end user. Outcomes can be framed in many ways—in addition to "think" (cognitive outcomes); "do" (behavioral outcomes), or "feel" (affective outcomes), participants may show particular types of engagement, or develop certain skills, or develop new interests in science (types of outcomes are elaborated in chapter 5).

If you'd rather skip reading about setting goals and objectives and jump right into project implementation, please first consider one more tale about the perils of ignoring best practices in project design.

TRUE STORIES

"Ready, aim, fire!" vs. "Ready, fire, fire!": Implementing a project without consensus on its goals.

[Rachel Hellenga] Early in my career I found myself on a project team that was racing to develop a series of exhibit components under the time pressure that we all typically face. We had hired an evaluator to satisfy the requirements of our grant. As an exercise to understand the team's vision, she asked each team member to review a stack of images and to choose those most representative of the exhibition content. Our choices were all over the map! We had not come to an early consensus on project direction or clearly articulated our project outcomes. As a result, exhibit components were chosen based on the advocacy of individual team members. Instead of "Ready, aim, fire," it was "Ready, fire, fire." Unfortunately, the exhibition never really jelled—although it did open on time!

A goal and objective-setting exercise is NOT the time to talk about outputs, which may be project deliverables ("we will produce a video") or logistics ("we will reach 5,000 students in the first year") or content ("the topic is earthquake modeling"). Rather, outcomes are the effects that you hope to achieve, and continuing to develop your program model requires spelling them out in detail. Not only will the process provide you with measures of success, it also will suggest strategies for project development that will help you achieve your outcomes.

TRUE STORIES

Defining Outcomes (Not Outputs leads to new program strategies and measurable results).

[Kathleen Burke] When developing a body image program for teen girls, answering the question "How will the participants be transformed as a result of this project?" was a useful exercise for staff at the Robert Crown Center for Health Education. The word "transform" conveyed the idea of important change rather than big deliverables, and focused their conversation about desired outcomes for the new program. What would girls think, do, or feel differently as a result of participating in our program?

One cognitive outcome defined by the team was for girls who participated in the project to "think critically about the media and realize they can refuse to imitate what they see." This desired outcome led to the development of an activity showing "before and after" pictures of magazine models, which revealed how photo manipulation made the models look thinner. Girls in the pilot program rated this activity very highly, and their responses to an open-ended question after participating in the pilot—I will not compare myself to people and magazines;" "I don't need to look like a model;" "Models don't look the same on TV as they do in real life;" and "Magazines are liars!!!"—showed that the girls were definitely thinking critically about the media.

Learn about your Audience: Front-end Evaluation

SPOTLIGHT ON PRACTICE

Proctor and Gamble

Market research looks a bit like front-end evaluation. [Rachel Hellenga] Market researchers have lots of techniques for learning about their audiences. For example, Proctor and Gamble has sent employees to live with families to understand custom needs and concerns. This Living It research program led P&G employees to notice that water was a scarce commodity in Mexico—often it was lugged to the household from wells or communal pumps. Many middle-class women washed their clothes by hand and a large percentage of them used fabric softener, which required rinsing the clothes extra times. This observation inspired P&G to develop a fabric softener that could be added directly after the wash cycle to reduce three rinse cycles down to one.

Front-end evaluation provides information such as the intended audience's general knowledge, questions, expectations, experiences, motivations to learn, and concerns regarding the topics or themes that will be addressed by your project. Some might argue that this phase of project design is not "evaluation" per se. However, understanding your audience is critically important in project design, and audience research is often facilitated by an evaluator as part of the overall project development process.

TRUE STORIES

Front-end evaluation can identify "hooks" for a tricky subject

[Rachel Hellenga] When replacing the genetics section of the Tech Museum's

LifeTech gallery, my team suggested removing a DNA crime-fighting component and encountered vigorous resistance from members of the marketing department. They were concerned that the exhibition would be all doom and gloom if we focused solely on high-tech strategies for addressing genetic disease. We pointed out that crime fighting didn't fit in a gallery about technology and health. They countered that the crime-fighting exhibit could be "extra"—it wouldn't do any harm to insert something fun to offset the gloomy content.

We addressed the "gloom and doom" risk by conducting front-end evaluation. The evaluators conducted interviews with visitors and presented various written scenarios such as 1) a woman deciding if she needed breast cancer screening based on her family history and 2) a boy who might benefit from human growth hormone. [Please note that the evaluators did not go out on the floor and ask visitors to vote on inclusion of a crime-fighting exhibit!] We learned that visitors were not turned off by the dark side of the topic. In fact, the personal stories were a "hook" that motivated visitors to absorb scientific information so they could perform diagnoses and advise patients.

Once Genetics: Technology with a Twist opened to the public, our summative evaluation found that two-thirds of interviewees described what they took away from the exhibition by naming exhibits related to personal stories. Some visitors expressed surprise at the extent to which genetic disorders impact families, while others found the information from the stories personally relevant. Front-end evaluation had given us confidence to pursue this direction and using the "hook" of personal stories had paid off.

Facilitate Good Decisions: Formative Evaluation

Formative evaluation provides information to improve a project during its design and development. Project leads may be tempted to rely on their own opinions about how a project is shaping up and skip formative evaluation with the argument that "it's too early in the process" followed shortly by the argument that "it's too late in the process." I believe that formative evaluation is always a good idea to do right now. Each round of data collection can provide you with a snapshot of one point in time, delivering rapid feedback that feeds into your project design.

TRUE STORIES

Learn a little before investing a lot.

[Rachel Hellenga] Sometimes an evaluator can help a team reach consensus before incurring the full expense of producing a polished deliverable. During the design phase of the Science Storms exhibition (Figure 2, see page 7) at the Museum of Science and Industry, the design team was planning a component that would allow visitors to race two solar-powered cars around a track. The design showed visitors using a joystick to tilt a solar panel into or out of the path of the light as a method to generate more or less power and control the car. The exhibit developer suggested an alternate strategy of covering and uncovering the solar panels to more clearly indicate how much light was reaching the solar panels. The "covering" approach seemed superior, but the "tilting" approach had already been detailed, so the team debated the relative importance of spending the time and money to change the component before spending tens of thousands of dollars on fabrication.

The developer produced a $20 mockup consisting of a regular light bulb, solar panel, and 3-inch diameter fan. The prototype was

Science Storms "Sunlight from Energy" $20 solar cell kit & final exhibit component. Figure 2

less than rudimentary; not only was it a fraction of the real size but it used a solar-powered fan instead of a car. Nevertheless, pairing it with exhibit text gave the evaluator enough to work with. Formative evaluation determined that visitors understood the concept much better when they were allowed to cover and uncover the solar panel instead of tilting it toward and away from the sun. The team changed the design.

What did they learn? Or do? Or feel? Summative Evaluation

Summative evaluation measures the degree to which objectives for the intended audience have been realized, which explains the importance of setting objectives that can actually be measured. While sometimes thought of as "final" evaluation, summative evaluation can begin long before a project is complete, and it can take many forms-tracking and timing for exhibits; pre- and post-surveys for community projects; analyses of journals for community science projects; interviews with project participants; and many, many more. The following stories about three diverse projects show the power of painting a clear picture of your destination at the outset of a project, not in the form of specific exhibit layouts or project materials or other deliverables, but as a shared vision of the desired outcomes for your visitors or project participants. If you take the time to define your vision of success, you certainly will increase your odds of achieving it.

TRUE STORIES

Summative Evaluation: Measuring cognitive outcomes (What will they learn?)

[Richard Hudson] As we embarked on the summative evaluation of Dragonfly TV, we wanted to measure whether watching the show successfully changed children's appreciation for and understanding of scientific inquiry.

In the planning stages for this show featuring "real kids doing real science," we had identified specific aspects of inquiry that we wanted to convey and incorporated those concepts into the project. Now it was time to see if the message got across. I worked together with our evaluator to craft evaluation questions to measure whether we had achieved this outcome. The dialogue between PI and Evaluator was essential to arriving at a solution. We landed on a strategy of asking kids questions such as "How important is it to you to write down information during an experiment?" and "How important is it to you to keep some things the same each time you repeat the experiment?" The kids ranked the importance of each aspect of inquiry before and after viewing episodes of Dragonfly TV, and through this simple set of questions we were able to demonstrate that watching our show resulted in a significant increase in children's understanding of the process of scientific inquiry.

Summative Evaluation: Measuring Behavioral outcomes (What will they do?)

[Rachel Hellenga] I was on a team that completed a radical overhaul of a popular construction exhibit with the aim of engaging more girls and ensuring that more visitors succeeded at the building activity.

Over the years the museum staff at the Chicago Children's Museum had seen all kinds of wonderful free-standing structures built by visitors to Under Construction every day, but a formal front-end study revealed that only 11 percent of the visitors observed created these structures; others were connecting pieces together or engaging in other activities before moving on to another gallery. One of our desired behavioral outcomes for the Skyline exhibition, to be measured through observation, was to increase the number of children who built a free-standing structure. This desired outcome became a driver for various design strategies such as the substitution of larger nuts and bolts to cut the assembly time in half. The evaluation team determined a set of criteria for "free-standing structure" and observed 100 children using the new exhibition: We were pleased to find that 40 percent of them built stable, free-standing structures.

A second objective was to increase girls' involvement in, as the exhibit was only engaging half of the girls visiting the original exhibit versus about two-thirds of the boys. One of the modifications inspired by this objective was the addition of fabric with appealing textures and colors to the array of available materials. Keeping those outcomes in mind during early stages of the design process helped our team achieve the results we were aiming for by the time we reached the summative evaluation stage. Once Skyline opened, observations conducted by the evaluation team determined that 71 percent of children became engaged in building with no statistically significant differences between males and females.

TRUE STORIES

Summative Evaluation: Measuring affective outcomes (What will they feel?)

[Holly Walter Kerby] Bouncemania integrates elements of a play and a science demonstration in a Wrestlemania-style match between two rubber balls to teach the relationship between polymer structure and functions. The show was created by Fusion Science Theater, a project led by the Madison Area Technical College, and features a live performer who presents the balls as characters: "In this corner, weighing in at 10.5 grams, the five time title winner—a ball with experience, a ball with bearing. Give it up for . . . B.B. the King Bouncer!" The performer shares information about the molecular structures of the balls and asks the audience to vote for which one they think will bounce higher. Then the performer has audience members act as atoms in a physical model of the structures, and once again, predict which ball will bounce higher. The demonstration itself is not that fancy—one ball bounces and one drops with a thud—but by that point in the show, the audience is cheering and exploding with excitement. Committing to a prediction motivates the children to apply what they've learned and to take a strong interest in the outcome of the demonstration.

In addition to looking for cognitive gains, we studied affective outcomes by measuring whether children experienced a positive change in their perception of themselves as individuals who are interested in and capable of learning and doing science. We believed that the application of a playwright's bag of tricks—dramatic irony, plot twists, and character development—would all contribute to engaging children's emotions at a deeper level than a typical science demonstration. Our in-house evaluator worked with us to develop

questionnaires for use before and after the performance. Children used Likert scales to indicate their interest in science and their confidence in their ability to learn science. Their ratings revealed that participating in the performance had a dramatic impact—no pun intended!

On a more informal level, we hold conversations with audience members afterward. One of my favorite conversations went something like this: I asked a boy what he liked about the show and he said "I like BB the King." I asked, "Why?" and he said "Because he won." I responded, "You like it when things win?" and his answer was "I liked it because I knew he was going to win."

Maximizing your Strategic Impact: Reporting and Dissemination

Not only can evaluation findings provide a picture of a project's effectiveness, but they can be valuable in identifying strategic next steps for an organization, in making the case for further funding, and informing the practice of others—but only if the information is embraced by the project team and shared with others who could benefit. Dissemination should be a fundamental part of your overall project evaluation plan and something you factor in from the start.

TRUE STORIES

Integrating evaluation with your strategic vision.

[Rachel Connolly & Pamela Rosenstein] In an ideal world your evaluators go beyond assessing project outcomes; they help to gather strategic information for your organization.

In anticipation of its 2011 The Fabric of the Cosmos documentary series, NOVA hosted a series of "Cosmic Café" around the country with the assistance of chapters of the Society of Physics Students (SPS). Based on the format of a science café featuring an informal discussion with an expert, the Cosmic Café addressed topics raised by the program. We needed our evaluation study to go beyond simply assessing what the attendees at these specific café slearned about Fabric of the Cosmos because we intended to use the evaluation findings to help shape the ongoing national network of science cafés.

It was important for the PI to articulate the big picture to our evaluation team and to plan for this broader scope of work from the start, so that the evaluation team understood it was an ongoing national network that needed to grow. We worked with the evaluators to frame questions about how these cafés could inform the field of informal science education. Armed with the big picture, the evaluators conducted the study and informed us that they saw great potential in further collaboration with professional organizations like the Society of Physics Students. The SPS was a very valuable partner in implementing the cafés; in return, the project helped the undergraduates learn how to present publicly and impressed on them the importance of communicating with the public as they advance in their careers. The recommendation to invest further in that partnership informed our strategy going forward.

TRUE STORIES

Disseminating evaluation findings can attract resources for your project.

[Ben Wiehe] In the early years of the Science Festival Alliance, a common critique that we heard was "we don't need a once a year party, we need a sustained effort."

When the Science Festival Alliance was in its second year, the evaluation covered a huge swath of events that had different formats, served diverse intended audiences, and took place at venues ranging from retirement homes and elementary schools to bars and tattoo parlors. The evaluation findings helped to demonstrate that these festivals are worth the energy and resources: the festivals were

reaching new and diverse audiences; changing attitudes; resulting in cognitive gains; and increasing awareness of the science happening in the region. Over 40 percent of the festival collaborators reported follow-up contacts from the public and 89 percent of the STEM practitioners surveyed indicated an interest in ongoing participation in public outreach for the rest of the year.

Perhaps the most significant finding was that interaction with a STEM professional was identified as the number one predictor of positive learning outcomes, whether in terms of science learning, increased interest, or the perception that science is fun. That was a unique finding and revealed a distinct strength of festivals. The festival organizers disseminated the evaluation results in a wide range of formats such as PowerPoint, conference poster presentations, online PDF documents, and video, all of which helped to get buy-in for annual festivals and supported new festival organizers in making the case to their communities.

Conclusion

Project planning, evaluation, and implementation are all parts of a whole, working best when they are synchronized and coordinated. At its best, evaluation works to answer questions that give a project team a deeper and richer understanding of its own practice—before, during, and after a specific project implementation. Therefore, an evaluation study should evolve and be guided by a project, while the project should be informed and guided by its evaluation. This Guide examines how you can implement and manage evaluation to inform your practice, facilitate decision-making on project-development teams, gather evidence of success, attract further funding, and most importantly, make a difference in the lives of all the visitors and project participants whom you touch through your work.

Chapter 2
Definitions and Principles:

Guiding Ideas for Principal Investigators to Know

LAUREN RUSSELL

Introduction

This chapter presents ideas to help you incorporate evaluation into your work. We include definitions, explanations, and principles that guide the use of evaluation in the process of designing, testing, and refining a project and then understanding its outcomes for visitors and participants. We'll define evaluation within the context of informal Science, Technology, Education, and Math (STEM) education, describe the purposes of evaluation, and discuss the challenges and opportunities that evaluation offers to practitioners of informal STEM education.

Evaluation plays an important role throughout your project. In the planning stages, an evaluator can help you clearly define your targeted outcomes and their connections to your project activities. As you implement your project, evaluation can reveal needed course corrections. At the conclusion of your project, evaluators can help you understand its outcomes.

Evaluators use a variety of data collection methods such as observations, interviews, focus groups, and surveys. These methods yield qualitative and quantitative data that can be used to make recommendations for project improvement, to assess project effectiveness, and/or to answer a wide range of project-specific questions.

The concrete end product of an evaluation is usually a formal report (or reports) that includes project background, study design and methods, data gathered from visitors and participants, and key findings and recommendations. You can view numerous examples of evaluation reports for ISE projects on *InformalScience.org/evaluation*.

At its best, evaluation works to answer questions that give a project team a deeper and richer understanding of its own practice—before, during, and after a specific project implementation. Therefore, an evaluation study should evolve and be guided by a project, while the project should be informed and guided by its evaluation.

Defining Evaluation

Evaluation has many definitions—you need to make it work for your project! Here are some formal definitions, along with more background on what questions evaluation can answer and a description of the differences between evaluation and research.

Evaluation in the Informal Science Education (ISE) field & other fields

The term "evaluation" encompasses a broad range of activities and purposes, so it's no wonder that ISE practitioners are often challenged to pin down a precise definition. Evaluation takes many forms—front-end, formative, and summative evaluation might be familiar terms—and each has a wide range of purposes and benefits. *Surrounded by Science,* a National Research Council report focused on learning science in informal environments, defines evaluation as a set of approaches and techniques used to make judgments about the effectiveness or quality of a program, approach, or treatment; improve its effectiveness; and inform decisions about

its design, development, and implementation (National Research Council 2010). In other words, evaluation for an ISE project generally provides information that can guide a project, suggest how it might be improved, and in the end provide evidence to demonstrate whether it worked as intended.

