Lost in Translation

It’s time for people in the nonprofit community to reclaim the language of nonprofit evaluation.

BY KARINA KLOOS & DANIELA PAPI

The field of nonprofit evaluation is shifting, and with that shift has come a change in vocabulary. Terms like “randomized controlled trials” and “social return on investment,” for instance, have become central to the conversation. We take a moment to explore who is driving this conversation and how the nonprofit evaluation discourse may perpetuate ideas, expectations, and values that can be incongruous with the realities of the nonprofit sector.

With nonprofits under increasing pressure to monitor and evaluate their work, many of the proposed solutions are strongly influenced by scientific principles and the logic of “measurement,” “data,” and “evidence.” Business principles, too, and the logic of “investment,” and “returns,” have led to a preferred form of assessment. Scientific terms like “statistically significant results” or business terms like “outcomes” seem benign enough, but when we discuss social impact using phrases like “outputs to outcomes,” or “controlled experimental designs” it can conjure images of widgets in a factory, or a sterile scientific laboratory rather than real people in real communities.

Our main concern is that the predominance of managerial and scientific discourse is crowding out nonprofit voices. This has led monitoring, measuring, and reporting to become goals in their own right, often at the expense of genuine learning. Our aim here is to engender more thoughtfulness about the language used to describe and measure nonprofit performance and reflect on how to better align the evaluation conversation with the empowering of nonprofits to actually become more effective. In order to do this we believe that nonprofits should more actively participate in the conversation and set the agenda on evaluation.

WORDS MATTER
How we talk about nonprofit performance both reflects and reinforces particular values and practices. To illustrate that point, we draw from a recent study on nonprofit evaluation discourse. Four years ago, a team of researchers from Stanford University launched an effort to explore who is contributing to that discourse and what they are talking about. (One of us, Karina Kloos, was a cofounding member of the research team.)

The Stanford team began by exploring broadly who is contributing to discussions about nonprofit evaluation and then meticulously identifying those who appear most central to the overall conversation. Their methods yielded about 400 individuals and organizations that include funders, foundations, nonprofits, consultants, bloggers, associations, universities, and governmental institutions.

The team then explored what these contributors were saying in relation to nonprofit evaluation. Applying the method of content analysis, the researchers explored the website of each entity to determine how and how often each site used certain kinds of language. They identified a vocabulary of nonprofit evaluation that includes approximately 200 words and phrases, and they categorized each of those terms as belonging to one of three cultural domains. The managerial domain includes terms that have roots in the culture of business, such as “key performance indicators,” “social return on investment,” and “logic model.” “Evidence-based practice,” “control groups,” and “quantitative measurement” have a basis in scientific culture. And terms such as “constituent voice,” “social change,” and “empowerment” originate in the associational culture that prevails in the nonprofit sector.

The researchers’ findings demonstrate that although terms that reflect associational values remain prevalent in the discourse on evaluation, language that indicates a managerial or scientific orientation is now also very common. Management terms such as “outcomes,” “monitoring,” and “accountability” each appear in the website content of roughly 90 percent of the sample entities. The terms “data,” “evidence,” “methods,” and “survey,” all of which come from the realm of scientific investigation, appear with a similarly high frequency. Overall, in fact, “data” is used nearly twice as often as “mission,” “purpose,” “values,” “justice,” “empowerment,” “social change,” and many other important terms that
The critical attitude that comes with a focus on “monitoring” moves us away from the kinds of trusting relationships that facilitate learning.

The strong influence of managerial and scientific discourse is not particularly surprising given the diversity of entities identified by the Stanford researchers as key to shaping the conversation. What is surprising, though, is how small a minority of these entities are operating nonprofits that are subject to the kinds of evaluation that the conversation targets. As such, we are concerned that the findings in the study represent a marginalization of nonprofit voices.

VOICE LESSONS

We believe that nonprofits should take a lead in setting the evaluation agenda. To help nonprofit management teams reclaim their position in shaping the conversation on evaluation, we offer a few guidelines.

Talk about purpose | Foremost, every nonprofit team should create their own definition of success and have clear reasoning on how they plan to get there: their chosen strategies, projects, and programs. In the Stanford study, 66 percent of entities in the sample include specific reference to a “theory of change.” Our view is that all nonprofits should have a clearly defined theory for how they will create change that connects their strategies and programs to the results that they anticipate. If they fail to develop and clearly communicate their purpose—along with a set of strategies, programs, and assessment plans to support it—then funders will define those elements for them.

Talk about people | Funders often require nonprofit assessment reports to include quantitative aggregations: the number of clients served, the number of people indirectly affected by the work of an organization, the number of lives saved, and so forth. Yet that quantitative focus can reduce the impact of nonprofit work to a series of numbers. Qualitative assessments that draw on conversations with people are often more consistent with how nonprofits operate, and they are also a methodologically valid form of evaluation. The Stanford researchers noted that 60 percent of entities in the sample mention use of “control groups.” For nonprofits, though, the use of focus groups—as well as client interviews and other forms of ongoing interaction with beneficiaries—is an equally relevant method of assessment that leverages nonprofits’ daily work and supports values of empowerment through participation and bottom-up problem solving.

Talk about the big picture | Each nonprofit should identify and articulate how its efforts fit into a broad system of social change. In order to support systemic progress, nonprofit evaluation should include how well each organization is contributing to collective transformations. In the Stanford study, the influence of competitive, market-based thinking was evident in the prevalence of terms such as “value proposition” (used by 55 percent of entities in the sample) and “expected return” (45 percent). Use of such terms reinforces the tendency to focus on organizational effectiveness and organizational outcomes, encouraging individual organizations to compete with each other by highlighting their own success. But as most nonprofit leaders know, taking full credit for changes that involve many factors (including the work of other nonprofits and community leaders) puts an inordinate emphasis on the organization as the unit of analysis and crowds out more systemic perspectives.

Talk about challenges | Discussion of failure, and of lessons learned from challenges, should be a central aspect of nonprofit assessment. Assessments shouldn’t be about proving if something worked or not, but rather understanding the context of successes as well as failures. Nonprofits should be able and willing to talk openly about tactics or projects that have fallen short of their goals, and the insights gained from the experience. Toward that end, we should encourage greater transparency rather than monitoring on its own. The vast majority of entities in the Stanford research sample mention both “monitoring” and “transparency” (around 90 percent for both terms)—yet, in total, monitoring appears more than four times as often as transparency. If nonprofits are honest about their missteps they will be in a better position to manage expectations and facilitate collective learning.

Talk about learning | To foster a spirit of learning, we would invite more listening—in particular, to the voices of beneficiaries and nonprofit workers. Too often, however, the most influential voices in evaluation come from funders who talk less about listening and more about “monitoring.” Consider the phrase “monitoring and evaluation.” It has become so widespread that people simply use the abbreviation “M&E.” (In the Stanford study, 41 percent of entities in the sample use that shorthand term.) The critical attitude that comes with a focus on “monitoring” moves us away from the kinds of trusting relationships that facilitate sharing, listening, and learning. Perhaps the more appropriate term is “learning and evaluation.” In fact, the bottom-line question in any process of nonprofit evaluation should be, “What are we learning from this evaluation and how can that be used to help improve our collective work?”

The conversation on nonprofit assessment should certainly draw on ideas from other fields, but people in the nonprofit community should drive that conversation. When talk turns to “randomized controlled trials” and “social return on investment,” without discussion of “collaboration,” “participation,” and “lessons learned,” the real purpose of nonprofit work gets lost in translation.
Copyright of Stanford Social Innovation Review is the property of Stanford Social Innovation Review and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.