THE THOUGHTFUL MUSEUM

Museum Evaluation without Borders: Four Imperatives for Making Museum Evaluation More Relevant, Credible, and Useful

HALLIE PRESKILL

Abstract   In this article, I invite readers to think outside of evaluation's current boundaries and to see the deep connectedness between what museums hope to achieve and how we evaluate the extent to which these aspirations may be realized. To do this, I present four imperatives for making museum evaluation more relevant, credible, and useful: 1) Link program activities with intended outcomes and hoped-for impact. 2) Take a systems-oriented evaluation approach. 3) Use affirmative data collection approaches based on assets and strengths. 4) Engage in courageous conversations.

It was my first big evaluation, early in my career. I was conducting a three-year evaluation of an innovative and highly publicized school reform effort in St. Paul, Minnesota. Having been educated and trained in evaluation approaches and methods by some of the most well known evaluators in the field, I felt confident that I had the knowledge and skills to do a good evaluation—one that would help the teachers, parents, district staff, and the school board make critical decisions about this innovative approach to education.

The evaluation's design and implementation were what I later came to call "traditional." It was framed as a formative evaluation, used multiple quantitative and qualitative methods, and resulted in a year-end report delivered in late August, at the beginning of each school year. However, throughout the three years, I found myself becoming more aware of, and then frustrated by, the boundaries of traditional evaluation practice. Having always believed that

evaluation should be helpful in making decisions, I came to realize that the evaluation results I was presenting each August were having limited impact on what was happening in the school.

There were several reasons for this. The school was innovative and progressive; it had received all kinds of state waivers to experiment with curriculums, teaching approaches, administration, and governance. It was even written up in *Time* and *Newsweek* magazines as the school of the future. The teachers were literally designing and developing courses, policies, procedures, and instructional practices every day. Nothing was static: the school was emergent, dynamic, and fluid.

As such, it was *too early for a formative evaluation*.

And each year-end report was delivered *too late to be useful*. There were few opportunities to share what I was learning (the findings) during the year; no attention was given to

Hallie Preskill (hallie.preskill@fsg-impact.org) is executive director of the Strategic Learning and Evaluation Center at FSG Social Impact Advisors, Seattle, Washington.

providing real-time data. As a result, the potential for using the findings for ongoing learning and programmatic refinements was severely limited.

In the school's first year, an interaction between a teacher and a para-professional resulted in claims of racism. Unfortunately, the school leader's efforts to resolve the issue were unsuccessful, and the organizational culture (for the instructional staff) became increasing mistrustful, hostile, and disruptive to the educational process. Although I was collecting data throughout this period, given my negotiated role as the formative evaluator, I felt powerless to help. The teachers made clear that I was the "historian" of the school and I should stay "objective," not attempt to help "fix the problem." While they hired three different organizational development consultants over the next two years, none was able to improve the relationships among the teaching staff and the resulting school culture.

This early experience had a profound and long-lasting effect on my thinking and career as an evaluator. When I look back at the topics I have researched, taught, written about, and presented on since then, I can see that I have been trying to push the boundaries of evaluation for many years. While some in the field may prefer to keep the borders around evaluation tight and fixed, I believe the future of evaluation is directly tied not only to reconfiguring the borders, but to making them more flexible, responsive, and permeable.

That much applies to evaluation in all kinds of educational nonprofit settings, not just museums. Now I'll turn to museums and offer some ways to think outside of evaluation's current borders and see the deep connectedness between what museums hope to achieve and how we might evaluate the extent to which these aspirations are being realized.

I believe there are four imperatives for making museum evaluation more relevant, credible, and useful:

- Link program activities with intended outcomes and hoped-for impact; start with the end in mind.
- Take a systems-oriented evaluation approach; it really is all connected.
- Use affirmative data collection approaches based on assets and strengths; problem-solving is like a dog chasing its tail.
- Engage in courageous conversations; it's the only way we will move forward.

IMPERATIVE #1

Link program activities with intended outcomes and hoped-for impact; start with the end in mind

Weil argued that the ultimate goal of a museum was to improve people's lives: "It is not the collection the museum houses, but what it does with its collection that matters."[1] Implicit in this statement is the notion that museums serve some social, intellectual, cultural, psychological, and/or economic purpose. But what exactly *is* that purpose? What role do museums play in communities? In society? What is it that museums *do* and to what end? And most importantly, from an evaluation standpoint, on what basis would a museum claim it is successful or is having an impact?

These questions could be asked of any social sector organization. Increasingly, they *are* being asked, in articles, newsletters, op-ed pieces, blogs, and conversations in a variety of sectors. Over the last year and a half, I have had the opportunity to meet with evaluators and program staff from more than 60 foundations

and non-profit organizations. Though I have been an evaluator for nearly three decades, I have been surprised at how many museum staff members have difficulty describing what the outcome would be if their work was successful.