Evaluation became a prevalent practice and growing academic field in the 20th century when it was used on a widespread basis to assess social programs in education and public health (Rossi 2004). Evaluations are now used in an array of contexts to assess diverse programs in education, social services, organizational development, and public policy initiatives. Many journals and professional organizations are devoted to the broad field of evaluation, including the American Evaluation Association *(www.eval.org)*, which publishes the journals *The American Journal of Evaluation* and *New Directions for Evaluation*.

Looking beyond the field of informal STEM education, evaluation can be defined as the use of social research methods to systematically investigate the effectiveness, value, merit, worth, significance, or quality of a program, product, person, policy, proposal, or plan (adapted from Fournier 2005, Rossi 2004). *The Encyclopedia of Evaluation* further explains that conclusions made in evaluations encompass both an empirical aspect (that something is the case) and a normative aspect (judgments about the value of something) (Fournier 2005). The normative aspect explains why recommendations are often included in evaluation reports.

The Questions answered by evaluation

Evaluation answers three questions: What? So What? Now What?

A common pitfall when designing evaluations is the instinct to start by identifying preferred evaluation methods, for example, "What I want is a series of focus groups conducted with youth in the science afterschool program" (Diamond 2009). Evaluation planning should begin not by choosing methods but by defining questions that frame what you want to know from the overall study (not questions that might be asked of participants) (Diamond 2009). Then your evaluation questions can guide the choice of data collection methods.

Michael Quinn Patton, in his *Utilization-Focused Evaluation* (2008 p.5) states that in the simplest terms, evaluation answers three questions: What? So what? Now what?

What: What happens in the program? What services and experiences does the program offer? What activities and processes occur? What outcomes and impacts result? What unanticipated outcomes emerge? What are the program's documented costs and benefits?

So what: What do the findings mean? Why did the results turn out as they did? What are the implications of the findings? What judgments can be made? To what degree and in what ways can the program be considered a success? A failure? A mixed bag of positives and negatives? How does this program compare to other programs? What sense can we make of the findings?

Now what: What recommendations can be made from the findings? What improvements should be made? Should funding be continued, expanded, reduced, or ended? Should others adopt the program? What do findings from this project suggest for other or future projects? In short, what actions flow from the findings and their interpretations?

The Difference Between Evaluation and Research

Let's discuss an important question—what is the difference between evaluation and research? Many practitioners are confused by this question because research and evaluation share many of

the same methods for collecting and analyzing data, and many professionals lead both research and evaluation studies.

However, the purposes and the units of primary interest for research and evaluation are usually different. Much of educational research is designed to study a characteristic of learning grounded in an academic discipline such as psychology or sociology, or to study a particular theoretical framework. Research traditionally is geared toward knowledge generation and usually includes dissemination of findings through publication in peer-reviewed journals.

In contrast, the primary purpose of evaluation is to assess or improve the merit, worth, value, or effectiveness of an individual program or project and to advance the field (in this case, informal STEM education) by deriving lessons for funders, policymakers, or practitioners. Evaluation studies are generally conducted for clients and in collaboration with various stakeholders who are invested in improving or assessing a particular intervention, event, program, or activity.

TRUE STORIES

The complementary roles of evaluators and discipline-based researchers.

[Judy Diamond] Learning researchers and evaluators are using complementary methods to study the Life on Earth exhibit, a multi-user touch-table featuring an interactive visualization.

This interactive visualization of data from the Tree of Life (a web-based hierarchy of phylogenies representing over 90,000 nodes in the evolutionary tree) was developed by Harvard University in partnership with Northwestern University, University of Michigan, and the University of Nebraska State Museum in order to study strategies for engaging museum visitors in exploring the relatedness of all known species.

Previous research has shown that museum visitors initially have reasoning patterns that reflect a combination of intuitive reasoning about how life changes with some evolutionary knowledge and religious reasoning. Results from research studies with Explore Evolution indicate that a single visit to the exhibition can help visitors significantly shift their reasoning patterns to include more evolutionary reasoning. Moreover, visitors appear to do so in a predictable learning trajectory. Preliminary results from the Life on Earth exhibit component suggest similar findings.

The Life on Earth research team is investigating whether the experience of interacting with the multi-touch exhibit moves visitors along a gradient toward using evolutionary explanations more often. The team's discipline-based researchers focus on specific types of learners (pairs of youth aged 9-14), and they use comparison studies of groups randomly assigned to different conditions: for example, one condition involves using the multi-touch exhibit while another involves viewing a video about the Tree of Life. In contrast, the evaluation team uses a more naturalistic approach to assess the impact of the exhibit on visitors' behavior and attitudes. The evaluators examine how visitors use the exhibit as designed and implemented to see what people do and say when the exhibit is installed in a museum environment. The evaluation findings thus help the team understand how a range of people use and interact with the Life on Earth exhibit, providing context for the researchers' findings.

Stakeholders

Even within informal science education, evaluation has many stakeholders.

Many stakeholders benefit from evaluation, including project developers, project participants and their communities, and project funders.

The primary stakeholder is often the project team. Evaluation can help a team build a reflective practice throughout project development,

understand what audience impacts are occurring, strategically improve a project, and plan for future work.

Project participants and their communities are stakeholders because they are typically the project's direct beneficiaries. Evaluation findings often describe participant experiences and may inform future services and programs that will be available to them. The American Evaluation Association's "Guiding Principles for Evaluators" explains that evaluators must "articulate and take into account the diversity of general and public interests and values that may be related to the evaluation" (AEA 2012).

Funders such as the National Science Foundation are also key stakeholders in evaluation of the projects they help to bring about. Funders recognize the value of integrating evaluation into project development for the benefit of all stakeholders. Evaluations also help funders understand and describe the impact of their project portfolios and inform strategic decisions about investments (Friedman 2008).

In some cases stakeholders may hold conflicting opinions regarding the purpose of an evaluation. For example, on-the-ground practitioners may be most interested in learning how to make a program better (improvement-oriented formative evaluation), while funders may prioritize summative or accountability-focused evaluations. Therefore, stakeholders and evaluators must have open conversations to agree on the goals and intended purposes for a project's evaluation. Then the evaluator will be able to determine the best approaches and methods to carry out one or more studies. Sometimes data collected from participants can be used for multiple evaluation purposes.

Three Main Types of Evaluation

Formative Evaluation

Improvement-oriented formative evaluation focuses on learning how to improve or enhance a project (Patton 2012). Formative evaluation gathers data about a project's strengths and weaknesses with the expectation that both will be found and that the information can be used to make improvements.

SPOTLIGHT ON PRACTICE

Avoid the wrong mouse hole.

[Rachel Hellenga] Front-end research can help you avoid the wrong mouse hole. We overturned one key assumption about visitors to the Tech Museum with just a little research. I was working with an advisory group of retired engineers to develop approaches to conveying the concept of digital communication, and the advisors were firm in their position that we needed to present "analog" in order to teach about "digital." For example, to explain how a laser sends signals over fiber optic cable with on/off flashes of light, they recommended making a comparison to Morse code. This seemed logical so we tried out a brief questionnaire to find out what kids knew about Morse code. The answer was—nothing. Most of the kids in the target age range of 10-16 had never heard of it.

We tried rewording the question and prompting the kids in a number of ways, but still no luck. While we were very frustrated, we had actually had accomplished something very important, as this information steered us away from elaborate comparisons of analog and digital. The result was an engaging exhibit that involved:

- flipping a row of 8 light switches on and off to select the 8-bit code for a letter of the alphabet
- watching the letter get transmitted via a blinking light over fiber optic cable

- repeating this process to build up a message on an LED screen

No Morse code required!

The Front-end evaluation

Information-seeking front-end evaluation focuses on gathering information that informs project planning and development (Diamond 2009). Front-end evaluation often takes the form of audience research as it gathers data about the knowledge, interests, and experiences of the intended audience.

SPOTLIGHT ON PRACTICE

Refining a program with formative evaluation.

[Kirsten Ellenbogen] Years before CSI became a television hit, school groups visiting the Museum of Science and Industry could use scientific imaging tools to determine the identity of a girl who had fallen comatose while in possession of a $20 million lottery ticket. During the design phase of this facilitated Mystery Lab, our team cobbled together simple versions of each forensics activity, and produced a fake newspaper article to set up the story, and headed over to a local elementary school to try it out. These scanty props inspired whispered debates about whether the story was real! Never underestimate your ability to evoke the spirit of your project with a mock-up. (Figure 3)

We uncovered a strong interest in the activity but didn't realize the mystery was almost too interesting until after we launched the program. Once we were up and running in a fully equipped Mystery Lab (including a "news" video featuring a local anchor), we initiated a new round of formative evaluation. Our evaluator's interviews with program participants revealed that much of their conversation revolved around extraneous plot details such as the dollar amount of the lottery ticket. We revised the script to draw clearer connections to the science underlying the MRI scans, scanning electron microscopy, fingerprint analysis, and face aging software involved in the mystery—and as a result we saw a marked improvement in the summative evaluation. During the formative stage we also had an opportunity to address a concern about equal participation of boys and girls which resulted from our evaluator's critical review of the facilitated program. An evaluator can be a valuable "witness" during formative evaluation!

Figure 3

Summative Evaluation

Judgment-oriented summative evaluation focuses on determining a program's overall effectiveness and value (Patton 2012). Summative evaluation is particularly important in making decisions about continuing, replicating, or terminating a project, or providing lessons learned about informal STEM education for the broader field. Summative evaluations are often requested or required by funders, including the National Science Foundation.

SPOTLIGHT ON PRACTICE

Conducting summative evaluation mid-project.

[Ben Wiehe] In the first year of the Science Festival Alliance, evaluation was conducted at festivals held in Cambridge, MA and San Diego, CA. A fair number of attendees reported that they had not had an opportunity to interact with a science professional or engineer during the festival. This was a little surprising considering the vast number of STEM practitioners who were central to the events. Our evaluator determined that the numbers correlated with people who reported they had never knowingly interacted with a scientist or engineer, and we realized that although many attendees actually were interacting with a scientist or engineer at the festivals, they were simply not able to recognize who was a STEM practitioner. Each festival responded in its own way. The San Diego festival added "Ask A Scientist" and "Ask an Engineer" areas that invited attendees to pop in with a burning question. Organizers of the new Philadelphia festival had stickers made up that said "I am a scientist" or "I am an engineer" for participating STEM practitioners.

Qualities of Informal STEM environments

Principal Investigators, evaluators, and project teams need to understand and consider the special attributes of informal STEM education when planning and implementing evaluation studies. Understanding this context helps a team design evaluations that leverage the strengths of ISE settings and set reasonable and realistic goals and expectations. Key attributes of informal learning environments and experiences that present opportunities and challenges for evaluation include complexity, social experience, variety, and the fact that informal STEM education is an emerging field.

Complexity

Informal learning environments and experiences are complex.

Many informal STEM experiences are short, isolated, free choice, and self-directed. Often they reach target heterogeneous public audiences whose members come to the project with unique prior knowledge, interests, and experiences, and individual audience members learn different things, not just different amounts (Friedman 2008).

Challenge:

- Separating the effects of a single experience from a variety of other factors that could contribute to positive learning outcomes can be challenging (National Research Council 2009, 2010). This is true with many education interventions, but particularly so with informal learning environments.
- Establishing uniform evaluation activities, approaches, and methods that do not sacrifice a participant's freedom of choice and spontaneity can be difficult (National Research Council 2009, 2010).
- Experimental designs, where participants are studied in both treatment and control groups, may not be practical or the most appropriate method for evaluating many ISE projects. Therefore, conclusively attributing specific outcomes to a set of specific experiences or interventions is a difficult, often inappropriate, task for evaluation in the ISE context (Friedman 2008).

Opportunities:

- ISE environments allow evaluators and practitioners to consider a wide range of potential outcomes, including some that may be unanticipated during project design.
- Given that ISE experiences are learner driven, evaluations in ISE environments can be

designed to be learner centered.

- Because of the complexity of ISE settings, evaluators must respond creatively and flexibly with new instruments, methods, and approaches, which can advance the field of evaluation as a whole.

Social Experience

Many informal learning experiences are collaborative and social.

"Doing well" in informal settings often means acting in concert with others (National Research Council 2010). Participants may be motivated to engage in ISE with the primary goal of having a social experience, considering learning goals secondarily or not at all.

Challenge:

Teasing apart individual assessment from group process and accomplishments, especially in light of unanticipated outcomes, can be difficult (National Research Council 2009, 2010).

Opportunity:

Evaluation in the ISE context helps us better understand socially mediated experiences across family and multi-age groups. These insights add richness and depth to our understanding of how people learn through interaction and conversation, which subsequently helps us to design experiences that better support social interaction.

Variety

Informal STEM Education environments and experiences are exceptionally diverse. Intended audiences, settings, delivery methods, depth, expected outcomes, and other dimensions vary, and experiences may include exhibits in museum environments, television and radio programs, casual investigations at home, or afterschool programs.

Challenges:

- Participants may or may not be able to articulate personal changes in skill, attitude, behavior, or other outcomes at any stage of an informal learning experience. Therefore, evaluators may need to design instruments or other evaluation techniques that do not require or solely depend on self-articulation (Allen in Friedman, 2008).
- Connecting the dots between various evaluations to make generalizations about learning or best practices is complicated because of multiple unique contextual factors.

Opportunities:

- Many ISE environments allow for nimble and flexible evaluation settings. Especially at museums, visitors are abundant, and most are willing study participants.
- Because of the diverse contexts that surround informal learning, ISE evaluation is well positioned to draw on and contribute to theory, knowledge, and methods from a broad array of academic disciplines including psychology, learning sciences, cognitive sciences, formal education, sociology, public health and anthropology.

ISE is an emerging field

Informal STEM Education is a young, emerging field.

This presents a challenge because everything is new! But this newness presents even more opportunities for creative research and evaluation.

- Recent efforts have yielded new tools and resources for evaluators and practitioners that help to integrate evaluation into projects. These include new products (like this Guide and the *User-Friendly Handbook for Project Evaluation*); efforts to establish more consistent language and categories of impact (Framework for Evaluating Impacts of ISE Projects); and initiatives to store project outputs and evaluation reports in accessible and consistent places (such as *InformalScience.org*).

- Evaluation in ISE contexts helps us broadly understand how lifelong and informal learning opportunities are contributing to an informed citizenry and scientific workforce, areas of increasing focus and importance from a policy perspective.
- Evaluation in ISE contexts provides a unique contribution to our understanding of how people learn, which parallels and complements current research aimed at advancing knowledge within ISE and related disciplines.
- Growing collaborations among ISE project developers and evaluators present tremendous opportunities to develop innovative evaluation methods, to understand and disseminate effective practices, and to develop unified ISE evaluation theory.

Guidelines for professional practice

Evaluators and evaluation are informed by guidelines for professional practice. The American Evaluation Association (AEA) Guiding Principles for Evaluators (Figure 4), the Visitor Studies Association's Evaluator Professional Competencies (Figure 5, page 20), and the Joint Committee Standards for Education Evaluation are described in Figure 6, see page 21.

Evaluator professional competencies

The Visitor Studies Association has developed a set of five competencies that evaluators should have or demonstrate. The competencies are relevant to evaluators working in a variety of informal settings such as media, technology, and youth and community projects.

Guiding Principles for Evaluators

The AEA principles are intended to guide the professional practice of evaluators and to inform evaluation clients about ethical practices that they can expect their evaluators to uphold. American Evaluation Association (2004).

A. Systematic Inquiry: Evaluators conduct systematic, data-based inquiries.

B. Competence: Evaluators provide competent performance to stakeholders.

C. Integrity/Honesty: Evaluators display honesty and integrity in their own behavior, and attempt to ensure the honesty and integrity of the entire evaluation process.

D. Respect for People: Evaluators respect the security, dignity and self-worth of respondents, program participants, clients, and other evaluation stakeholders.

E. Responsibilities for General and Public Welfare: Evaluators articulate and take into account the diversity of general and public interests and values that may be related to the evaluation.

Figure 4

Ethical standards for Evaluation

The final set of professional guidelines, the Joint Committee Standards for Educational Evaluation (JCSEE 2011), include five ethical standards focused on the evaluation, as opposed to the AEA principles, which are focused on the evaluator.

Each standard is articulated in sub-statements and descriptive text, but in brief, the five standard categories are listed in Figure 6, see page 21.

Evaluator Professional Competencies

Visitor Studies Association (2008)

Competency A. Principles and Practices of Visitor Studies

Evaluators should be familiar with the history, terminology, past and current developments, key current and historic publications, and major contributions of the field. Evaluators should also be familiar with educational theory, environmental design, developmental psychology, communication theory, leisure studies, and marketing research.

Competency B. Principles and Practices of Informal Learning Environments

Evaluators must understand the principles and practices of informal learning, the characteristics that define informal learning settings, and an understanding of how learning occurs in informal settings. An understanding of the principles, practices, and processes by which these experiences are designed or created is required in order to make intelligent study interpretations and recommendations.

