

THE CHRONICLE OF
PHILANTHROPY
Connecting the nonprofit world with news, jobs, and ideas

Opinion

April 5, 2007

How to End the Evaluation Tug of War

By Suzanne Callahan

ALSO SEE:

LIVE DISCUSSION: [Read the transcript](#) of an online discussion about how to evaluate nonprofit results with Suzanne Callahan, who has worked for grant makers and charities.

In the fall of 2005, a program I run called the National College Choreography Initiative announced a grant to Tulane University to bring two choreographers, Sara Pearson and Patrik Widrig, to its New Orleans campus.

They would work with students to create a dance about the environment, in collaboration with the university's Center for Bioenvironmental Research, to be performed on the Mississippi River.

Then Katrina hit, the levees broke, and that didn't happen.

On the spur of the moment, with the help of faculty members at the University of Texas at Austin, students were transported to Texas to attend classes there.

The choreographers could have abandoned the project, but instead re-envisioned it as *Katrina, Katrina: Love Letters to New Orleans*.

Through their art, they expressed their great love for the Big Easy, and students were led through a process that helped them cope with being torn from their homes, friends, and college.

When, during rehearsals, the choreographers asked students to comment on the terms "lost" and "found" in ways that related to Katrina, their thousands of journal entries became part of the performance itself.

Audiences were enthusiastic, and others in the arts world took notice. The Kennedy Center booked it — twice — for its free performance series and then made it part of the center's national outreach program. It then toured the country, including performances at schools in upstate New York that inspired students there to volunteer in New Orleans. And now it just might go to Japan.

One doesn't need to commission a major study to confirm the resounding success of this performance.

But a traditional evaluation would probably have shown the project as an utter failure. After all, the environmental project never happened, the goal changed, most of the original partners didn't participate, and the stated results weren't achieved.

Instead, the innovation on the part of artists and faculty members led to a transformative experience: They took stock of the situation and created an experience that would be meaningful to people during a time of dire concern about a national disaster.

Had Dance/USA, the organization that distributed the money, or the National Endowment for the Arts, which provided the funds, insisted on rigid accountability, the grant would not have been made, but Dance/USA released the funds and trusted the faculty members and artists.

How should we evaluate this project? Would its impact be captured by saying that 350 people attended, and that the budget balanced?

Or should we design a longitudinal study to find out how many kids returned to New Orleans? Or finished college?

These varying methods and questions illustrate the polarized thinking about evaluation that is affecting most areas of the nonprofit world.

We live in times of extremes: good, bad, right, wrong. More is better.

The syndrome of extremes is alive and well in philanthropy. The longer the list of results that grant recipients promise, the better. It's achieving them that's the problem.

Unrealistic expectations have set up a tug of war over the need for evaluations.

At one end are some nonprofit groups that wish they could avoid evaluation altogether. Resentful of the requirement, they call an evaluator after the program has finished to place an order for positive comments to insert in a final report.

Their chant is "isn't it enough if we did what we said we would?" Then there are the many organizations that overpromise in hopes of getting and keeping money.

Like hamsters on a wheel, staff members believe that somehow they can accomplish it all if they just work a little harder.

Their well-intentioned discussion about evaluation usually takes place right before a grant-proposal deadline and is then forgotten because there's a performance to produce, a national walkathon to coordinate, a disease to be cured, or a house to be built.

After all, that's why they are in business — it's their charitable mission.

Caught in the middle are well-intentioned grant makers who are sandwiched between nonprofit groups that can't deliver what they promised (or can't find convincing ways to talk about their work if they have) and a powerful foundation board that wants solid evidence of big results.

They call for detailed studies, lots of numbers, and seismic change that imply that the immediate results of a program — those that occur within a few months or a year — are not enough.

Their own predetermined measures may be adapted from the corporate world; they focus on the financial bottom line and assume that things such as joy, prevention, transformation, and beauty either need to be translated into a spreadsheet or don't matter.

One example: An after-school arts program for second graders should increase the reading scores for the school. The level of assurance that some grant makers want would require a crystal ball rather than an evaluation plan.

Somewhere in the shuffle are the ranks of evaluators who are supposed to bring more credibility and be the silver bullets to the accountability conundrum — or at least a quick fix to getting another grant. But they know enough about long-term studies to realize the kinds of change that can — and cannot — be attributed to a modest program.