Most people can easily articulate in general terms what they want to achieve overall, but they nonetheless tend to talk at that 30,000-foot level, and often have a hard time landing on specific outcomes and indicators of what success would look like tomorrow, next week, or even next year. It's as if no one has ever asked them to be explicit about what outcomes they expect or hope to see as a result of their efforts. Instead, there is a great deal of emphasis on the *doing* part—the activities that make up our everyday organizational experiences.

There are many reasons we have tended to focus on activities instead of outcomes and impact, among them the fact that most non-profit staff are constantly playing the mole game: We give everyone a hammer, and as one problem pops up, we attempt to fix it—to tap it down. Then another problem pops up, and again, we hammer it down. Our days are spent being reactive rather than taking time to slow down the process to think, reflect, or engage in conversations that matter. Museum professionals I have talked with have lamented that they are constantly "putting out fires," feel "like hamsters on a wheel," or are on a "nonstop treadmill." As a result, we often fail to consider what we are trying to achieve. Why is this exhibition so important? What kind of impact do we hope it will have? On whom? In what ways? What does success look like—and would we know it if we saw it?

Korn suggests that museums should be intentional about their work. She writes, "Museums that strive for intentionality operate from a set of carefully crafted intentions that are derived from and reinforce the museum's mission; they define and describe what the museum wants to achieve. They reflect and describe the essence of the museum and its unique value and potential impact on its community" (2007, 257). This issue—the tactical versus the strategic—falls squarely on the question, what is your theory of change?

It appears that museums, not unlike other nonprofit organizations in the social sector, have spent a lot of time and effort measuring the *outputs* of their work. It is not unusual to hear museum staff talk about attendance figures, or how many children have visited, or how many objects were acquired, or how many dollars have been raised, or how many people have visited a museum website. But, as Korn writes, "the value of museums, however, extends well beyond outputs" (2008, 1). For example, we can count how many people attended a museum or particular exhibition. However, attendance does not describe how visitors experienced the exhibition, or how their thinking or lives were affected, or what they will do differently as a result.

As evaluators we have to ask ourselves what we have to consider: What is our theory of change about an exhibition or other museum experience? What do we *expect* to happen? What do we *want* to happen? What do we *hope* will happen?

This means we need to be more strategic and intentional about the outcomes and impact we believe are possible through museum work. For example, what kinds of learning and behavioral change do we hope will occur as a result of a science museum's exhibition on energy conservation? What outcomes do we hope to achieve with an after-hours program for teens? What assumptions do we have about how an exhibition on the cultural heritage of a particular population will influence visitors' sense of self, pride, and place in a community? The

imperative for evaluators is to find ways of guiding stakeholders through one or more conversations and processes that will help them articulate their theory of change that makes explicit the desired outcomes and impact of their work.

One cautionary thought about outcomes is in order. While we would like to think that a museum's program or exhibition caused a change in the attendee or participant, we need to be thoughtful and realistic about the issue of contribution vs. attribution. This is especially important for evaluation, since the questions we might ask about an exhibition's contribution are very different than questions about causal effects (attribution). In most cases, proving a causal relationship between the work of museums and social impact is nearly impossible at the aggregate or population level. What is more achievable and meaningful is to show how museums *contribute* to the desired outcomes via results within individuals. However, to understand "contribution," we need to understand the context or system in which a museum exists.

IMPERATIVE #2

Take a systems-oriented approach: It's all connected

The notion of connectedness is critically important for measuring any kinds of outcomes, short- or long-term. For evaluation to be most useful, it needs to account for and connect to the system in which the program and institution operates. Thus, systems-oriented evaluations strive to do the following:

- Engage stakeholders throughout the evaluation.
- Use multiple methods of data collection.

- Provide real-time feedback when needed/appropriate.
- Collect data from a range of sources.
- Look for alternative explanations.
- Observe unanticipated outcomes and consequences (because we can't always predict outcomes in advance).
- Look inside and outside to gauge outcomes and impact.
- Acknowledge that cause and effect are often delayed; they are not always near in time and space.
- Communicate and report findings and learnings in ongoing and accessible ways.

A systems-oriented evaluation also pushes the boundaries of more traditional evaluation practice. Instead of looking at a program's component parts, we pay attention to how the parts influence each other and contribute to the program or project's overall impact. In addition, a systems-oriented evaluation evolves as the program changes in response to its internal and external conditions, and to the relationships between and among different actors. Consequently, an evaluation's design must be responsive to shifts in the initiative's focus, resources, needs, opportunities, and challenges (Coffman 2007; Parsons 2007; Williams and Iman 2007; W.K. Kellogg Foundation 2007).