Competency C. Knowledge of and Practices with Social Science Research and Evaluation Methods and Analysis

Evaluators must not only understand but also demonstrate the appropriate practices of social science research and evaluation methods and analysis. These include: research design, instrument/protocol design, measurement techniques, sampling, data analysis, data interpretation, report writing and oral communication, human subjects research ethics, and research design, measurement, and analysis that shows sensitivity to diversity and diversity issues.

Competency D. Business Practices, Project Planning, and Resource Management

Evaluators must possess appropriate skills for designing, conducting, and reporting evaluation studies. They should demonstrate their ability to conceptualize an evaluation project in terms of scheduling, budgeting, personnel, and contracting.

Competency E: Professional Commitment

Evaluators should commit to the pursuit, dissemination, and critical assessment of theories, studies, activities, and approaches utilized in and relevant to visitor studies. Through conference attendance and presentations, board service, journals and publications, and other formal and informal forums of communication, evaluators should support the continued development of the fields of informal science education and evaluation.

Development and implementation of the Visitor Studies Professional Competencies was supported in part by grant No. 04-43196 of the Informal Science Education Program of the National Science Foundation.

Figure 5

Conclusion

While the positive benefits of evaluation are enormous, practitioners and evaluators must wrap some perspective around the limits of evaluation. For the most part, evaluation findings address only the questions that were originally asked by a project team. However, Principal Investigators, practitioners, and evaluators must interpret evaluation data, results, and findings in light of broader circumstances and contexts. Evaluation findings themselves do not directly make recommendations or decisions; rather, they inform recommendations and decisions. These limits of evaluation point again to the importance of collaboration and communication between evaluators and practitioners before, during, and after a project is designed and implemented.

Joint Committee Standards for Educational Evaluation (JCSEE)

(2008)

Utility Standards are intended to increase the extent to which program stakeholders find evaluation processes and products valuable in meeting their needs.

Feasibility Standards are intended to increase evaluation effectiveness and efficiency.

Propriety Standards support what is proper, fair, legal, right, and just in evaluations.

Accuracy Standards are intended to increase the dependability and truthfulness of evaluation representations, propositions, and findings, especially those that support interpretations and judgments about quality.

Evaluation Accountability Standards encourage adequate documentation of evaluations and a meta-evaluative perspective focused on improvement and accountability for evaluation processes and products.

Figure 6

This page intentionally left blank.

Chapter 3
Choosing an Evaluator

Marching Project Needs with Evaluator Skills and Competencies

MARY MARCUSSEN

Introduction

This chapter is about locating an evaluator well matched to your project needs. Whomever you select will influence the shape of your evaluation plan and the results that you get from it. We'll start by looking at some choices you'll need to make while searching for an evaluator. Next we'll provide an orientation to the expertise and skills that your evaluator should possess. After that we'll share some issues to keep in mind when adding an evaluator to your project team. Finally, we'll suggest some ways to locate an evaluator.

Identify your project requirements

Before we get started, imagine setting up your best friend with a blind date. First you critique each candidate's credentials, background, interests, job experience, credibility, social connections, and, of course, interest in dating your friend. These factors combine to shape your perception of candidates and, ultimately, your final selection. After the big night, you are eager to learn whether your friend found a good match. You have invested significant time and effort in your role as matchmaker, and you hold yourself personally liable for the outcome of the relationship. Your friend is depending on you, and you desperately need this blind date to "work out."

Choosing an evaluator who matches the needs of your project is much like setting up a blind date. Before considering the skills and qualifications of a potential evaluator you must first establish some basic parameters for your project. To skip this step would be like selecting a date for your best friend without knowing what type of person interests them. Or like interviewing candidates for a job without a job description.

For example, you will need to decide if your evaluator should be experienced in formative or summative evaluation or should be someone who can do both types. You'll also want to understand your options for hiring an internal (in-house) as opposed to an external evaluator; for choosing an evaluation firm vs. an independent contractor; and for working with a local evaluator vs. someone located across the country. You also need to consider any special skills you may require: "do you need an evaluator who speaks another language? Or one who is familiar with a particular culture or subject matter?

If this process seems a far stretch from setting up a blind date, think of your project as your best friend (for whom you will do anything) and the evaluator as the date (whom your friend needs to be happy). After you read the following example of a project team choosing an evaluator, we will look critically at each of the decisions made throughout the process.

TRUE STORIES

Articulate your project's evaluation needs before hiring.

In a project to revitalize static wildlife dioramas created in the 1950s, the Oakland Museum of California aims to connect visitors more deeply to urgent environmental issues. To help minority audiences in particular make such connections, the museum involved a diversity of visitors in co-design and development of the

renovated exhibits.

Hotspot California: Bringing Dioramas to Life Through Community Voices is a project of the Oakland Museum of California that is funded by the Informal Science Education program at the National Science Foundation (DRL 09-15778). The project's evaluation needs were many and varied. First, like most informal science education projects, this one required multiple stages of evaluation, and for the project's summative evaluation, the team required an evaluator with experience in exhibit design and a broad perspective of the museum community. Second, to build institutional evaluation capacity, the museum wanted to involve its staff in the evaluation process. Third, the project team wanted to conduct formative evaluation with exhibit prototypes in a variety of museum settings beyond Oakland. Finally, the project team wanted to examine the effects of visitor involvement in the process of exhibition design.

With a tight budget and timeline, the project team made some critical decisions. First, they selected an independent evaluator to 1) conduct front-end evaluation by collecting baseline data from members of the intended audience, 2) conduct formative evaluation to determine the effectiveness of the exhibit prototypes in helping visitors develop deeper connections to their local environments, and 3) train museum staff to assist with the front-end audience study and with exhibit prototyping and formative evaluation. This particular evaluator had experience with the museum's intended inner city audiences.

In addition, the project team selected a second independent evaluator to conduct summative evaluation. This evaluator was widely recognized for her expertise in evaluating museum exhibits.

Let's tease apart the process by which the project team made its selections.

Four Fundamental Choices

Your project will have certain qualities and requirements that will influence your choice of evaluator. Here are some of the parameters to consider:

1. What type of evaluation do you need?
2. Do you need an internal or external evaluator?
3. Do the requirements of the project lend themselves to an independent contractor or an evaluation firm?
4. Will your project be better served by a local or out-of-area evaluator?

Consideration #1: What type of evaluation do you need?

Different types of evaluation require different sets of evaluator expertise and skills. Depending on the goals of your project, you may need to find one "date" who runs cross-country marathons and another who holds season tickets to the ballet. Or someone who does both. In particular, front-end, formative, and summative evaluation each require different approaches. Some evaluators conduct all three types of evaluation while others are particularly skilled at one type. These three types of evaluation were defined in Chapter 2 and they bear repeating here in the context of choosing an evaluator.

Front-end evaluation

Front-end evaluation gathers background information that informs practitioners as they plan and develop a project. Such information-seeking evaluation often takes the form of audience research, gathering data about the knowledge, interests, and experiences of a intended audience.

So, it's important for a front-end evaluator to have an understanding of or prior experience with your intended audience.

Formative evaluation

Formative evaluation focuses on learning how to improve or enhance a project. Such improvement-oriented evaluation gathers data about a project's strengths and weaknesses with the expectation that both will be found and the information can be used to make improvements.

It can be helpful if your formative evaluator has experience in your particular type of project, for example, film or exhibit production, or educational technology, or professional development for informal STEM education practitioners.

Summative evaluation

Summative evaluation determines a project's overall effectiveness and value. Such judgment-oriented evaluation can be conducted at interim points or at the end of a project, and is particularly important in making decisions about continuing, replicating, or terminating a project. Summative evaluations are often requested or required by funders, including the National Science Foundation.

Deciding which data are essential and determining how much data can be collected requires experience and careful thought.

TRUE STORIES

Choosing evaluators skilled in a specific type of evaluation.

In the case of Hotspot California, the Oakland Museum of California chose to establish formative and summative evaluation as separate roles.

Hiring separate evaluators for the different phases of evaluation is not always necessary, and working with multiple evaluators requires developing a collaborative evaluation plan. For this project, however, the two evaluators brought complementary skills and experience. Both understood natural history dioramas and had evaluated them previously. However, one evaluator understood the intended inner-city audience while the other had a deep experience with the type of summative evaluation that the museum wanted to conduct.

Consideration #2: Should you choose an internal evaluator or an external evaluator?

Another important decision is whether to engage an internal (in-house) or external evaluator or to employ an internal evaluator who works with an external consultant. While a funder may dictate this decision—the National Science Foundation, for example, typically requires an external evaluator for summative evaluation—Principal Investigators and their institutions often spend a great deal of time weighing the pros and cons of each option.

Internal evaluator: pros and cons

Internal staff are typically more accessible for team meetings and their cost may be covered by an institution's operating budget (CPB 2011). In-house evaluators also are familiar with the culture of the organization and the project team and may have working knowledge of the project's subject matter. On the other hand, internal evaluators may be invested in the outcome of the evaluation, caught up in internal politics, or hampered by supervisory relationships.

External evaluator: pros and cons

External evaluation is typically more expensive. However, the results often bear more weight because the evaluator is considered to have no vested interest in the project outcomes. An external evaluator also may be able to work more independently from the producers of the project deliverables (Kellogg 2004).

Combining internal and external evaluators

Some projects blend these approaches by having internal staff conduct evaluations under the guidance of an independent evaluator who reviews the evaluation design and assesses the validity of the findings and conclusions. This approach can maintain external expertise and impartiality along with the benefit of an internal person's first-hand project knowledge.

TRUE STORIES

Teaming up internal and external evaluators.

The Oakland Museum hired external evaluators, and had them train museum staff to become deeply involved in the evaluation process during the course of the Hotspot California project. If this approach helps the museum to develop internal evaluation expertise, in the future the museum may be able to limit the role of external evaluators to independent oversight.

Comparison of internal and external evaluators.

The corresponding chart weighs a variety of factors when considering internal versus external evaluation. Comparison of Internal vs. External Evaluators (adapted from Conley-Tyler 2005, Kellogg 2004, and Patton 2008). (Figure 7, see page 28.)

Consideration #3: Should you choose an evaluation firm or an independent contractor?

The question of whether to hire an evaluation firm or an independent contractor hinges on several factors.

Evaluation firm: pros and cons

A robust evaluation can require a cadre of well-trained staff (data collectors, transcriptionists); the necessary equipment (data storage, statistical software, cameras, recorders); and infrastructure (office space, supplies). A project team that requires such support may look to an evaluation firm and should factor in the potential for higher institutional costs.

Independent contractor: pros and cons.

Conversely, you may be looking for the nimbleness of an individual contractor who is your direct contact for all aspects of the project operates with lower overhead. In that case you may need to determine whether your project can be evaluated without the more extensive personnel and resources needed for complex studies.

TRUE STORIES

Deciding on an independent contractor vs. an evaluation firm.

In our example from Hotspot California, the Oakland Museum needed a formative evaluator who could respond quickly and efficiently to project developments and who could travel last minute for timely data collection. The museum also required a summative evaluator who could perform data analysis in cooperation with the formative evaluator. In this case, the nimbleness of two independent contractors allowed for timely, efficient, and cooperative evaluation.

Consideration #4: Should you choose a local or an out-of-area evaluator?

When choosing an evaluator you also need to consider the location of potential evaluators relative to your institution and any project partners. If your project is intended to reach your local or regional community, working with a nearby evaluator who understands your local audience and issues can be sensible. However, many informal STEM education projects are collaborations between multiple institutions located in different states or regions of the country. In this case you may not need to be constrained by location in choosing your evaluator as long as you build in the resources that are necessary to support a long-distance evaluator to travel to your project location or locations. Communication among team members via wiki, teleconferences, or Skype can also allow evaluators located cross-country to participate actively in the team.

For the project at the Oakland Museum of California, the project team did not see proximity as an issue when selecting two evaluators from out of state. While the issue can boil down to the cost of a local vs. long-distance evaluator's travel to conduct the work, this issue should be considered in the context of factors that may override location.

Comparison of Internal vs. External Evaluators
(adapted from Conley-Tyler 2005, Kellogg 2004, and Patton 2008)

	INTERNAL EVALUATOR	EXTERNAL EVALUATOR
Expertise	Internal evaluators work in the environment in which the project operates and may have firsthand knowledge of the project, content, and organizational policies and practices.	External evaluators may possess special skills or exposure to a wide range of methods and practices that would be useful to incorporate.
Perceived bias	There may be a perception of bias if the internal evaluator is "too close" to the subject matter.	Perceived impartiality is a strong argument for the use of external evaluators.
Availability	Staff evaluators are readily available for project meetings or spontaneous data-collection activities.	Local evaluators can be readily available or can use telecommunications when needed.
Cost	Internal evaluators on salary can have an advantage over external evaluator fees. However, it can be expensive to maintain an idle staffer in between projects.	External evaluation fees can be high compared to salary, but can be cost-effective when the evaluation is needed only part time or for a limited duration.
Organizational investment	Over time, an internal evaluator can build an organization's capacity to support evaluation. However, this might not be a priority for an organization that will conduct evaluation on an infrequent basis.	External evaluators can acquaint staff with the value and methods of evaluation and can train staff in data-collection techniques. This can build a culture of evaluation within an institution.

Figure 7

Evaluator Qualifications

Once you have a handle on some of the fundamental choices involved in selecting an evaluator, you are ready to begin thinking about the actual evaluator you need. But before scheduling interviews with prospective evaluators, sit down with your project team and identify the skills that are particularly important for your project. Experience shows that the most important overall characteristics of a successful evaluator are the abilities to remain flexible and to solve problems (Kellogg 2004). In addition to these key traits, consider the following criteria: education and background experience; content expertise and experience with similar projects; ability to handle multiple project deliverables; experience with the audience that you are serving; and any unusual aspects of your project which might demand special skills.

Education and background experience

Most professional evaluators have at least a Master's Degree in evaluation or a field related to evaluation, including science and social science. Many have variable backgrounds such as Peace Corps volunteers, educational researchers for technology firms, or senior staff of major conservation organizations. Such experience can enrich your project's content development or audience connections. Look for knowledge and attitudes regarding evaluation that suggest a compatibility with your project and evaluation goals.

Query potential candidates about their experience in the areas of evaluation design, data collection, and data analysis. Many evaluators have expertise in specific areas such as ethnographic research, statistics, outcomes-based evaluation, timing and tracking, focus groups, bilingual evaluation, or participatory evaluation. This information is crucial to know, depending on the needs of your project and the roles that you have envisioned for your evaluator. Focus group facilitators need to be able to manage groups. Interviewers must be supportive and skilled listeners. Ask candidates about their expertise in qualitative, quantitative, and mixed methods evaluation. While some evaluators have a preference for one particular technique, a combination of approaches is likely to provide the most useful information.

Ask prospective evaluators what they need to know about your project goals, objectives, and desired outcomes before they can determine appropriate evaluation approaches and methods. Beware of an evaluator who assures you that they know how to evaluate your project before learning enough about it to determine an evaluation plan!

Content expertise and experience with similar projects

Locating a top-notch evaluator with solid credentials is more important than finding an evaluator who knows the specific content area of your project. Consider generalists who are able to grasp your project quickly or specialists who are aware of their subject biases.

Regardless of content expertise, your evaluator should be trained in the evaluation of projects similar to yours or have a track record of completing successful evaluations of similar projects. For example, evaluators will have differing experiences in youth and teen programming, citizen science, family learning, social technologies, virtual worlds, radio, gaming, planetarium shows, live theater, or communities of practice.

Ability to handle multiple project deliverables

Some informal science education projects combine several types of deliverables. Cecilia Garibay (2008) discusses the complexities of evaluating the outcomes of such projects in the Framework for Evaluating Informal Science Education Projects (Friedman 2008, p. 96). It is common, for example, to see NSF-funded exhibitions that include related educational programming, or a television series with an accompanying

educational website, or collaborative projects among organizations that include components for both public and professional audiences. In some cases a suite of integrated components is designed to work together as a whole to achieve impact. If you are developing such a project, your evaluator must devote resources to evaluating each of the pieces. Understanding how the components interact is critical to developing appropriate evaluation strategies that accurately measure the outcomes of each. For example, your evaluator might suggest that one innovative deliverable is critically evaluated while others receive less review.

Cultural Competence

Evaluators interact with a broad range of people from many political, religious, ethnic, language, and racial groups and need special qualities to conduct culturally competent work. Frierson, Hood, Hughes, and Thomas state in *The 2010 User-Friendly Guide to Project Evaluation* (NSF 2010a, p. 75): "Culturally responsive evaluators honor the cultural context in which an evaluation takes place by bringing needed, shared life experiences and understandings to the evaluation tasks at hand and hearing diverse voices and perspectives. The approach requires that evaluators critically examine culturally relevant but often neglected variables in project design and evaluation. In order to accomplish this task, the evaluator must have a keen awareness of the context in which the project is taking place and an understanding of how this context might influence the behavior of individuals in the project."

