FACING EVIDENCE FEARS

FROM COMPLIANCE TO LEARNING

CHRISTOPHER SPERA

While there are hundreds of programs across the United States to tackle a variety of health and social policy topics, as Jon Baron, formerly of the Arnold Foundation states, "U.S. social programs, set up to address important problems, often fall short by funding specific models/strategies ('interventions') that are not effective."[1] A program is a strategy or intervention that has been provided to a group of people to achieve a desirable consequence; in short, programs are developed and implemented to solve a social problem. Given this, the question becomes how society at large can implement more high-quality program evaluations to ensure the public is receiving interventions that are proven to work or, at a minimum, hold promise to work. More specifically, the question becomes: How can the field use a learning framework (versus a compliance framework) to evaluate programs and continuously build evidence with a research and development mindset? Here, I tackle this by looking at three core issues: 1) common challenges to conducting evaluations; 2) viewing evaluation in a learning versus an accountability lens; and 3) increasing internal capacity of organizations to evaluate their own programs and/or oversee a high-quality external evaluation. I will discuss practical experiences and observations I have had in my time as an evaluator in the field for over twenty

years at ICF International (formerly Caliber Associates), AmeriCorps, and Abt Associates.

Common Challenges to Conducting Evaluations. While there are numerous challenges to conducting evaluation, in my experience, there are four core ones that need to be tackled to advance the field. The first common challenge is the notion or feeling from several stakeholders that evaluation is too expensive and will divert resources from program delivery. I have heard this notion repeatedly, especially during my time as the director of research and evaluation at AmeriCorps (formerly the Corporation for National & Community Service). The general feeling is that social programs are underfunded from the get-go, so why would you take additional funds away from program delivery to evaluate the program? My response to folks who ask this is to ask them whether they would invest their personal savings in investments without any data on whether those investments would generate a return. They often then quickly seem to understand. In general, we need to continue to look for ways to generate investments in evaluations—which include everything from congressional set-asides to philanthropic investments to enhanced internal capacity within organizations that require evaluation. In short, there is proportionately very little money for evaluation compared to the amount of money put into program delivery; my estimate is less than a tenth of 1 percent, but that is just a guesstimate. It should be somewhere between 3 to 5 percent to really have a cadre of evidenced-based programs to turn to when we need them.

The second limitation is the belief by stakeholders that evaluation is too complex and too hard to understand to put to practical use. Randomized controlled trials are somewhat easy to understand in the sense that one group gets the intervention/program and the other does not and you compare the differences in outcomes. However, once you enter into the world of quasi-experimental designs that require statistical matching and techniques to generate impact estimates that control for extraneous variables, it becomes hard to understand for policymakers and others not trained in evaluation techniques. In short, we need to find a way to educate folks or develop reports that provide more concise evaluation findings and methods in a way that is easier to understand. There have been pushes by federal agencies, such as the Centers for Medicare and Medicaid Services (CMS), to make some progress here in terms of presentation of findings in infographics. Despite this, researchers are incentivized to publish in

academic journals, and sometimes the incentives to push toward action-able evidence in user-friendly language are less than needed.

A third limitation is the belief that evaluation will burden staff. It is true that program staff, especially those doing service delivery, are overburdened, but evaluations often can be done with minimal interference with their daily duties and responsibilities. I have found that evaluations that rely on staff administering measures to participants often can result in too much for staff to handle, which leads to issues; however, sometimes this is neces-sary to complete an evaluation.

The most common challenge to the use of high-quality evaluations is the fear by program staff and funders of what happens in the event of "negative" or "null" findings. The concern is related to the implications for the program—up to and including elimination—if the evaluation shows it generates no impacts. This creates the feeling that the program is taking a risk by conducting an evaluation. The counterforce to this over the last several years has been program funders simply requiring evaluations to continue their funding, making the risk-reward scenario very different. In addition to this counterforce, viewing the program evaluation from a learn-ing versus an accountability lens becomes very important, which is our next topic.