With these ideas in mind, we need to consider:

- What is the system in which museums live?
- What are the boundaries of that system, and of the institutions within it?
- How do various relationships affect the outcomes of museum programs?
- What do museums hope to affect and how?
- What other actors or forces within our system can support or inhibit our work and its impact?

One way to answer these questions is to actually draw your system with a group of stakeholders, and wonder together what the implications for an evaluation may be. The answers should inform your evaluation questions, designs, and methods.

IMPERATIVE #3

Use affirmative data collection approaches based on assets and strengths; problem solving is like a dog chasing its tail

My commitment to Imperative #3 comes after many years of trying to help others believe in the value of evaluation, to not see it as threatening, burdensome, or a waste of time and resources. While there are many reasons for people's mistrust and fear of evaluation, I started to wonder if evaluation's value was being undermined, in part, because it is often viewed as a "problem-solving" activity: it reflects a deficit-based mental model. While I think most evaluators would say that they do not focus on the negative or look specifically for problems when conducting evaluations—that they are neutral—it is true that many of our stakeholders and evaluation participants have negative perceptions about evaluation and/or have had bad evaluation experiences. Regardless of how we describe it, evaluation is often seen by others as threatening, anxiety producing, and potentially punishing. In part, this is based on the language we use (as discussed earlier), but it also relates to where we turn our gaze—what we pay attention to. For example, evaluations often describe program and organizational deficits such as poor interpersonal relationships, low morale, ineffective leadership, silos, low attendance or participation rates, who is at fault, and burnout, to name a few.

Perhaps this emphasis on what's missing or broken is because of our evolutionary history, in which we needed to focus on problems for survival purposes. Or maybe it is a result of the influence of seventeenth- and eighteenth-century Western philosophers such as Locke, Hobbes, and Rousseau, who looked at social problems, enumerated the ills of society, and set out to fix them, one-by-one. Whatever the reasons, deficit-based language has permeated our vocabularies and affected our thinking for many years. Consider what was found in a recent review of the psychology literature. The researchers discovered that since World War II, approximately 200,000 articles have been published on the treatment of mental illness; 80,000 on depression; 65,000 on anxiety; 20,000 on fear; and 10,000 on anger; but only about 1,000 on positive concepts and capabilities of people (Luthans 2002, 697).

In response to this imbalance, the field of positive psychology has emerged. Researchers are finding "that when individuals or teams hear five positive comments to every negative one, they unleash a level of positive energy that fuels higher levels of individual and group performance" (Roberts, et al. 2005). Others, in the field of neuroscience, have found that when we shift from problem-focused thinking to thinking about possibilities, new cognitive connections (circuits) are made, and that when we intentionally work on feeling optimistic, our neural connections become strengthened and develop "muscles of optimism."

As evaluators, we are learning that we have a choice about whether we adopt a problem-oriented, deficits-based perspective or an affirmative perspective based on strengths and assets. The fundamental premise is that if you look for problems, you tend to find them, and create more problems. Conversely, if you look

for success, you are more likely to find and create more success.

There are some, however, who believe that a set of numbers and metrics will bring order, predictability, and control to our often-chaotic environment. The idea is that there is one truth, and the facts will set us free. However, the reality is that our thinking is guided by deeply embedded narratives, not numbers. When a numerical fact doesn't fit our conceptual frames or the metaphors we use to make sense of the world, we reject it. Moreover, we know that change is often inspired by emotional appeals rather than factual statements. Many evaluators can tell stories of times when action was taken on a policy or programmatic finding based on a story (a quote, a picture, a video) that stirred the stakeholders' emotions, rather than the rationally presented quantitative findings. The point here is not that we as evaluators should collect only qualitative data. Rather, I'm arguing that we need to design our evaluations using mixed-methods approaches that collect both quantitative and qualitative data, and to tell effective stories in our reports and presentations. There are a number of innovative approaches for collecting stories. In my own work, I have adapted Appreciative Inquiry processes to a variety of evaluation settings (Preskill and Catsambas 2006).

I've learned that when we engage stakeholders in an evaluation process that asks them to share stories about their experiences with the program being evaluated, amazing things can happen. Participants often learn that they have much in common while also discovering that their differences are valued. As a result, they are able to co-construct new understandings of their programs, their constituencies, their

organizations, communities, themselves, and each other. Consequently, those stories inform us about what works when, why, and how, and they help frame the evaluation's questions and design.