The American Evaluation Association affirms the significance of cultural competence in evaluation, stating: "To ensure recognition, accurate interpretation, and respect for diversity, evaluators should ensure that the members of the evaluation team collectively demonstrate cultural competence. Cultural competence is a stance taken toward culture, not a discrete status or simple mastery of particular knowledge and skills. A culturally competent evaluator is prepared to engage with diverse segments of communities to include cultural and contextual dimensions important to the evaluation. Culturally competent evaluators respect the cultures represented in the evaluation throughout the process." (AEA 2011)

Your evaluator will need to develop a trusting relationship with the audience for your project and should have experience in doing so. You will want to locate an evaluator who has developed an understanding of your intended audience and the context in which your project will be implemented.

Cosmic Serpent, a project of the Indigenous Education Institute and UC Berkeley Space Sciences Laboratory, provides an excellent example of a culturally responsive evaluation.

TRUE STORIES

Blending Native American and Western evaluation methods.

A national evaluation firm collaborated with a Native American consultant to provide evaluation for Cosmic Serpent (DRL 07-14629 and 07-14631). The project was conceived to build capacity among museum educators to bridge native and western science learning in informal settings. The team developed an evaluation design using the Diné (Navajo) model, in which native and western evaluation methods are equally valued and respected. All aspects of front-end, formative, and summative evaluation were then conducted collaboratively between the evaluation firm and the Native evaluator. Instrument design, data collection, analysis, and interpretation were all conducted collaboratively, with validation provided by each party.

This collaborative partnership allowed the team to take multiple viewpoints into account and to explore issues surrounding the cultural context of educational evaluation. This process increased the capacity of the indigenous evaluator, who is experienced with the evaluation

of Native populations, and of the researchers at the evaluation firm, who have experience in museum and professional development evaluation and the assessment of informal science learning. The front-end and summative evaluation reports for Cosmic Serpent are posted on the Informal Science database.

Other special situations and skills

In addition to planning for culturally responsive projects and projects with multiple components, evaluators often need special skills and competencies to deal with challenging situations (Kellogg 2004).

Familiarity With Standards and Guidelines for the Field

Your evaluator should demonstrate familiarity with the best practices in designing and evaluating informal science learning. Guideposts include the *Framework for Evaluating Impacts of Informal Science Education Projects* (Friedman 2008), *Learning Science in Informal Environments* (NRC 2009), *Surrounded by Science* (National Research Council 2010), and the *User-Friendly Handbook to Evaluation* (Westat 2010). Familiarity with the competencies of evaluation articulated by the Visitor Studies Association (VSA 2008) and the principles of evaluation espoused by the American Evaluation Association (AEA 2004) and the Joint Committee Standards (JCSEE 2011) also are important. An evaluator knowledgeable about these resources will be able to help you articulate the objectives and outcomes for your project and develop a project logic model as part of an appropriate evaluation plan.

To get a clear sense of an evaluator's work, request evaluation reports that he or she has previously prepared. Are they readable and understandable? Do they meet your expectations and standards? You also can evaluate an evaluator's professionalism by factors such as the appearance of their website and the quality of their written communications. And check references!

Challenges Requiring Special Evaluator Skills

(Adapted from M. Q. Patton, 1997, Utilization-Focused Evaluation, p. 131. In W.K. Kellogg Foundation 2004, W. K. Kellogg Foundation Evaluation Handbook, p. 61.)

SITUATION	CHALLENGE	SPECIAL SKILLS
Highly controversial issue	Facilitating different points of view	Conflict-resolution skills
Highly visible project	Dealing with project publicly; reporting findings in media-circus atmosphere	Tolerance for ambiguity; rapid responsiveness, flexibility, quick learner
Highly volatile project environment	Adapting to rapid changes in context, issues, and focus	Cross-cultural sensitivity, skilled in understanding and incorporating different perspectives
Evaluation attacked	Preserving credibility	Calm, able to stay focused on evidence and conclusions
Corrupt project	Resolving ethical issues/upholding standards	Integrity, clear ethical sense, honesty

Figure 8

Communication style with colleagues and funders

Methodological knowledge is not sufficient to conduct and report on a high-quality evaluation. Evaluators also require skills in stakeholder involvement, contract management, and written and oral communication. Concise summaries and creative use of electronic media are important means of delivering evaluation findings. Evaluators also should demonstrate willingness to have their work vetted by colleagues and to respond to their critiques.

Perspectives of an evaluator

Up until this point, as matchmakers we have focused on the needs of your best friend, i.e., your project. But what are your potential dates looking for? Both parties must contribute to a perfect match. It's time to consider what evaluators seek in a project.

Above all is the ability to conduct their work within a professional context in which the project developers have clear goals and objectives and understand enough about the process of evaluation to support and value their work. Team cohesiveness is another factor. Your evaluator needs team members who can facilitate the project work plan, develop the deliverables, nurture the project partners and advisors, and communicate with funders. Some evaluators look for projects that will further their professional interests, for example early childhood science education, educational media, or cultural science learning. A firm plan for dissemination in terms of publications and conference presentations enables them to share their work with the field.

Locating an evaluator

Now that you have a good search image for your evaluator, it's time to complete your role as matchmaker by finding one. There are many ways and places to locate qualified candidates. While not quite as easy as Match.com makes it seem, below are some strategies related to informal science education.

Evaluator databases and listservs

Look at evaluator databases and related listservs. There are several online resources which can assist you in locating an evaluator. Note that the Informal Commons web infrastructure developed by CAISE (*informalcommons.org*) provides a universal search engine across many of these resources.

- InformalScience.org maintains an updated list of ISE evaluators. Each entry includes the evaluator's affiliation, professional bio, and interest and expertise descriptors, along with selected research and publications.
- Visitorstudies.org maintains a "find an evaluator" database.
- The American Association of Museums Committee on Audience Research and Evaluation (AAM-CARE) publishes a directory of evaluators for AAM-CARE members.
- The Visitor Services in Museums Listserv (VSMUS) operates as a forum to bring together museum professionals and others concerned with the quality of the visitor experience in museums.
- EVALTALK is a listserv of over 2,000 members hosted by the American Evaluation Association.

Network with other NSF grantees

Principal Investigators who have implemented projects similar to yours may be able to suggest evaluators who will be a good fit with your project. A strong personal recommendation and a discussion of an evaluator's strengths and weaknesses from someone who has worked with that individual can be extremely useful (NSF 2010a, pg. 128). Several resources can be pursued:

- InformalScience.org project database and member directory, ISE Project Spotlights and the ISE Evidence Wiki
- Exhibitfiles.org

Read the evaluation literature

Find evaluation studies of projects like your own. Peer-reviewed journals with articles based on evaluations of informal science education projects include: Curator: The Museum Journal, Journal of Museum Education, International Journal of Science and Education, International Journal of Learning and Media, Afterschool Matters, Cultural Studies in Science Education, Journal of Research in Science Teaching, Journal of the Learning Sciences, Science Education, Studies in Science Education, Visitor Studies, American Journal of Evaluation, Evaluation, New Directions for Evaluation, and Evaluation and Program Planning, among others. The Research2Practice.org website contains a set of briefs summarizing recent peer-reviewed educational research. The briefs are written with the interests, needs, and institutional settings of informal science educators in mind.

Review the conference presentations

Finally, look through agendas of professional conferences that focus on informal science education, for example, those of the Association of Science-Technology Centers, American Alliance of Museums, Association of Zoos & Aquariums, American Evaluation Association, and Visitor Studies Association. Contact the evaluators to discuss your project.

Putting the team together

It is unlikely, but not impossible, that you will find a blind date who meets all of your best friend's dreams. It is also unlikely that you will find an evaluator who is representative of your intended audience, knowledgeable about your specific content area, experienced with your type of proposed deliverables—and available. Most important is locating an evaluator whose skills and experience, along with those of the other people involved with the project, create a cohesive and well-rounded team. And be sure that you and your team can work with your evaluator and enjoy the experience!

TRUE STORIES

Rounding out a project team with the right evaluator.

An example of efficient team building is illustrated by a professional development project of the Astronomical Society of the Pacific: Astronomy From the Ground Up (DRL 04-51933) was designed to build the capacity of informal STEM educators in science museums to deliver astronomy to their visitors more effectively. The PIs included astronomers and astronomy educators from the Astronomical Society of the Pacific and the National Optical Astronomy Observatory, along with the Director of Exhibitions, Research, and Publications for the Association of Science-Technology Centers. So the science content, pedagogy, and representation of the intended audience were covered. What the project needed was an evaluator familiar with the culture of informal STEM educators who could gauge their experience with the project and determine the impact on their daily work. The evaluator also needed the capacity to assess the project at multiple sites and on multiple levels including on-site and distance learning workshops.

The team selected an evaluator whose background as a planetarium educator and whose experience as an educational researcher in science museums provided the expertise necessary to work closely with astronomy educators. This evaluator—assisted by a team of researchers at his firm—conducted the front-end, formative, and summative evaluation, with independent validation of the summative evaluation design provided by an external consultant.

Conclusion

Once you've put together a good team, you are well on your way to carrying out an accurate and valuable evaluation. Keep in mind that an important part of an evaluator's job is to assist in building the skills, knowledge, and abilities of other project team members and stakeholders.

This page intentionally left blank.

Chapter 4
Working as a Team:
Collaboration Through All Phases of Project Development

JESSICA J. LUKE, STEVEN YALOWITZ AND SASHA PALMQUIST

Introduction

Let's assume that you have chosen an evaluator to join your project team. What does the evaluator need from you to achieve a successful evaluation? What should you expect from your evaluator? What roles will each of you play throughout the evaluation process?

Working together through-out all phases of your project

To answer these questions we will walk you through the life of a project as we discuss how Principal Investigators and evaluators work together at key phases, including:

1. Proposal writing
2. Project start-up and front-end evaluation
3. Formative evaluation
4. Summative evaluation
5. Dissemination of evaluation results

As we move through these five phases we'll offer a list of PI and evaluator roles and responsibilities gleaned from evaluating dozens of informal science education projects over many years. We'll explore how PIs and evaluators can work together to effectively manage relationships, leverage expertise, set realistic expectations, and ensure effective communication.

Proposal writing

Ideally you should identify an evaluator before you begin writing a proposal to develop your project. When you do, you can work with your evaluator to design a rigorous evaluation plan that serves as a key element of the proposal, is tightly aligned with project activities, and will provide feedback to inform project development and the achievement of targeted project outcomes.

Principal Investigator Responsibilities:

- Conceptualize and write the proposal; clearly articulate the project design, goals and outcomes, intended audiences, activities, and implementation plan
- Position the evaluator as a team member from the start; include the evaluator in team meetings and discussions; share iterative drafts of the proposal
- Work collaboratively to review the evaluation plan (i.e., ask questions, ensure the plan will provide the data that the project needs, ensure reporting will happen at key decision points).

Evaluator Responsibilities

- Help the PI to articulate and refine project goals and outcomes and to ensure that they are well-aligned with the project activities (this often takes the form of a logic model)
- Design an evaluation plan that articulates a) questions that will guide the evaluation; b) the

approach that will frame the evaluation activities; c) data collection methods that will be used to answer evaluation questions; d) a budget and timeline for the evaluation; e) a plan for working with the project team; and f) a dissemination plan for sharing information with the PI and the public

- Bring knowledge of trends in the field and literature that can help to align the proposal within relevant field(s).

TRUE STORIES

Collaboration during proposal writing.

LEAP into Science is a partnership between the Franklin Institute Science Museum (FI) and The Free Library of Philadelphia.

The project integrates science content and inquiry into an existing afterschool program at the Library, with three overarching goals: 1) to increase the capacity of afterschool facilitators for science teaching and learning; 2) to increase the capacity of libraries for science teaching and learning; and 3) to understand the ways in which science and literacy can be connected to promote children's learning and family engagement in both subject areas. During the proposal writing phase, FI worked closely with an independent evaluation firm to design a project evaluation plan that was grounded within a relevant theoretical framework. At that time the project was targeting youth, so the evaluators used a framework from the positive youth development literature to conceptualize evaluation measures, which in turn informed the articulation of project outcomes. The evaluators had many phone conversations with FI staff as the proposal took shape, and they reviewed two different drafts of the proposal as it moved into final form. When the project was funded the evaluators were ready to dive into it because they'd been part of the team from the outset.

Working together during Project Start-Up and/or Front-end Evaluation

As a project begins it's common to hold a kick-off meeting that allows your team to flesh out and adjust project activities and intended deliverables. At this time you and your evaluator must establish expectations and procedures for working together as project development progresses. Will the evaluator attend team meetings? If so, which ones? Will you hold standing meetings to check in on the evaluation? To what extent do you want to play a role in evaluation implementation? What does the evaluator need from you in order to begin work?

Once these questions have been addressed, work often begins on front-end evaluation that guides project development—for example, by examining potential participants' understanding of content, current behaviors, and misconceptions.

Principal Investigation Responsibilities

- Ensure that the evaluator has the most up-to-date copy of the project plan (sometimes plans change during the proposal-negotiation process) and any adjustments to goals, objectives, budget, timeline, or staffing
- Establish the team dynamic and articulate communication strategies
- Support the evaluator by encouraging his/her active participation within the project team and by facilitating team buy-in to the evaluation process
- Review the evaluation plan with the evaluator and make any needed adjustments

Evaluator responsibilities

- Review the evaluation plan and ensure its continued alignment with project activities and outcomes. In coordination with the PI, make any needed adjustments

- Develop a detailed work plan for the first phase of the evaluation that specifies when evaluation activities will occur, who will be involved, and how results will be communicated
- Encourage decision making that is grounded in data or best practices, research and evaluation literature, and/or relevant projects that have emerged during the time between proposal and award letter/start date
- Identify any assumptions being made by the project team, and encourage discussion about their implications for the project
- Confirm data sources and availability

TRUE STORIES

Collaboration during front-end evaluation.

To illustrate how PIs and evaluators can work together during the front-end phase consider The YardMap Network: Social Networking for Community Science (DRL 0917487) led by the Cornell Lab of Ornithology. This citizen science project is designed to cultivate a richer understanding of bird habitat for both professional scientists and people concerned with their local environments (www.yardmap.org). Participants can map their yard or other public spaces, showing buildings, yards, and landscaped areas, and can indicate tree and plant species as well as bird-friendly features. For this project the evaluators conducted front-end evaluation to determine the extent to which the intended audiences of birders and gardeners were interested in the idea of YardMap and to provide specific feedback to inform both project design and content. As a result some of the team's assumptions about the project were confirmed while other findings showed that planned approaches needed to be modified to fully motivate participants. Many design decisions about website appearance and functions were also directly informed by the front-end findings, including an online tutorial, appearance of the main page, instructions, and options for drawing yards.

The front-end study was a collaboration between the evaluators, PI, and the rest of the project team, ensuring that the results would be useful for all. For example, the original project plan included a front-end web survey with a sample of 300 to 400 participants. However, the PI saw an opportunity to use experimental methods to assess whether references to climate change influenced individuals' interest in participating in YardMap. This approach required significantly customizing the web survey and increasing the sample to more than 3,000 participants. The experiment was successful, and the PI and evaluators co-authored a journal article using the results.

Working together during formative evaluation

Formative evaluation focuses on ways of improving and enhancing projects. In this phase you and your evaluator need to work hand-in-hand as project components and/or activities are developed, tested, implemented, and reflected upon. The cycle of development and evaluation may be repeated several times as you make refinements.

Key roles and responsibilities in the formative evaluation phase:

Principal Investigator Responsibilities

- Develop products or program elements and clearly identify what needs to be learned about how they are received (i.e., does the intended audience understand the main ideas? Do the participants find the activities engaging? Do the products or elements function as intended?).
- Create forums for the evaluator and project developers to work together with shared purpose
- Identify internal evaluation expertise that might be leveraged for the project
- Clarify where building institutional capacity is a priority so the evaluator can work to train staff in data collection.

Evaluator responsibilities

- Be responsive to PI's needs relative to what they want to know and what they want to test, with whom, and when
- Engage in an iterative testing process that provides feedback to the PI in a timely and useful manner
- Support data interpretation and provide broader context for results as needed.

TRUE STORIES

Collaboration during formative evaluation.

Cosmic Serpent serves as a useful example of how PIs and evaluators can work together during the formative phase. This collaborative research project funded by the National Science Foundation (DRL-0714631 and DRL-0714629) focuses on building respectful, sustainable relationships between science museums and Native communities and/or tribal museums. Cosmic Serpent aimed to support museum practitioners in connecting Native and Western science learning in informal settings; creating awareness of the value and integrity of Native science paradigms among museum practitioners; and nurturing a process of "collaborating with integrity" that embraces and values multiple worldviews. The primary components of the project were a series of intensive weeklong professional development workshops, a culminating conference, and a legacy document that shared project outcomes and lessons learned with the broader field. A joint evaluation approach was used to model the type of cross-cultural collaboration that the project itself was designed to support.