And, at the end of the day, no one knows if, or how, the information is being used.

The solution, I think, lies in the middle. By lowering our expectations of evaluation efforts, we would actually raise our effectiveness at doing it.

Finding a middle ground would give reasonable assurance to grant makers that goals will be met, but allow for progress (or lack thereof) to be discussed in an honest environment. Flexibility is key, as Michael Quinn Patton, a prominent evaluation consultant, told the Stanford Social Innovation Review last fall: "Innovations by their nature are adapting solutions to conditions of uncertainty. When you want people to be innovative, what you need is a process that allows you to try out things, alter things, change things, get feedback, and adapt."

The process of meeting in the middle starts by taking a reality check. Nonprofit groups must develop measures that are appropriate and are based on their own shared values, collective experiences, and the context for running a program.

But it also means that grant makers must learn about and accept those measures, rather than to impose their own.

To do that, we must start with a realistic understanding of how programs operate, and what is possible — and what is not. We must look at what the project's sphere of influence is, relative to the organization's mission and resources, and resist the temptation to promise that any project by itself will change the world.

Then grant makers and nonprofit officials need to work together to gather relevant information to see if we got there or not. In the middle ground, nonprofit groups must make evaluation a priority, by allocating real staff time and resources to it, and conduct evaluations systematically. And grant makers must support the real costs for measurement, which adds a lot of extra work for their grantees.

The elephant in the room is the need for a deeper understanding about evaluation. It means that at least one person on the staff of the foundation and the nonprofit organization needs training. If grant makers understood evaluation, they'd realize that they sometimes are calling for a \$50,000 research process on a \$5,000 grant.

In the end, nonprofit groups that bite the bullet and conduct evaluations would probably realize they were more effective than they thought. I've never done an evaluation in which the end results weren't more powerful than initially assumed, or in which the staff members didn't hear thoughtful feedback about how to improve programs. We have to be ready to hear the truth. But grant makers can't punish grantees for not being fully successful.

An example of what can happen when grant makers and nonprofit groups meet in the middle can be seen in the effort to gauge the success of a collaboration between the Walker Art Center, in Minneapolis, and the Wexner Center for the Arts, in Columbus, Ohio.

When the two organizations commissioned a new performance from London's Improbable Theatre, they decided to take a simple approach: They asked audience members to answer surveys and participate in focus groups.

The results offered presenters valuable feedback on how audience members regarded the cultural groups that sponsored the theater performance, the artists learned what resonated with audiences in their performances, and the marketing department got useful ideas for promotion.

When we evaluated the National College Choreography Initiative program — including the Katrina project by Ms. Pearson and Mr. Widrig — we gathered statistics from all 35 projects. We learned that more than 10,000 students participated, 60,000 people attended 174 performances and 226 outreach activities, and more than \$665,000 was raised to match the money initially provided.

But the numbers don't tell the whole story.

We systematically reviewed the stories that came from faculty members at all 35 colleges. From them, we identified more than 30 indications of what had changed since the grants were made.

We know, for example, that two of the grants led to professional opportunities for students, including a job with a ballet company, and eight involved opportunities for a program to spread beyond the campus grounds, such as the creation of dance curricula for elementary and secondary schools.

Colleges also reminded us of their problems: Securing matching funds can be a challenge.

Each year, Dance/USA creates a publication that documents not just the statistics, but also the stories showing the signs of progress from each project. Being highlighted in a national publication builds credibility on a campus. And, the National Endowment for the Arts knows of the impact of its grant, and what was to be a one-time effort has now received four rounds of financial support.

Nonprofit groups and grant makers have to find ways to realistically evaluate how money was spent. Otherwise, the existing — but inappropriate — standards for other fields will continue to be applied to measure nonprofit success, and the anxiety about evaluation will prevail. Let's transform the evaluation craze from a burden into an opportunity.

We need to roll up our sleeves and do the hard work of evaluation, but also talk openly and realistically, and reconsider the way we measure change.

Suzanne Callahan founded Callahan Consulting for the Arts and the Laboratory for Arts Evaluation, in Washington.

Copyright 2010. All rights reserved.

- The Chronicle of Philanthropy
- 1255 Twenty-Third St, N.W.
- Washington, D.C. 20037