> **THE AIM OF THIS AFFIRMATIVE APPROACH TO EVALUATION IS TO HELP CREATE CONTEXTS IN WHICH PEOPLE FLOURISH. SUCH EVALUATIONS BUILD ON PARTICIPANTS' SUCCESSES AND POSITIVE EXPERIENCES.**

The aim of this affirmative approach to evaluation is to help create contexts in which people flourish. Such evaluations build on participants' successes and positive experiences, and identify times when their interests, creativity, and passions resulted in pride and excitement, so that the organization can change to do more of the things that made these experiences possible. Not surprisingly, this entails changing the nature of the evaluation questions. Diana Whitney and her colleagues have written, "If you truly wish to change your world, you must change your way of asking questions. It could be that the moment you do so, a totally different world will take shape around you" (Whitney 2002, x).

IMPERATIVE #4

Engage in Courageous Conversations: It's the only way we will move forward

I recently attended a systems-thinking conference and had the good fortune to hear a talk by Peter Senge, who helped translate the theory of organizational learning and the learning organization into practice with his book

The Fifth Discipline nearly 20 years ago. In his presentation, Senge emphasized the importance of engaging in courageous conversations, and explained that the definition of courage is "openings, tearings of the heart." Senge suggested that courage is not a passive act, and that courageous conversations require energy. Moreover, he said, courageous conversations require one to be vulnerable, to expose oneself. Taking a stand is actually about being vulnerable and coming to grips with our own inconsistencies.

I believe that to do good and meaningful evaluations, we need to have courageous conversations. These conversations should take place with executive leadership, program staff, boards and trustees, museum members, community leaders, partners, funders, and other stakeholder groups. Those conversations would answer questions such as:

- What does this organization stand for?
- What really matters to me about this organization?
- What evokes profound respect?
- What does the museum seek to conserve?
- What does the community seek to conserve?

I believe that the answers to these questions will have profound effects on what, how, and for whom we evaluate the outcomes and impact of museum work. Who knows—if these courageous evaluative conversations actually take place, then perhaps we will achieve the future as envisioned in the Future of Museums study, which predicts that in 2034:

More museums will be places of cultural exchange in their communities; they won't have any other choice. Museums will be primary sites for civic dialogues about community interests and the policies that can affect communities. They will be one of the most powerful agents in helping all children understand the future and ensuring they are prepared to take leadership roles in various sectors (Chung, Wilkening and Johnstone 2008, n.p.).

It is time for museum professionals and others to think more strategically about evaluation and the ways in which it can help inform key decisions and actions. If we were to link program activities with intended outcomes and hoped-for impact, adopt a systems-oriented evaluation approach, use affirmative data collection approaches based on assets and strengths, and engage in courageous conversations, we might be surprised at how credible, relevant, and useful evaluation can actually be.   END

NOTE

1. See http://www.lukeweil.com/_pages/stevePage.html.

REFERENCES

Chung, J., S. Wilkening, and S. Johnstone. 2008. *Museums and Society 2034: Trends and Potential Futures*. American Association of Museums through the Center for the Future of Museums (December). Version 1.0.

Coffman, J. 2007. *A Framework for Evaluating Systems Initiatives*. Accessed at http://www.buildinitiative.org/files/BuildInitiative-fullreport.pdf.

Korn, R. 2007. The case for holistic intentionality. *Curator: The Museum Journal* 50 (2): 255-264.

_____ 2008. Transforming Museums—To What End? Proceedings: Transforming Museums Conference, Seattle, WA. Accesssed Oct. 2010 at http://www.randikorn.com/docs/transforming_to_what_end.pdf.

Luthans, F. 2002. The need for and meaning of positive organizational behavior. *Journal of Organizational Behavior* 6: 695-706.

Parsons, B. A. 2007. The state of methods and tools for social systems change. *American Journal of Community Psychology* 39(3-4): 405-409.

Preskill, H., and T. C. Catsambas. 2006. *Reframing Evaluation through Appreciative Inquiry*. Thousand Oaks, CA: Sage.

Roberts, L. M., G. Spreitzer, J. Dutton, R. Quinn, E. Heaphy, and B. Barker. 2005. How to play your strengths. *Harvard Business Review* 9 (Jan. 1): 1–6. Accessed Oct. 2010 at http://hbr.org/2005/01/how-to-play-to-your-strengths/ar/1.

W.K. Kellogg Foundation. 2007. *An Overview: Designing Initiative Evaluation. A Systems-oriented Framework for Evaluating Social Change Efforts*. Battle Creek, MI.

Whitney, D., A. Trosten-Bloom, D. Cooperrider, and B. Kaplin. 2002. *Encyclopedia of Positive Questions, Volume 1: Using Appreciative Inquiry to Bring Out the Best of Your Organization*. Lincoln, NE: Lakeshore Communications.

Williams, B., and I. Iman. 2007. *Systems Concepts in Evaluation: An Expert Anthology*. Point Reyes, CA: EdgePress of Inverness.