At the heart of a Native worldview is relationship, which served as the guiding principle of Cosmic Serpent. To create a balance of multiple perspectives and to ensure validity of the data, evaluators needed to understand the importance of their own relationship to the community being served by the project and to the project being evaluated. Therefore, the evaluation team participated in almost all project activities and became an integral part of the project community, gathering feedback from project participants (Fellows), reflecting the community voice back into the planning process, and sharing insights and processes as participant-observers. The success of Cosmic Serpent's core team collaboration also depended upon building relationships through presence, participation, openness, and trust. To support this process, PIs, advisors, and evaluators created more in-person meetings than originally anticipated. However, these meetings were critical to the success of the project because they enabled different cultural worldviews to be expressed both through verbal and non-verbal communication strategies. Through this ongoing process of relationship building and reflection on the project's pathway, the evaluation team was able to share their learning with the PI team in a deep and integrated way, while the PI team was able to share their views on the project goals and objectives and why implementation needed to happen in specific ways. Key lessons learned included the need for immersive, participatory experiences to engage Fellows in Native paradigms; allowing time and space for the emotional aspects of working across worldviews; and providing Fellows with examples to inspire their own work, particularly examples that could be integrated into existing programs and exhibits.

Working together during summative evaluation

Summative evaluation determines a project's overall effectiveness. Here again, you and your evaluator must work together closely to clarify evaluation methods and measures, and to ensure that the resulting data will be useful to the team and the informal STEM education field.

Key roles and responsibilities during the

summative evaluation phase include:

Principal Investigator Responsibilities

- Ensure that the definition of success is clearly articulated
- Clarify targeted outcomes and match them with intended audiences
- Communicate with funder if targeted project outcomes have significantly shifted as a result of logistics, project management, staffing turnover, or front-end and formative evaluation findings
- Ensure that during summative data collection, project activities are as consistent as possible
- This is not the best time to try something new and explore the impact that it might have on participant experience—ideally such modifications should be explored during formative evaluation.

Evaluator Responsibilities

- Check that the summative evaluation will answer the questions put forth in the proposal and promised to the funder; carefully revise questions as necessary
- Determine whether the evaluation methods will answer the evaluation questions
- Develop items and instruments that align with appropriate methods; before data collection, make sure that results provided by the instruments will answer the evaluation questions
- Allow time for PI to review instruments
- Establish reasonable data collection timeline
- Conduct the study or studies

TRUE STORIES

Collaboration during summative evaluation.

Life Changes was a collaborative education and research effort funded by the Informal Science Education program at the National Science Foundation (DRL 0540152). This project explored whether learning experiences in a museum exhibition could productively address the lack of basic understanding of the biology of evolution and the challenges of teaching this complex topic. Exhibit designers from the New York Hall of Science, Miami Science Museum, and North Museum of Natural History & Science worked closely with researchers from the University of Michigan to produce exhibit components that introduced five basic evolutionary concepts: Variation, inheritance, selection, time, and adaptation (VISTA). Based on these concepts, the Life Changes team developed a 1,500 square foot traveling exhibition called Charlie and Kiwi's Evolutionary Adventure.

To successfully complete both the planned learning research and the summative evaluation, careful coordination was required around the development of instruments, timing of data collection, and decisions about analysis and data interpretation. The team created a complementary research evaluation design that supported an investigation of the impact of the exhibition experience on young children's understanding of basic evolutionary concepts. Summative evaluation determined that children's basic evolutionary thinking and reasoning were influenced by exposure to VISTA concepts in a museum context. Following their experiences in the exhibition, children were more aware that species can change over time, that dinosaurs and birds are related, and that these relationships have evolutionary explanations.

Working together during evaluation reporting/ dissemination

As the project wraps up, you and your evaluator should ensure that project stakeholders have access to the evaluation findings and that the results are shared with the field as appropriate.

Key roles and responsibilities during the reporting phase include:

Principal Investigator Responsibilities

- Monitor funder reporting guidelines and coordinate the collection of necessary information from project team members to be included in funder reports
- Create annual reports; Fill in monitoring system data forms
- Create and manage online project identity
- Verify that all abstracts and online descriptions (e.g., NSF Fastlane abstract) are an accurate reflection of the project
- Create project page on InformalScience.org with summary
- Create project page on ExhibitFiles.org if appropriate
- Ensure project representation at PI summit meetings

Evaluator responsibilities

- Provide necessary data to support completion of funder reporting requirements
- Review summaries of evaluation results included in annual reports, monitoring, online project pages for accuracy
- With final permission from the PI, post summative evaluation reports to InformalScience.org
- Work with PI and other project staff to target webinars, conferences, and publications that would be appropriate mechanisms for sharing evaluation findings.

TRUE STORIES

Collaboration during reporting and dissemination.

Asteroids! is a project of the National Center for Interactive Learning at the Space Science Institute, funded by the Informal Science Education program at the National Science Foundation (DRL 0813528). This multi-faceted informal STEM education initiative encourages public engagement and understanding of the dynamic structure of the solar system through investigations of asteroids, comets, and meteors. The centerpiece of this project was the development of the traveling exhibition Great Balls of Fire. The evaluators worked with the project team to provide front end, formative, and summative evaluation as well as original research associated with the exhibition design and development process. Each year the evaluators provided evaluation reports and data summaries ready for submission to the National Science Foundation and coordinated updates to the project logic model that reflected ongoing refinement of project activities and outcomes.

This project generated data and results that have been used by a range of project staff as well as the evaluation team to support presentations at annual meetings of the Association of Science-Technology Centers and the Astronomical Society of the Pacific. In addition, lessons learned from this project about the impact and implications of incorporating youth in the exhibition design development process motivated a week-long online discussion forum hosted by ASTC Connect. More than 100 participants signed up for the forum, which was coordinated by Asteroids project advisors, staff, and evaluators and featured contributors from other projects focused on supporting positive youth development and increasing youth engagement with STEM learning.

Conclusion

We have walked you through one approach to making sense of the life of a project, but it's important to realize that one size does not fit all when it comes to navigating the evaluation process. In reality projects are more complex than the linear process described in this chapter. The most effective way to ensure a productive working relationship is to establish expectations early, deliver on them to the best of your ability, and design regular communication mechanisms that support the ability to respond and adapt to changing project needs and result in effective collaboration.

Chapter 5
Planning for Success:

Supporting the Development of an Evaluation Plan

TINA PHILLIPS AND RICK BONNEY

Introduction

"Would you tell me, please, which way I ought to go from here?" "That depends a good deal on where you want to get to," said the Cat. (1)

The Cat, who utters this famous line in *Alice's Adventures in Wonderland,* knew the key to planning a successful project: You have to know exactly what you want your project to accomplish before you can decide what you'll do to accomplish it. Then, you and your evaluator can develop a plan to determine whether your project has been successful. More specifically you need to articulate clear project goals and measurable targeted outcomes. Goals tend to be lofty and visionary; outcomes are specific and describe the changes that you expect people to undergo as they experience your project.

With clearly articulated goals and outcomes in hand, your evaluator can develop a plan that serves as the roadmap for project evaluation and provides a window into the evaluation process. The plan should provide information about the purpose and context of the evaluation, who will be involved, and how evaluation data will be collected and reported. The plan also should include evaluation questions that align to the goals and outcomes and frame the entire evaluation. Finally, the plan should include a detailed timeline, budget, reporting strategies, and other logistical considerations such as the means for obtaining Institutional Review Board (IRB) approval for working with human subjects.

While the evaluation plan should be comprehensive it also needs to be flexible, so that it can reflect changes in project needs or circumstances as project development gets under way. For example, an evaluation plan may change owing to logistical hurdles such as access to participants or budget limitations; new project directions; or to explore unexpected outcomes as they emerge. Consider the evaluation plan to be a working document that you and your evaluator share and which evolves as stakeholders offer their perspectives and insights on the developing study (Diamond 2009).

The remainder of this section will guide you along the evaluation highway.

Key elements of an evaluation plan

Background information

- Project overview, intended audience, and stakeholders
- General information for your evaluator

Project goals and outcomes

- Logic model, theory of change, or other description of outcomes

Evaluation Questions

- Identification of what is to be evaluated
- Evaluation questions (refined and prioritized)

Indicators of success

- Measurable indicators of success
- Links between goals, outcomes, and indicators

Methodology

- Design strategy
- Data collection strategy
- Data analysis strategy
- Interpretation strategy
- Reporting strategy

Logistics

- Timeline
- Budgets
- IRB approval

Project Background

Project Overview, Intended Audience, and Stakeholders

Right from the start, your evaluator will seek to obtain as much information as possible about your project. He/she will want to know more than what is presented on a website or informational brochure. You'll need to provide information about your project's overall goals, intended audience, and project staff and partners. You'll want to describe the development and implementation plan for your project along with its targeted outcomes and deliverables. And, you'll need to describe all "stakeholders"—the people and/or institutions that will be interested in the evaluation process and results. These include funders, collaborators, program participants, administrators, and policy makers.

Information to share with your evaluator

Providing your evaluator with previous evaluations or reports about related projects will be invaluable in helping him/her understand your audience. If no prior evaluation reports are available, provide your evaluator with whatever demographic data you have about your target audience. You also may wish to include the organizational, cultural, and historical context for your project. For instance, it's helpful to share information about how the project and team like to work, your organizational structures and expectations, and any contextual information that may influence the evaluation design.

Finally, if your project operates under an existing program theory, be sure to share that with your evaluator. Articulating program theory can be done both formally and informally. For example, if your project operates in afterschool settings, you can find plenty of literature that describes the research in this area, which may provide a guiding theoretical framework for how these types of projects are intended to succeed. More often, project staff simply provide information to the evaluator about what the program is actually supposed to do and how it is supposed to do it. Either way is fine for describing how the program works, and the more information that you can provide to your evaluator at the start, the more efficient he/she can be with her time, allowing her to focus on developing a comprehensive evaluation plan.

Once the necessary information has been obtained, your evaluator should feel comfortable and well versed with your project and its intended audience. His/her understanding of the project should be evidenced in what is written as the background or overview of the evaluation plan. If you sense misunderstanding about your project, sort it out at the beginning!

Project goals and outcomes

Most likely you developed goals and outcomes for your project while preparing a project proposal or development plan. In an ideal world, your intended goals are achievable and your targeted outcomes are specific, measurable, and relevant to your project participants. As you begin working with your evaluator to develop an evaluation plan, however, you may discover that your goals were a bit too ambitious or that your targeted outcomes were vague. In short, your goals and outcomes may fall under the technical term of "fuzzy" (Patton 2008). If so, your evaluator's knowledge and experience can help you refine your goals and clearly

articulate your outcomes. For example, your evaluator will examine whether each of your outcomes identifies the intended audience(s) and how the audience is expected to change (e.g., increasing knowledge, developing a more positive attitude).

During this process you can expect your evaluator to ask you exactly what you are attempting to achieve with your project. Probing questions are not meant to make you feel uncomfortable; they are intended to help your evaluator clarify your goals and objectives, because learning whether they are being met is what the evaluation process is all about. And unlike other aspects of your project, which can change and adapt, changing your goals and objectives halfway through a project can mean starting all over with a new evaluation plan. Your evaluator also will check that all stakeholders agree on project goals and objectives, and if different stakeholders have different goals, the evaluator will set priorities or look for agreement that multiple goals will be evaluated.

Seeing how goals and objectives fit into project development can be challenging. Experienced evaluators are skilled at visualizing complexity, and their expertise will help you see the big picture and graphically present what you hope to provide and how your program will achieve its intended outcomes. Evaluators have many tools for visualizing complexity; below we discuss two that are widely used.

Logic models

A logic model is a visual depiction, often presented in matrix or mind-map form, of how a project works. You can think of a logic model as a graphical representation of your program theory. Logic models link outcomes (both short- and long-term) with project activities/processes and the theoretical assumptions and principles of the project (W.K. Kellogg Foundation 2004). Logic models also help evaluators focus their study on the most critical project elements (National Science Foundation 2010).

Logic models should be developed collaboratively between the project team and the evaluator. This process will contribute to a unified project vision including shared understanding of project goals, agreement on targeted project outcomes, and expectations about when those outcomes will occur.

While the PI leads decision-making for the logic model's content, evaluators often bring skills in facilitating and supporting the development process. They can help to distinguish and refine elements of the model and make sure that the full extent of potential outcomes are considered. The evaluator can also identify any outcomes depicted by the logic model that cannot be easily or reliably measured (this may lead to a discussion of whether those outcomes should still be included in the model).

While logic models should be developed early in project development, they are not static tools. As projects evolve to reflect changes in underlying assumptions and theory, logic models must be updated to reflect this thinking (National Science Foundation 2010). Some projects create a poster of their logic model and then, as the work progresses, use sticky notes to update and "check off" tasks within the model. Or they create an online tool that all team members can access, discuss, and modify. Logic models come in many types and formats and no single strategy is "best" for creating them. Often, however, a logic model is portrayed in graphic form with the following key elements: Inputs, activities, outputs, outcomes, and impacts.

Inputs

Inputs are resources that are made available to the project. They include funding sources, staff time, volunteer/user interest, and project or technological infrastructure.

Activities

Activities refer to things that the project will develop, conduct, or make available for use by the

intended audience. They can be broken down into activities conducted by project staff and those done by the intended audience. Staff activities could include providing training workshops, creating educational materials, recruiting partner organizations, or developing exhibits. Participant activities might include attending trainings or events, visiting a web site, collecting and submitting data, and communicating with others.

Outputs

Outputs are the direct products or services of the activities and typically are easy to quantify, for example, the number of training workshops that staff deliver, the number of people that participate in a project, or the number of web pages that a project produces.

Outcomes

Outcomes are the changes to individuals, groups, or communities as a result of project participation or experience. Outcomes are often described as short-term, occurring within a few years of the activity; medium-term, happening within 4–7 years after the activity; or long-term, happening many years after an activity has commenced (W.K. Kellogg Foundation 1998).

Impacts

Impacts are essentially long-term outcomes. They tend to be broad in scope and provide expanding knowledge or capacity for a particular segment of society. While desired impacts are often presented in logic models they are rarely measured because of their inherent complexity and because their timeframe usually lasts past the time of project funding.

Sample logic model

This sample logic model was adapted from a citizen science project (Figure 9,see page 48).

Theory of Change

Some evaluators may ask you to articulate your "Theory of Change," i.e., how you think each of your project activities will lead to your desired outcomes. A theory of change does not have to be based on documented theories but can be based on your prior experiences, assumptions, expert knowledge, or even wishful thinking. Once you make your theory of change explicit you need to communicate it to other members of your team and, in turn, have them share how they think the project activities will lead to desired outcomes.

Once your team's assumptions are made explicit, you can begin to test them by creating statements that link your activities with short, medium, and long-term outcomes. A theory of change will describe the strategy or set of actions to be implemented by the project as well as the desired outcome from those activities. The easiest way to do this is by using "if . . . then" statements. For example, let's say that you are implementing an afterschool program aimed at increasing interest in science careers. For this outcome, begin by listing your assumptions: We assume that exposing kids to science will increase their interest in science careers. Then describe the activities as they relate to the outcomes with "if . . . then" statements. You may find that you need to provide additional activities or supports to reach the outcome.

EX 1: If we provide fun, compelling science related activities, then we will increase interest in science careers.

Are there holes in example 1? Are there assumptions that need to be addressed? Could it be improved? Let's try another one...

EX 2: If we provide science-based activities, and describe how they relate to science careers, then students in the afterschool program will have knowledge of some different science careers. If students know about different science careers, then they may seek out additional information about a particular career. If they seek out more information on a career, then they may show increased interest in pursuing a science career.

	INPUTS	ACTIVITIES (PARTICIPANT)	OUTPUTS	SHORT-TERM OUTCOMES	MEDIUM-TERM OUTCOMES	LONG-TERM IMPACTS
Logic Model Example	■ Project leaders, scientists, support staff, educators ■ Volunteers' interest, time, skills, prior knowledge, and motivation ■ Partnering institutions ■ Funding sources	■ Learn about project protocol ■ Collect and submit data ■ Provide feedback to project staff ■ Communicate with others via groups, listservs, etc.	■ Amount of volunteer-collected data ■ Publicly accessible database ■ Individuals engaged with program ■ Interactive and educational website	■ Increased engagement with science and nature ■ Increased knowledge of science content ■ Improved data collection skills ■ Improved species identification skills	■ Increased public access to scientific institutions ■ Sustained change in the way participants collect high-quality data ■ Participants serve as project ambassadors to promote project	■ Increased public support of science ■ Increased public appreciation and stewardship of nature
Indicators	■ Number of staff ■ Available resources ■ Baseline data of participants' interest, knowledge, skills, etc.	■ Participants express understanding of project protocol ■ Participants demonstrate confidence in collecting & submitting data ■ Web analytics of project website ■ Quality & quantity of communication between staff and participants	■ Number and quality of data submitted ■ Number of people accessing database ■ Number of people engaged; frequency, duration, and intensity of engagement ■ Web analytics of project website; quality of web-based educational materials	■ Number of hours spent with science and nature activities ■ Measures of change in knowledge content from pre- to post-tests ■ Self-reported increase in data collection skills ■ Demonstrated increase in species identification skills from pre- to post-quiz	■ Change in quantity of communications between scientists and participants ■ Detection of changes in long-term data submission records ■ Data is of higher quality and more valuable over time ■ Amount of publicity and project exposure my participants	■ Increased private and public funding for science institutions ■ Improved environmental conditions

Figure 9

The set of statements in example 2 makes it much more clear how the activities are linked to the desired outcomes. As project developers we are often too embedded in programs to see and identify assumptions about audience needs and interests or to envision the explicit mechanisms that must be in place through project activities to influence change. Working with your evaluator to develop logical "if . . . then" statements can help uncover and address these assumptions so that activities and outcomes are aligned.