Viewing Evaluation through a Learning versus an Accountability Lens. When I teach program evaluation each fall semester at Carnegie Mellon Universi-ty's Heinz School of Public Policy, I tell my students early in the semester that there is an inherent tension between using evaluation for learning and program improvement versus accountability and funding decisions, describ-ing it as a tug of war. This is not a new concept as Michael Quinn Patton (2008) described utilization-focused evaluation as designed to answer spe-cific questions raised by those running the program,[2] and Michael Scriven (1980) described judgment-focused evaluation as the technique evaluators use to make determinations about the value or worth of programs.[3]

I recall when I first started in my role as the head of program evaluation at AmeriCorps, program directors of national programs ran the other way when seeing me in the hallway for fear I would evaluate their program and hold them accountable. I quickly changed this by meeting with them and asking them questions like: What do you want to learn about your program? What keeps you up at night? How can I help? This quickly changed their attitude toward evaluation, and they began to see it as a tool for learning and

program improvement rather than an up-or-down vote on their program. This takes trust that can be built only over time, but it worked then.

I also know that several other federal agencies have begun to develop and implement learning agendas designed to have evaluators create a collective roadmap for helping programs better understand their progress. While there always will be an accountability lens for programs due to the funding going into them, using the learning lens is more productive and can be used to generate more evaluations in the long run to determine what works. I would argue that program evaluation should be used for program elimination only when programs have run their course trying to improve fidelity and program delivery and cannot make any progress with impact. Very few are at this point, and viewing evaluation as a learning tool can help make progress.

Increasing Internal Capacity of Organizations to Evaluate their own Programs and/or Oversee a High-Quality External Evaluation. A final challenge to nurturing more high-quality evaluations is the lack of internal capacity within organizations that need to evaluate their program. I see this routinely in my work at Abt Associates and saw it also in my former work at AmeriCorps. For a program to implement an evaluation, it requires an individual(s) within the organization who can design an evaluation solicitation, hire a local or national evaluator, and oversee their work. External evaluators also need help from internal teams and stakeholders to design a strong evaluation and interpret results when they are available. An alternative option to an external evaluation is for a unit within the organization that manages monitoring, evaluation, and learning to be responsible for conducting an internal evaluation. Either way, this requires evaluation capacity within the organization. In my experience, there is a paucity of evaluation professionals within the federal and state governments relative to the need (except for a few exceptions, like the Department of Education) and even more so in the philanthropic sector. The recently completed Social Innovation Fund (SIF) requires internal capacity within organizations, which is a promising shift. Building internal capacity will be key moving forward.

In conclusion, in my twenty years within the program evaluation field, I have seen significant progress. More programs have been evaluated, evaluation demand has spiked and surged, especially during the Obama administration, and new techniques such as rapid cycle evaluation have been woven into the fabric of the field. With that said, when you compare the progress the field of program evaluation has made in twenty years and compare it

with the progress made in technology, for example, I feel like progress has been slow. Too many policy and programmatic decisions are made based on personality and politics versus evidence and data. As we move into the next decade of "big data everywhere," I am hopeful the field will begin to gallop ahead and the use of evidence to drive major program and policy decisions will become the norm versus the outlier. As a field, we need to embrace this new wave of data everywhere, regardless of whether the data was initially gathered for research purposes, and more quickly harness it to improve programs that help the lives of our fellow citizens. I am excited and optimistic for the future!

NOTES

1. Prepared Statement of Jon Baron, President of the Coalition for Evidence-Based Policy, at a hearing before the U.S. House of Representatives, July 31, 2013, https://www.govinfo.gov/content/pkg/CHRG-113hhrg81981/html/CHRG-113hhrg81981.htm

2. Michael Quinn Patton, *Utilization-Focused Evaluation* (Newbury Park, CA: Sage Publications, 2008).

3. Michael Scriven, *The Logic of Evaluation* (Iverness, CA: Edgepress, 1980).