Sample Theory of Change

A theory of change can also be depicted graphically as a "results chain," as demonstrated in Figure 10, see page 49.

Whether you and your evaluator develop a logic model, theory of change, or some other representation of your project, remember that you are the driver and primary decision maker for setting the project direction. Your evaluator complements and supports your project. In other words, it is the PI's job to decide what a project should

Graphical representation of a "theory of change."

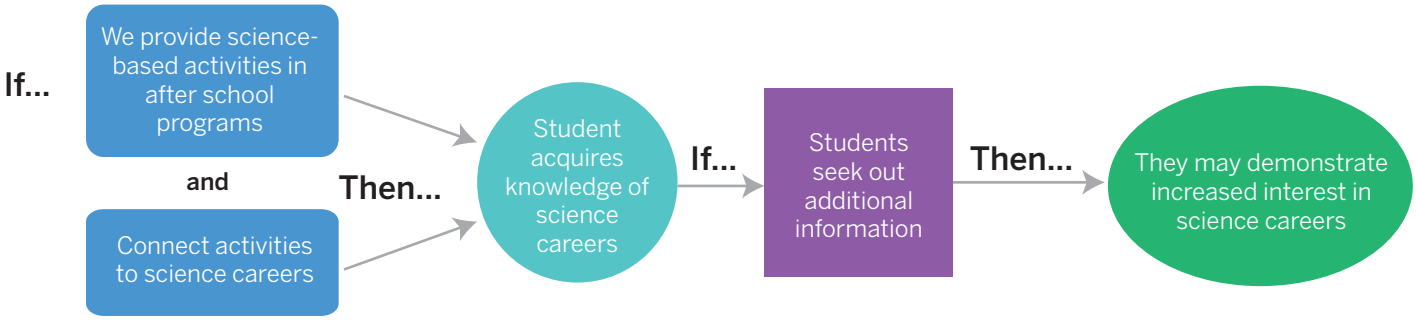


Figure 10

do and the evaluator's job to determine whether the project did it. That said, evaluators often support PIs and project teams by facilitating the process of focusing project goals, intended audiences and outcomes, and by providing expertise in clarifying ideas to build cohesive project designs and conceptual frameworks to drive the evaluation forward.

Common pitfalls when developing goals, outcomes, and indicators

It is not enough to develop a program and then assume that participants will achieve the outcomes that you intend for them to achieve. Below are pitfalls that we often see in program development:

- "Wishy-washy outcomes"- outcomes that are not specific, not measurable, and not relevant to the project.
- Targeted outcomes not aligned to project activities—for instance you say that you want your project participants to increase their data interpretation skills, but your project does not actively support data interpretation as an activity.
- Expecting too much— You want your project to have far-reaching and lasting impacts, but the truth is that your resources are limited. You need to be realistic about what your project can actually influence.

Evaluation questions

Your evaluation questions form the backbone of your design strategy and everything that follows. It is helpful to begin by clarifying what you intend to evaluate and understanding what will not be evaluated. Next, you will generate questions that can be answered during front-end, formative, and/or summative evaluation. We have included sample questions in each of those broad categories as thought-starters. And finally, as you refine your set of questions, you will want to shape them and prioritize them according to a variety of criteria described below.

Identifying what is to be evaluated

With goals, outcomes, and a logic model for your project in place, the next step is to explicitly articulate the main reason or reasons for your project evaluation, the specific aspects of the project that will be evaluated, and the specific audience for the activities or products that will be evaluated. The phases of evaluation discussed earlier in this guide can be used to frame the evaluation plan:

Front end:

- Determine audience needs and interests
- Acquire contextual information about the political, social, and cultural environment of a particular program

Formative:

- Monitor a project on an ongoing basis through regular data collection
- Describe how a project functions
- Provide recommendations to improve project functionality
- Clarify program purpose or theory

Summative:

- Gauge whether targeted outcomes have been achieved
- Summarize learning from the evaluation and any unintended effects that were documented
- Identify project strengths and weaknesses
- Determine overall value or worth of a project
- Determine cause and effect relationships between an intervention and outcomes

Additional goals of evaluation can include:

- Obtain additional funding or support
- Increase organizational evaluation capacity building
- Compare outcomes across projects
- Conduct a cost-benefit analysis between project costs and outcomes

Just as important as identifying what will be evaluated is deciding what will not be evaluated. Defining boundaries for the evaluation as the project begins—whether such boundaries are specific audiences, time frames, locations, or individual project elements—will minimize any surprises later in the process. Too often PIs arrive at the end of project development and wonder why something was not evaluated simply because the boundaries of the evaluation were not explicitly discussed. Avoiding this problem is easy if you take responsibility for communicating boundaries to your evaluator as the evaluation plan is developed. You'll also want to check with your evaluator to see whether he or she foresees any constraints that might affect the overall evaluation.

Developing evaluation questions

The next step in developing the evaluation plan is to frame appropriate evaluation questions within the context of desired outcomes and the purpose of your evaluation. Evaluation questions should be broad enough to frame the overall evaluation yet specific enough to focus it. Articulating well-formed questions (those that frame the overall study, not questions that might be asked of participants) will help your evaluator determine the overall study design and approach and selection of methods (Diamond 2009). You and your evaluator can work together toward developing questions that will address what you need to know to determine if you are reaching your desired outcomes. Answers to the evaluation questions must be relevant, meaningful, evidence-based, and useful to the project stakeholders.

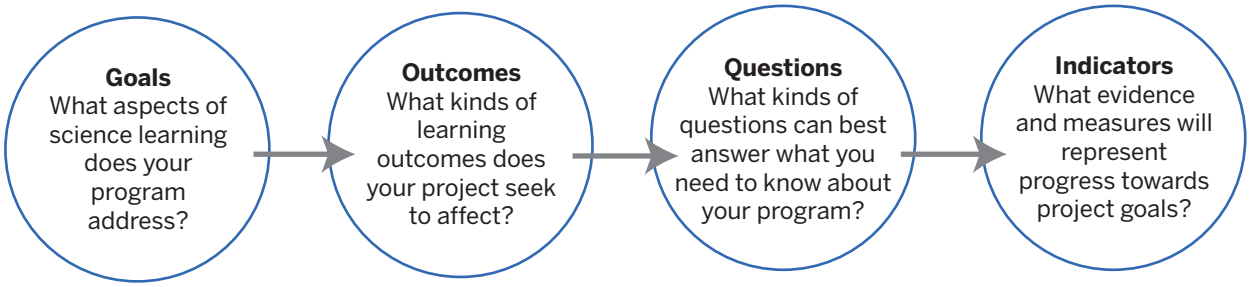
Sample evaluation questions

For example, a **front-end evaluation** interested in better understanding a project's audience might ask the following types of questions:

- What does our audience already know about this particular topic?
- What misconceptions exist among our audience regarding this topic?
- How interested is the intended audience in this new emerging topic?

Formative evaluation questions, which focus on understanding the extent to which a project is functioning as expected, may ask:

- What, if any, were the barriers to participation?
- Were project participants satisfied with their experience? Why or why not?
- What lessons were learned about developing and implementing the project?
- Were participants engaging in activities as planned? Why or why not?



Linkages between goals, outcomes, questions, and indicators when designing an evaluation plan.

Figure 11

Summative evaluations, where the emphasis is on determining if projects have met their goals, may ask the following questions:

- Was there evidence of an increase or change in knowledge as a result of interacting with this exhibit? For which participants, and what level?
- Did participants improve their skills in data interpretation after participating in the project?
- Was there evidence that participants changed aspects of their consumer behavior as a result of viewing this television program?
- What was the value, if any, of participation in this project for the intended audience?

Qualities of effective evaluation questions

You will likely come up with a large number of questions for which you would like answers, but remember that not all questions can be answered given the allotted time and resources, and not all questions will have the same importance to all stakeholders. Also, multiple data sources can be used to answer individual evaluation questions; similarly, single data sources can contribute to answering multiple evaluation questions.

Your evaluator will work with you to ensure that your evaluation questions are 1) answerable; 2) appropriate for the various stages of evaluation; 3) aligned to the desired outcomes; and 4) address stakeholders' information needs. In addition to these criteria, your evaluator also will help you prioritize the questions that are most critical to address by considering the following aspects:

- The resources needed to answer the question
- The time required
- The value of the information in informing the evaluation purpose

As each question is examined through the lens of these criteria, some will present themselves as high priority while others will be eliminated altogether. At the end of this process you should feel comfortable knowing that the questions you focus on will demonstrate measurability, relevance, and feasibility, while setting the stage for the rest of the evaluation roadmap. (Figure 11)

Indicators of Success

Now you have project goals, outcomes, a logic model, clearly expressed reasons for conducting your evaluation, and clearly articulated evaluation questions. The next task that you and your evaluator will tackle is developing indicators, which are criteria for measuring the extent to which your targeted outcomes are being achieved. Effective indicators align directly to outcomes and are clear, measurable, unbiased, and sensitive to change. For instance, if an outcome relates to knowledge gains, the indicator should measure knowledge gains as opposed to, say, participant interest. An indicator answers the question: How will you know it when you see it? And while indicators are measurable, they do not always need to be quantifiable. Indicators can also be qualitative and descriptive, i.e., "Participants will describe that they . . . "

Identifying realistic, feasible, and valid indicators is probably the most difficult step in designing an evaluation. The constraints of time, funding, and reach can restrict the types of data that you can collect. At the same time, the easiest things to document may not always be the most salient or compelling issues. Sometimes data are not feasible or available for certain indicators; in this case lack of data sources should be acknowledged in the evaluation plan as a limitation of the study.

Links between goals, outcomes, questions, and indicators

Template for articulating goals, outcomes, and indicators

In the template for articulating goals, outcomes, and indicators provided in Figure 12, you will note that for each goal we provide space for developing several outcomes and indicators. There is no "correct" number of outcomes or indicators, and each project will vary in the number that it attempts to achieve and measure. Working through this worksheet will be an extremely valuable exercise in developing a project and its associated evaluation plan. And if you include an outcomes development sheet as part of a grant proposal, you'll help readers better understand the chain of effects that you're hoping will result from your project.

Evaluation Methodology

Matching the Study Design to your Questions

As we continue our journey down the evaluation highway we arrive at a critical juncture: What strategy will we use to design the overall evaluation? The answer should reflect the types of questions you need answered, the reason for conducting the evaluation, the methods that best address the evaluation questions, the amount of resources you can commit to the evaluation, and the information that project stakeholders hope to learn.

Many different evaluation approaches and study designs exist, and it is beyond the scope of this guide to describe them all. Different study designs are better suited for different types of evaluation questions. If your question is concerned with comparing outcomes for participants directly before and after project

Figure 12

<table>
<tr><th>TARGETED OUTCOME</th><th>EVALUATION QUESTION</th><th colspan="2">INDICATOR (INCLUDE AUDIENCE)</th></tr>
<tr><td rowspan="5">High school visitors will increased their interest in the moon.</td><td rowspan="5">To what extent does the exhibit change high school students' interest in the moon?</td><td colspan="2">During their visit, high school students will engage their parents in conversation about specific phenomena featured in the exhibit.</td></tr>
<tr><td>Data Collection Method:
Observational tracking studies</td><td>Timeline, Personnel:
Three days, lead evaluator</td></tr>
<tr><td colspan="2">High school students visiting the museum will indicate that the exhibit increased their interest in learning more about the moon and/or a related topic.</td></tr>
<tr><td>Data Collection Method:
Interviews with students after viewing the exhibit.</td><td>Timeline, Personnel:
Three days, lead evaluator</td></tr>
</table>

participation, then pre-post designs will likely fit the bill. Questions that seek to answer causal processes where you can include control groups and random assignment are best suited for experimental designs. Many evaluators will combine these approaches to achieve mixed-methods designs, and will incorporate both quantitative and qualitative techniques as a method of enhancing the strength of various data collections methods and increasing the validity of results through a triangulation of findings (Creswell 2003). For example, if one of your questions is best answered by broad representation of a population and data are easy to acquire through questionnaires, then quantitative survey methods work very well. If one of your questions requires gathering robust information on participant experiences and you can gain easy access to participants, then qualitative interview or focus group methods are appropriate.

Data Collection Strategy

A common pitfall in designing evaluation studies is the instinct to start by identifying preferred methods; for example, "What I want is a series of focus groups conducted with youth in the science afterschool program" (Diamond 2009). Discussion of data collection methods should come only after your goals, targeted outcomes, evaluation questions, indicators, and study design have been clarified and agreed upon. Then, for each indicator, you and your evaluator will need to determine:

1. Who is the intended audience and what specific information do you hope to get from its members? (This discussion should be led by the PI.)

2. What method of data collection is best suited for obtaining the information that you need from this audience? (This discussion should be led by the evaluator.)

3. When will the information be collected and by whom? (This discussion should be led by the evaluator with input from the PI.)

The possibilities for data-collection strategies are nearly endless. In choosing methods your evaluator will consider issues such as the potential trade-offs in collecting rich, in-depth qualitative information versus information that has a high level of statistical precision, the need to collect standardized data, the cultural attributes of the audience, and the availability of contact information for the sample. These issues will also help your evaluator determine the population to sample and the appropriate sample size.

Sample data collection strategy

Data analysis strategy

Data analysis involves the process of examining, cleaning, and transforming data so that conclusions can be reached about whether targeted outcomes were realized. Data analysis can take many different forms and relies on different methodologies depending on the project need, audience, how the information will be used, and your evaluator's expertise. If the evaluation is going to rely heavily on qualitative data, i.e., data derived from text or images, then data reduction will be required to transform lengthy documents into succinct, useful information (usually in the form of common themes or categories). If the evaluation is going to be primarily quantitative, i.e., collecting various numbers or scores, your evaluator will need to use statistical methods to transform the data into charts, graphs, and tables that assign meaning to all the numbers and provide comprehensible information. Your evaluator may be skilled in analyzing both qualitative and qualitative data, thereby leveraging the strengths from both of these methodological approaches.

Before getting to the data analysis phase it is critical that you understand and are comfortable with the approach that your evaluator will use for collecting data, as this will most certainly shape the way in which he or she analyzes data. Regardless of the approach used, you should feel comfortable asking about the overall quality of

the data set and the measurements used, and if the appropriate data were collected in order to answer the evaluation questions.

Data interpretation strategy

Evaluation is both an art and a science, and nowhere is that more evident than in the data interpretation phase. Just as no two people will interpret a painting in the same exact way, no two evaluators will interpret data (either quantitative or qualitative) in exactly the same way.

Your evaluator should have expertise in interpreting the kind of data that you plan to obtain through your evaluation and should be able to explain how the interpretation will describe outcomes that were and were not realized and why. Data interpretation also should help to clarify whether limitations of the study design, data collection process, or other circumstances contributed to the results. In some cases, unintended outcomes and how these could be incorporated into future project improvements may be revealed. Some evaluators may also plan to compare results from your project with those from similar programs. Evaluators may also plan to reflect on project outcomes, the broader context of the project, and future research possibilities. If these are issues that you would like to have included in the data interpretation be sure to spell them out!

Data reporting strategy

Once data have been analyzed and synthesized, your evaluator will need to write an evaluation report. The report may be the most tangible product of your evaluation process and will be shared with all stakeholders interested in your project impacts. This phase of project evaluation is so important, and holds so many possibilities, that we have included an entire chapter of this guide on the subject (see Chapter 6).

In developing the evaluation plan your evaluator should describe not only what will be in the report but also how and when the information will be shared. For example, some evaluators provide continuous feedback about data being collected through interim reports or via regular meetings. Other evaluators prefer to wait until data collection is complete before analyzing or sharing information with you. Make sure that you are comfortable with the reporting strategy described in the plan.

Logistics

Your evaluator can assist you in laying out a budget and timeline for your evaluation design and ensuring that it meets requirements for Institutional Review Board (IRB) approval. It is helpful to maintain an open dialogue with your evaluator about the costs and time frames associated with different aspects of your evaluation study in order to shape a design strategy that is aligned with your budget and schedule.

Timeline

The evaluation plan should include a timeline that provides anticipated start and end dates for completing key tasks and meeting established milestones. (Figure 13, see page 55.)

Timelines are often presented in calendar format:

Be sure that the timeline seems reasonable given what you know about your project and its audience. For example, if the evaluator is conducting formative usability testing of a web-based application that your staff will develop, does the timeline align with your team's development schedule? If the evaluator plans to collect data for summative evaluation through a survey of participants, does the timeline allow sufficient time to recruit willing respondents? While timelines often change, starting with one that seems realistic will help to avoid later frustrations.

Budgets

The evaluation plan also needs to provide a budget. Complete evaluations typically make up about 10 percent of an overall project budget,

Timeline for Evaluation Activities for a Mixed-Methods Study

	YEAR 1				YEAR 2				YEAR 3			
	Sum	Fall	Win	Spr	Sum	Fall	Win	Spr	Sum	Fall	Win	Spr
Hold Initial meetings with stakeholders	X											
Identify Key people	X											
Obtain IRB approval	X											
Draft Scope of Work and budget outline	X											
Review all project materials		X										
Refine Goals, Outcomes, Indicators		X										
Develop Logic Model, Theory of Change		X										
Develop & prioritize evaluation questions		X										
Determine indicators		X										
Develop study design and data collection strategy			X									
Deliver draft evaluation plan to stakeholders, refine if needed			X	X								
Develop draft instruments				X								
Share instruments with stakeholders, refine if needed				X	X							
Pilot draft instruments					X							
Refine instruments					X							
Administer pre-test					X	X						
Conduct phone interviews						X						
Gather & clean pre test data						X						
Analyze interview data							X	X				
Administer post-test								X				
Gather & clean post-test data									X			
Analyze pre-post data									X			
Review other pertinent data sources										X		
Discuss preliminary findings with stakeholders										X		
Draft and disseminate final report											X	X
Secure data according to data management plan												X

Figure 13

but this figure can vary greatly depending on the evaluation complexity. For example, evaluations that incorporate experimental designs with control groups are generally more costly than those that rely on pre-post project surveys with no control groups. Considering qualitative designs, interviewing, transcribing, and analyzing information from 50 people for 60–90 minutes each can also be very time intensive and thus expensive. Recruiting participants can be costly depending on how you intend to engage them, particularly if incentives are required to ensure their participation. In discussing the plan with your evaluator, he/she will give you a sense of what is feasible at different cost levels, and together you can develop a budget that is appropriate for the project.

Institutional Review Board approval

Most Institutional Review Boards (IRBs) require you to submit a detailed description of your project, your audience, the methods you will use, any surveys, observation guides, interview guides, or other instruments you intend to use, and how you intend to recruit people into your study. They will also want to see a copy of a consent form as well as a description of how you will minimize risk to your participants and ensure their confidentiality. Typically, independent evaluators do not have direct access to an IRB and must rely on a college or university IRB to acquire approval. Be sure to check with your organization to determine what is required so that together you and your evaluator can complete the necessary training and submit the required documents well ahead of implementing your evaluation.

Conclusion

When the evaluation plan is complete, it will be up to you to make sure that it will meet your project needs. You may need to go back and forth with your evaluator a few times—indeed, constructing an evaluation plan that is relevant, feasible, and effective requires regular and iterative communication between the project team and evaluator. Remember that you are driving this process and it is up to you to make sure that the evaluation serves your project's long-term interests and helps answer questions that will guide your future planning and management goals. And, as you're learning about evaluation, your evaluator is learning about your project and your organization. There's a lot for you both to learn through this process, so clear communication, patience, flexibility, and a good sense of humor are all necessary elements in developing a strong and collaborative evaluation plan.

Chapter 6
Reporting and Dissemination:
Building in Dissemination from the Start

SASKIA TRAILL AND RACHEL HELLENGA

Introduction

"What Makes a Great Evaluation Report" was the topic of two workshops organized by the Visitor Studies Association for the 2012 National Science Foundation PI Meeting. The workshops were professional development opportunities, but they doubled as reconnaissance missions in service of this chapter. The workshop organizers assigned spies—oops, we mean scribes—to capture the insights, real world stories, and concrete examples shared during these sessions. Consider this intelligence-gathering to be a form of benign industrial espionage, aimed at unearthing trade secrets for the benefit of our readers. Your co-authors encourage you to borrow heavily from these strategies and recommendations, and even to copy them outright. In the name of dissemination, let's kick off this chapter with a quote overheard during the discussion.

TRUE STORIES

The critical thing is not that we succeed, but that we generate useful findings.

That pithy quote struck us as an excellent introduction to this chapter. Projects aimed at advancing the field of informal STEM involve experimentation and innovation. And with innovation comes risk. Whether or not your project goes as planned, it is important to communicate results of the project effectively to people who have the potential to extend, replicate, build on, or learn from your work. This chapter will look at the many ways in which you can use evaluation findings to achieve broader impacts for your project.

Getting a useful evaluation report: Tips from your peers

So how do you get an evaluation report you can use? A better question to ask might be, "How do I get an evaluation report everybody can use?" The variety of stakeholders in each project can result in many different possible goals and purposes for your evaluation study, as outlined in more depth in Chapter 2. Your stakeholders might typically include your funders; your internal project team; staff and administrators at your institution; your project's participants or consumers; colleagues in your field; colleagues in tangential fields; and future collaborators.

Creating a dissemination plan early can prompt you to think ahead about the needs of your stakeholders so that when it's time to produce your report, you and your evaluator will have gathered the relevant data and other documentation (e.g., photos, videos, or other visual evidence) to support various reporting formats.

Eavesdropping on discussions about "What Makes a Great Evaluation Report?" turned up several recommendations and strongly held convictions:

Define the summative report in the contract scope of work

As you define the scope of work, specify your expectations for the summative report. This is the time to ask for an Executive Summary and any custom report formats that you might need. Looking at a table of contents from other reports might be helpful to you in preparing for this conversation. For example, you might realize

you want to see evaluation instruments included in the appendices. In addition, the Summative Evaluation Report Checklist can serve as the basis for a conversation with your evaluator to make sure there are no surprises about what you need in the report.

The Summative Evaluation Checklist

This checklist can help you to plan for the elements that will make your report useful to your peers. Given that an evaluation report starts out as a tool for communication between you and your evaluator, it's possible that information known to your internal team may be omitted from the report unless you explicitly ask for it to be included. When you wrote "educators" did you mean informal or formal educators? Where did the evaluation take place? Don't leave your readers in the dark. (Figure 14, see page 60.)

In addition to addressing the summative evaluation report in the scope of work of your evaluator's contract, you'll want to factor it into your schedule. Let your evaluator know that you will expect to see and comment on a draft of the summative report before it is finalized. Look ahead at your timeline for reporting to stakeholders (e.g., board meetings, funder progress reports) and build in time for review and revision. For example, if a funder's final report is due in September, it makes sense to have the final evaluation report in late July or early August, so you can report findings in the final report.

Give special attention to the Executive Summary

An Executive Summary is an important part of a summative report. Many funders may refer primarily to this shorter document and it can be a tool to move the institution forward. Ask your evaluator to include highlights that are powerful: perhaps you experienced success due to a design strategy or faced a challenge that is relevant across the field. You should be confident about making suggestions for the Executive Summary, such as including specific points or describing tie-ins to previous work. Of course, the summary must accurately portray the results! It's not a place to sugarcoat negative findings or issues.

Keep in mind that when you send a final report or Executive Summary to one individual, it may go to others without the supporting documents that you intend to accompany it. For this reason, each portion of a report that you send out must be able to stand on its own. The Executive Summary should be a fair distillation of the research that conveys both the nature of the project and what you discovered.

Shape your summative report with your evaluator

Several workshop participants described the reporting process as a dynamic exchange in which the evaluator and client work together as co-authors. They suggested thinking of reporting in two phases: internal, then external. You have an important role to play in the interpretation and presentation of the data, so don't file draft reports away for later! It's important to build in time for your internal team to review drafts and offer prompt feedback to the evaluator.

TRUE STORIES

Including the project team's perspective when reporting results.

One PI related that staff in her institution originally saw the summative report as "test results," like getting a grade. It didn't occur to them that they could discuss the ideas and findings that would be presented. Her message: have the confidence to engage in respectful dialogue with your evaluator about what will go in the report. For example, the Children's Discovery Museum of San Jose's "Secrets of Circles" exhibit team first met with their evaluator to review the summative report in a fairly raw format. The evaluator learned what the team found to be the most exciting, surprising, and meaningful results, and emphasized these in the formal report. The staff

The Summative Evaluation Checklist

1) Basics: Report Title; Project title; NSF ID Number (if applicable); Dates (Year of written report, funding start date, funding expiration date); Author; Evaluation organization's name; PI; PI Organization's name
2) Evaluation Type: Front-end, Formative, Remedial, Summative, Other
3) Audience for the project: Children (specify age range of sample); Adults; Adult-Child groups; Informal education professionals; Formal Education Professionals, etc.
4) Project type: Educational program, Exhibit, Afterschool program, Broadcast media, Camp, etc.
5) Project setting: Aquarium or Zoo, Anthropology Museum, Radio, community-based organization, etc.
6) Science content area: Anthropology, Biological Science, etc.
7) Evaluation design: Mixed method, quasi-experiment, post-test, ethnographic study, longitudinal, etc.
8) Evaluation questions: what questions about the project were you trying to answer?
9) Description of the overall evaluation design, implementation and analyses
10) Data collection method(s): Survey, Interview, Focus Group, observation, etc.
11) Evaluator type: Internal, external
12) Sample descriptions and sample sizes (n)
13) Include Evaluation Instruments
14) Indicate whether the evaluation instruments were based on other validated instruments or how piloting and validation was conducted
15) Indicate whether the evaluation design, methodology, instruments, or specific questions were linked to research or literature in the field
16) What analyses were performed on data collected?
17) Recommendations or suggestions for improvement (These are recommendations from the evaluator, not a question about improvements in a survey or interview)
18) If funded by NSF, indicate which NSF impact areas were addressed: Awareness, knowledge, understanding, engagement, interest, attitude, behavior, skills, other
19) Provide an executive summary

Figure 14

reviewed the report and Executive Summary and made suggestions and comments before distributing it more widely.

Don't whitewash results

Taking a joint-authorship approach to the final report can offer many benefits, but it is important to note that your goal is to have an accurate and helpful report that is supported by data. Some evaluators have had to defend the wording of a report in the face of a client or funder who wanted to rewrite the summary with a more positive spin. In fact, the *American Journal of Evaluation* is conducting a study of this issue, including tips for avoiding misrepresentation of findings. This is just a reminder that innovation involves risk and possible failure, and all project outcomes are valid and worth reporting accurately. Trying to make the report more useful is different from trying to make the findings sound better than they were.

Paint a vivid picture

Several workshop members noted the value of capturing and reporting findings such as unintended outcomes that do not fit neatly into the original program logic model. Regardless of the methodology that was used for your evaluation, many PIs noted the importance of including qualitative descriptions of your project. Taking the time to describe the context and share impressions as part of the report can help paint a vivid picture and lead to additional insights for the project team and peers in the field.

Capturing unexpected outcomes.
The evaluation of a calculus exhibition at the Science Museum of Minnesota revealed that the exhibition was a powerful evoker of memory for visitors who had studied math.

The team had not articulated an intended outcome related to prompting positive memories of math, but inclusion of qualitative evaluation methods and a flexible approach to the reporting allowed them to uncover and document this unexpected outcome.

Creating a "highlights" document. In another example, the Children's Discovery Museum of San Jose invited members of the Vietnamese community to visit the museum as a group and to share their feedback on an exhibition. The format of data collection was not consistent with the larger study, so the results were called out separately. This component of the evaluation turned out to be the most valuable to the team and to the field. A report on this work was distributed widely in the form of a "highlights document," a polished presentation featuring graphic inserts calling out implications and direct quotes.

Key stakeholders and how to reach them

Now that we've shared the most urgent recommendations and strongly held convictions from your peers, we want to come back to the question of who will receive your evaluation results and how you will get this information out to them. Let's explore some of the key stakeholder categories and some strategies for reaching them.

Funders

Public and private funders represent a primary audience for evaluation findings, so you should know exactly what they require or expect before you or your evaluator generate reports. Some funders are hands-on when it comes to evaluation; for example, helping craft the right research questions or offering ideas for selecting an appropriate evaluator. Hands-on funders will stay engaged in a conversation about your evaluation, so you are less likely to be surprised by a sudden request, but you will have to devote resources to managing that dialogue. And, you may be sharing a lot with them—descriptions of methodology, data collection progress, metrics, initial findings, and a final report.

Some funders are more hands-off, but may still expect to see the evaluation report when it is completed. For both types of funders, it is important to make sure you are clear about their

expectations and what you will be submitting to them. Your external evaluator may already have worked with this funder and know their expectations. If you're not sure, ask.

Co-workers

One of the most overlooked audiences can be your co-workers. Disseminating findings internally is a way to build your organization's capacity and support its efforts to be a research-driven institution. Staff who aren't directly involved in this project may not take the time to read the full methodology and findings section, so it is not enough to simply forward the full report. You might consider sharing the Executive Summary, your own summary of findings, or some other custom presentation of information for your co-workers. For example, co-author Saskia hosts brown-bag discussions at TASC (The After School Corporation) to present findings and discuss impact. The discussions allow a free flow of ideas and questions suitable to brainstorming design changes. Her team has also presented at staff meetings or shared the highlights of an evaluation via e-mail.

Consumers

The consumers of your funded project—families of children in after-school programs, radio listeners, participants in public research projects—often do not hear what was evaluated and what was found. It is possible, however, that they would be interested in the research and the findings. You can reach them with many of the vehicles described below, such as social media, or through alternative methods such as policy briefs, brochures, and annual reports.

Sharing evaluation findings and results can expand the knowledge base of the field, build relationships, and even increase your own credibility. Your evaluation report might spark new ways of thinking and ignite change in practice or policy throughout a particular area of informal science. Ask yourself these questions:

- What does it mean that your evaluation came out the way it did?
- Have you found evidence for something interesting that could change other people's practice?
- Does it shed light on a trenchant problem for the field?

If you find something to say on these issues, you have a basis for starting a meaningful conversation with your colleagues. It is OK to mix the findings with your own message as long as you can be clear about the research versus your opinion. Consider collaborating with your evaluator to present together at key conferences, proposing solo conference presentations, writing journal articles, and designing Association of Science-Technology Center (ASTC) RAP sessions (Roundtables for Advancing the Professions), professional development workshops, or university courses.

We're not asking you to quit your day job and go on the road as a motivational speaker, but it can be well worth your time to go one or two steps past mailing the report to your funder. For example, informal communication via blogs and Twitter can help you convey what you're learning in order to support similar projects. If your project is funded by NSF, you'll certainly post your evaluation report on InformalScience.org (because you have to); and even if you don't have NSF funding, InformalScience.org welcomes evaluation reports from all relevant projects. When uploading full reports, remember to review the Summative Evaluation Checklist at the end of this chapter to ensure your reports will be understood by people unfamiliar with your project. Exhibition projects can also be profiled in a case study uploaded to exhibitfiles.org; as described in Chapter 3, Exhibit Files is a social media site aimed at exhibit developers and designers and maintained by ASTC. We would also like to challenge you to ask yourself who else is part of this

larger effort, whether or not they use the same terms to define the boundaries of your shared field. Think about who those unlikely field members are for your work.

TRUE STORIES

Beyond the usual suspects.

When a team at TASC started speaking with Science, Technology, Engineering, and Math (STEM) advocates in New York State, those advocates believed they were including after-school providers because science museums were a part of their outreach efforts. TASC encouraged them to go beyond museums to reach out to youth-serving organizations that offer diverse after-school programs with high-quality STEM activities. These after-school providers are also a part of the larger field, but would have been missed without more exploration about what types of institutions fall within the boundaries of the field.

Policymakers

Informing policymakers via personal contact and policy briefs is an important part of building sustainability. These officials can remove barriers and redirect public funds to support informal STEM education.

Elected and appointed officials can use their offices to highlight your successes and encourage the public to take an interest in your work. Evaluation findings are also a great reason to get back in touch with a policy maker's staffer who keeps a file on science issues or on your institution or organization. Data of any kind, along with a compelling story about the work you are doing, is powerful stuff for policymakers and influencers. Providing them with this information helps them with speech writing, plus they see you as an expert in your area. This outreach helps build important relationships.

America's scores on international tests of science have dropped just as we are seeing rapid increase in demand for science-literate members of the workforce; as a result, there is unprecedented interest in science education among policymakers at all levels. Sharing project findings with policymakers helps to make the case for increased funding (or against decreased funding) for federal agencies (like NSF and NASA) and programs such as Advancing Informal STEM Learning (AISL).

TRUE STORIES

The After-School Corporation.

TASC has implemented a "grassroots" and "grasstops" strategy for embedding science activities into comprehensive after-school programs in New York City. At the "grassroots" level, TASC trained after-school workers in how to use an engaging science curriculum and built up their confidence as science facilitators. At the "grasstops" level, TASC organized institutes which brought together New York City leaders of science, after-school, and education. During the institutes, leaders learned about specific strategies for integrating science into after-school programming along with evaluation findings that showed an increase in confidence about science among after-school educators and their students. Leaders from the Department of Youth and Community Development participated in the Institute and later added a requirement that grantees providing after-school programming include two hours per weeks of science or literacy activities. While many factors beyond a single institute certainly played into this decision, the staff at TASC saw it as a victory due in part to dissemination of evaluation results.

Potential collaboration partners

So often dissemination feels like due diligence in getting the evaluation to the people you know who do similar work. But what if you got your evaluation into the hands of your next collaboration partner? What if Bjork read your evaluation findings and decided to make an interactive science album? It's worth taking some time to think about how you might use the evaluation as a way to start

or deepen a conversation. Here are a few ideas:

- Use Twitter to pose a question to the "twitterverse" and see what comes back.
- Present at a conference that is nontraditional for you.
- Take your most surprising finding and imagine who might find it unsurprising.
- Ask your evaluator who would be interested in these findings.
- Ask your Program Officer who might want to know about these findings.
- Look at those who comment on your blog post.
- Present the findings at a local funders' group meeting.
- Reach out to other continents of the informal STEM education world—consider museums, youth-serving organizations, public media, or universities, to name a few.
- Reach out to the offices of elected officials.
- Talk to local or state education or youth development agency leaders.
- Reach out to formal educators and formal educational institutions.

More strategies for presenting and communicating results

Your summative report has many purposes and you may need to present the findings in multiple formats to accomplish your objectives. You don't necessarily need to spend your limited evaluation budget paying your evaluator to produce these additional presentations. You can ask for the content and use it to create the documents you need. It's best to make these requests up front when you negotiate your evaluator's scope of work. Some useful format variations are detailed below.

Alternative report formats: Social Media Strategies

A variety of communication strategies are described throughout this chapter, but Internet and social media strategies deserve a dedicated summary. Your evaluation might not go as viral as an English kid biting his big brother's finger, but social media can be an effective tool to get your evaluation out to a large and diverse audience. If you aren't familiar with the mechanics, don't throw up your hands and ignore the medium altogether. Take a look at the examples below for ideas about how your evaluation findings might fit with a social media strategy. You can put together materials such as Word docs, PowerPoint, pictures, and videos and then work with your marketing department to get the word out via the Internet and social media.

Tools Section: Summative Evaluation Checklist

This Summative Evaluation Checklist is derived from an extensive analysis of all evaluation reports posted to InformalScience.org. The Building Informal Science Education (BISE) network, which aims to create deeper connections between evaluation and practice, conducted the analysis as part of its efforts to identify insights that can inform the field as a whole. Think of the cross-cutting questions we could ask if we had the ability to slice and dice the database to look at specific audiences or subject matter across all project types. Or conversely, an in-depth look at all reports pertaining to a specific project type such as "exhibitions" or "media" would also be informative.

The BISE network's first step was to review all of the existing reports and code them from the ground up. This initial analysis revealed that the summative evaluation reports often omit surprisingly basic information, making it harder to categorize reports by target age or other factors that might cut across reports from different projects. Frankly, missing information can make it hard to understand the report at all, which is why we urge you to think about all of the possible audiences for your report from the very beginning. Summative evaluation reports are often written by the evaluator with only the PI

in mind, so information they both know sometimes doesn't get documented in the reports. The checklist below can help you to cover the basics. Consider adding it to the scope of work when you first establish your evaluation contract, and consult it again when it's time to post it at *InformalScience.org*.

Sample reporting and dissemination formats

Figure 15, see page 66.

Conclusion

Here are just a few organizations with annual national conferences that you might consider in getting the word out about your project and its evaluation. Use these as food for thought, not as an exhaustive list.

Suggested Annual National Conferences

American Education Research Association
American Evaluation Association
Afterschool Alliance
American Association of School Administrators
Association of Science-Technology Centers
Coalition for Science After School
United States Conference of Mayors
Council of the Great City Schools
Ed Trust
Education Commission of the States
Grantmakers for Education
National Afterschool Association
National Association of Broadcasters
National Association of Elementary and Secondary Principals
National Council of State Legislatures
National Dropout Prevention Network
National League of Cities
National School Boards Association
National Science Teachers Association
National Summer Learning Association
National Council of La Raza
PTA Conference
Society for Research on Adolescence (biannual)
TED

Thoughtful use of your evaluation findings will put you well on your way toward maximizing the impact of your project. We hope that you will think ahead about the audience for your project results and build dissemination into your evaluation plan from the start. Consider coming back to these suggestions, checklists, and examples at key points in your project. We have assembled them to help you plan for an evaluation report that gives you a clear picture of your project results and serves as a springboard for your dissemination efforts.

Now that we've armed you with a smorgasbord of tips and tools for making the most of your evaluation findings, we'd like to wrap up this section with a thank you to the many colleagues who, in the true spirit of dissemination, shared their hard-earned insights and wisdom to advance the success of future projects such as yours.

Sample reporting and dissemination formats

FORMAT	AUDIENCE	LINKS TO MODELS/SAMPLES
Executive summary	Funders, Project Team, ISE field (InformalScience.org)	Summative Evaluation of the Skyline Exhibition
Full Report	Funders, Project Team, ISE field (InformalScience.org)	Secrets of Circles Summative Evaluation Report
Project Highlights	Funders, Colleagues	http://www.cdm.org/i/Resources/Secrets-of-Circles.pdf http://www.philasciencefestival.org/node/548
PowerPoint	Co-workers, Funders	View Powerpoint
Conference Papers and Presentations	ISE field (colleagues)	http://www.museumsandtheweb.com/mw2009/papers/schaller/schaller.html http://evolution.berkeley.edu/UToL/evans2010.pdf
Conference or seminar dedicated to project dissemination		http://sciencefestivals.org/conference.html
Peer-reviewed journal articles	ISE field (colleagues)	http://itestlrc.edc.org/sites/itestlrc.edc.org/files/EcoScience%20TechTrends%20article.pdf
Tweets	ISE field (colleagues)	http://twitter.com/informalscience
Blogs		http://blogs.edweek.org/edweek/curriculum/2012/04/governors_urged_to_tap_into_in.html
Policy Brief	Policymakers and Influencers; Consumers	http://www.nysan.org/files/3664_file_TASC_NYSAN_STEM_Policy_Brief_2012.pdf
Facebook		View Facebook Posts
Brochures and Annual Reports	Funders, Consumers	http://www.nysci.org/media/file/NYSCI_AnnualReport_2011.pdf

Figure 15

Author & Editor Biographies

Rick Bonney

Rick is the director of program development and evaluation at the Cornell Lab of Ornithology, where he has worked since 1983. Some people think he was born there. He is co-founder of the Lab's citizen science program, and since 1991 has been PI, co-PI, consultant, advisor, or evaluator on more than 40 projects funded by the National Science Foundation. As a result he has extensive experience in developing partnerships between practitioners and evaluators to design and execute evaluation plans and disseminate their findings. Rick has been deeply involved in CAISE since its inception and was lead of the CAISE inquiry group that produced the report Public Participation in Scientific Research: Defining the Field. He is also on the board of directors of the Visitor Studies Association and is co-chair of VSA's communications committee. Rick received his BS and MPS degrees from Cornell University's natural resources department.

Kirsten Ellenbogen

As co-Principal Investigator of CAISE, Kirsten works in collaboration with the NSF to strengthen and advance the field of informal STEM education. Her work in evaluation and learning research has included service in several positions: Founding officer of the Informal Learning Environments Research SIG-American Education Research Association; affiliated researcher of the Museum Learning Collaborative; project director at the Center for Informal Learning & Schools, King's College London; senior associate at the Institute for Learning Innovation; and senior director for lifelong learning at the Science Museum of Minnesota. She was appointed to the National Academies of Science committee that produced the book Learning Science in Informal Environments and is past-president of the Visitor Studies Association, a network of professionals committed to understanding and enhancing visitor experience in informal learning settings through research, evaluation, and dialogue. Currently, Kirsten is President of Great Lakes Science Center in Cleveland, Ohio. Kirsten holds a Ph.D. in Science Education from Vanderbilt University and a B.A. from University of Chicago.

Leslie Goodyear

Leslie is passionate about the value that evaluation can bring to program planning, decision-making, organizational learning, and our understanding of human endeavors. She holds an MS and PhD in Human Service Studies from Cornell University, where her concentration was in program evaluation and research methods. Leslie's evaluation work has focused on building stakeholder capacity to manage and use evaluations effectively and on helping evaluators communicate their findings in dynamic and credible ways. She has evaluated programs ranging from HIV prevention curricula to services for adoptive families to civic engagement programs for youth to international youth media programs and afterschool initiatives. Most recently her work has focused on STEM education initiatives in both formal and informal settings. From 2009-2012 she served as a program officer in the Division of Research on Learning at the National Science Foundation, where she worked with the Informal Science Education (ISE), Innovative Technologies for Students and Teachers (ITEST), and Promoting Research and Innovation in Methodologies for Evaluation (PRIME) programs. She also contracted and managed evaluation studies for the DRL programs. She has served in leadership positions in the American Evaluation Association and is currently the Ethics Section editor for the American Journal of Evaluation.

Rachel Hellenga

Rachel develops exhibitions and performs exhibit-specific strategic planning and fundraising. She is a Chicago-based consultant and self-professed "Sam-I-Am" of evaluation owing to her many positive experiences working with professional evaluators over the course of a twenty-year career in the museum field. She has an insatiable appetite for visitor input, which has been reinforced by the results of integrating evaluation into projects such as the NSF-funded Inventing Lab and Skyline exhibitions at the Chicago Children's Museum, featuring a flying machine tower and construction materials replicated by other museums around the country; and the Science Storms exhibition at the Museum of Science and Industry, winner of the 2011 AAM Excellence in Exhibitions Award and the ASTC 2011 Roy L. Shafer Leading Edge Award. Rachel received her B.A. in psychology from Harvard University, and her particular areas of interest include education research in encouraging persistence; tinkering/making/engineering themes; Reggio-inspired design; bullying prevention; and novel uses of technology in exhibitions.

Jessica J. Luke

Jessica works in the Museology Graduate Program at the University of Washington, Seattle, where she teaches about and studies the ways in which museums can enhance quality of life. Jessica has a Ph.D. in Educational Psychology from the University of Maryland and a Master's degree in Museum Studies from the University of Toronto. She spent 15 years as a learning researcher and professional evaluator at the Institute for Learning Innovation, Annapolis, MD, where she designed and implemented dozens of evaluation studies in art museums, children's museums, science centers, and natural history museums across the country. In particular, her evaluation work has focused on the long-term impact of museum experiences for youth and families, as well as the development of critical thinking skills within the museum experience. Jessica has worked with a multitude of project PIs, clarifying project outcomes and developing strategies for enhanced communication of evaluation process and results. She also has conducted extensive evaluation training through graduate level courses and national and local workshops and seminars.

Mary Marcussen

Mary is a national grant writer and project design specialist with a reputation for high standards and professionalism in project and proposal development. Her record includes more than 40 successful proposals to the National Science Foundation to support museum exhibitions, planetarium shows, large format films, informal learning research, community and youth programs, and projects to build infrastructure for the field of informal science education. She is the former National Grants Manager for the California Academy of Sciences, prior to which she served as a systematic biologist and senior educator for the museum. With a B.A. in Biology, she has conducted field research for a variety of nonprofit and government agencies. Her development experience includes capital campaigns and high production corporate, foundation, and government grant work. She works with Principal Investigators to effectively manage both the people and the process involved with proposal development including research, project design, partnerships, and evaluation.

Sasha Palmquist

Sasha is a learning sciences researcher and professional evaluator of informal learning experiences. Over the last ten years, Sasha's work has focused on understanding how prior knowledge, interest, engagement, and personal identity shape learning opportunities and experiences in out-of-school and informal learning environments such as natural history museums, science centers, children's museums, and amusement parks. Sasha earned a BA in Psychology from the University of Pennsylvania as well as MS and PhD degrees in Cognitive Psychology from the University of Pittsburgh. As a Senior Research Associate at the Institute for Learning Innovation, she conducted studies that explored the development of scientific reasoning in complex domains including evolution and climate change. She has investigated the impact of children's interest and knowledge on family learning conversations in museums, identified challenges associated with developing and maintaining online communities of practice, and measured the impact of participatory design experiences on middle school students' STEM knowledge, interest, and engagement. Throughout these efforts, Sasha developed strategies for improving communication between researchers and practitioners that supported productive collaboration, facilitated evidence-based design decisions, and informed project outcomes.

Tina Phillips

Tina has extensive experience in developing, managing, and evaluating informal science education projects, with a particular interest in public participation in scientific research (PPSR). She is currently the Evaluation Program Manager at the Cornell Lab of Ornithology, where she is leading an NSF-funded project called DEVISE that is committed to building evaluation capacity within the PPSR field. As part of this effort, she is working collaboratively with evaluators and practitioners to provide guiding frameworks and contextually appropriate instruments for evaluating individual learning outcomes. She has written many articles on evaluation and was one of the authors of a landmark CAISE report: Public Participation in Scientific Research: Defining the Field and Assessing its Potential for Informal Science Education. Additional areas of concentration include formative and summative evaluations of machine learning experiences, website usability testing, and emerging research on understanding socio-ecological outcomes of PPSR. Tina holds a Master's in Education from Cornell University and is currently a PhD candidate at Cornell examining the relationship between citizen scientists' participation and outcomes related to knowledge, skills, and behavior.

Lauren Russell

Lauren specializes in leading coalitions of stakeholders toward envisioning and implementing informal science education projects. Currently director of grants and strategic partnerships at the Oregon Museum of Science and Industry, she has nearly a decade of experience working in science centers where she has worn many hats: Project manager; educator; external evaluator; professional development facilitator; and fundraiser. Lauren's belief in the value of partnerships between museums and their local scientific communities focuses much of her work on engaging the public with current research. Prior to joining OMSI, she led the award-winning Portal to the Public initiative at Seattle's Pacific Science Center. With funding from the National Science Foundation, this effort brings together scientists and science center visitors in personal, activity-based learning experiences—a framework now being implemented at science centers across the country. Lauren values cultivating effective working relationships among evaluators and practitioners. With an eye to both project success and advancing the field, these relationships support the shared ability of teams to complete evaluations that are both meaningful and useful to all stakeholders.

Saskia Traill

Saskia is the vice president of policy and research at TASC, a New York City-based organization reinventing learning opportunities for STEM and other disciplines. Saskia ensures that TASC's evaluation of more than 60 afterschool programs and expanded learning initiatives drive evidence-based policy and practice. Saskia also leads research and policy efforts for TASC's ExpandED Schools, a reinvention of urban public schools that brings together all members of the school and community to expand the day and increase learning options for students, including the integration of formal and informal science learning. She has co-authored articles, policy briefs, and reports on a range of issues, including engaging kids in STEM, how to fund innovative education strategies, and family economics. Saskia has served on the Public Policy Committee of the Society for Research on Adolescence and co-led the CAISE Policy Study Inquiry Group. She received her B.A. from Columbia University and a Ph.D. in research psychology from Stanford University.

Steven Yalowitz

Steven is a Principal at Audience Viewpoints Consulting, an evaluation and audience research firm specializing in informal learning environments such as museums, zoos, aquariums, and similar institutions. Prior to starting Audience Viewpoints Consulting, he spent four years as a Senior Researcher at the Institute for Learning Innovation, working on a variety of evaluation and research projects at a variety of institutions. He earned an M.S. in Experimental Psychology and a Ph.D. in Applied Social Psychology from Colorado State University, and spent seven years as the Audience Research Manager at the Monterey Bay Aquarium in Monterey, California, directing evaluations in the exhibits, marketing, programs, and guest services departments. While Steven's research and evaluation interests are broad, he has particular expertise in attitude and behavior change, visitor satisfaction, cognition and affect, climate change, and bilingual experiences in Spanish and English. He has worked extensively with science centers, natural history museums, art museums, and aquariums and zoos. In addition, he has worked on many technology-based evaluations of high-tech interactives, hands-on exhibits, labels, and web sites.

caise | center for advancement of informal science education

The Center for the Advancement of Informal Science Education (CAISE) works in collaboration with the National Science Foundation (NSF) Advancing Informal STEM Learning (AISL) Program to strengthen and advance the field of professional informal science education and its infrastructure by providing resources for practitioners, researchers, evaluators and STEM-based professionals. CAISE also facilitates conversation, connection and collaboration across the ISE field — including in media (TV, radio, and film), science centers and museums, zoos and aquariums, botanical gardens and nature centers, cyberlearning and gaming, and youth, community, and out of school time programs.

Established in 2007 with support from the National Science Foundation, CAISE operates as a partnership between core staff at the Association of Science-Technology Centers (ASTC) in Washington, D.C. and co-principal investigators at informal science and academic institutions across the country. CAISE is housed at ASTC's Washington, D.C. offices. ASTC is an international organization of science centers and museums dedicated to public engagement with science among increasingly diverse audiences.