Center for
Evaluation Innovation

Systems Change Evaluation Forum
Executive Summary

By Tanya Beer, associate director, Center for Evaluation Innovation

In May of 2017, the Gordon and Betty Moore Foundation convened staff and external evaluators for a Systems Change Evaluation Forum. In addition to establishing a common understanding of systems change evaluation, the participants grappled with four practical dilemmas that systems change efforts create for evaluation:

1. How do we better understand our unique contribution to change when there are so many other actors influencing the system?
2. How can we select and capture meaningful interim outcomes to understand systems change progress when we are operating in short time horizons and using strategies that have indirect impact?
3. How do we set realistic boundaries around the aspects of the system that our strategy and our evaluation should focus on?
4. How can we embed continuous, intentional learning into our strategy and evaluation work?

In our experience as evaluators who work in the philanthropic sector, these dilemmas are common among systems change funders who are also committed to evaluation. Foundations' organizational structures, practices, and processes often feel at odds with the lengthy time horizons and unpredictability of systems change. Boards tend to want initiatives that will produce durable systemic change but also request definitive, quantitative, easy-to-digest data about impact within short time frames. Program staff have to build a robust case that a particular strategy will result in specific changes at pre-determined points in time, even though they are uncertain about which of many possible strategies are most likely to produce those results and when. Boards and staff alike are committed to achieving results, yet find that the traditional accountability mechanisms historically provided by monitoring and evaluation are not up to the task of providing assurance that they're on the right road.

Although handling the dilemmas that come along with systems change work requires a broader shift in mindsets (e.g., getting comfortable with uncertainty), there are also practical tools and approaches that can help foundation program and evaluation staff manage them. Below is a short summary of the observations, tools, approaches and advice for doing so that emerged from the convening at the Moore Foundation.

Dilemma #1: How do we better understand our unique contribution to change when there are so many other actors influencing the system?

System change strategies are often composed of many different interventions that are intended to work *together*—or to complement work supported by others in the field—rather than working in isolation to

create change. Experimental or quasi-experimental research to isolate and establish causality is arguably ill-suited for many types of systems change interventions. It is difficult, if not impossible, to create robust counterfactuals under conditions where many actors are influenced by a wide range of forces in often unpredictable and invisible ways. In cases where we cannot attribute change to a single actor or intervention, what standard of evidence should we establish with our evaluation designs, and what methods and approaches can help us confidently conclude how the foundation has affected change?

Potential solutions to explore

Evaluation of systems change efforts requires rigorously testing whether a plausible case for <u>contribution</u> can be made, rather than attempting to isolate and attribute a systems-level change to a single intervention or initiative. Several evaluation designs and methods exist to accomplish this:

- **Process Tracing**[1]: A case-based approach to causal inference which focuses on the use of clues within a case (causal-process observations) to adjudicate between alternative possible explanations.
- **Modus Operandi**: A theory-based approach in which the evaluator builds a picture of predicted effects and supplements/tests this picture through research.
- **Propensity-score matching**: A quasi-experimental approach to estimate the difference in outcomes between beneficiaries and non-beneficiaries based on a sampling method that matches the characteristics of each person or case who received the intervention as closely as possible to those of a person or case who did not (the 'control').
- **Ripple effect mapping:** A participatory method that is best conducted for in-depth program interventions or collaborations that are expected to produce broad or deep changes in a group, organization, or community and to have "knock on" effects that are important to understand.
- **Contribution Analysis:** An approach for assessing causal questions and inferring causality that reduces uncertainty about the contribution the intervention is making to the observed results through an increased understanding of why the observed results have occurred and the roles played by the intervention and other internal and external factors.
- **Qualitative Comparative Analysis:** A means of analyzing the causal contribution of different conditions (e.g. aspects of an intervention and the wider context) to an outcome of interest across a number of cases.
- **Outcome harvesting:** A multistep inductive evidence gathering process designed for situations where strategists have a high level of uncertainty about what outcomes might occur as a result of an innovative or highly emergent set of interventions.

Understanding a foundation's unique contribution to change requires more than looking for confirming data on indicators identified in advance (a common practice for evaluators). Those indicators alone tell us little about *why* they occurred, or about what else occurred at the same time, either positive or negative. Instead, plan to cast a wide net for both intended and unintended outcomes and systems changes, and then use an evaluation approach that allows you to test and explore competing explanations for what is occurring. Consider establishing principles for what kinds of evidence will pass muster, given the nature of the strategy.

---

[1] All descriptions of methods, tools, and approaches in this summary are quoted or paraphrased from www.betterevaluation.org where possible, or other introductory sites, with the relevant page hyperlinked to the text so that readers can easily learn more about it.

Build capacity and knowledge among staff on the basics of methods and design beyond experimental designs and quantitative data to increase the range of options to choose from to find a fit with your strategy.

Dilemma #2: How do we select meaningful interim outcomes to understand systems change progress in short time horizons and for strategies that have indirect impact?

System changes efforts rarely have direct impact on their ultimate beneficiaries. Instead they focus on changing the behavior of other institutions or actors in the system. (Consider, for example, focusing on changing payment incentives in the health care system to shift how providers deliver care to improve patient outcomes.) Additionally, when the larger context is shifting around a particular strategy –economies boom and bust, who has political control changes with each election, major funders come and go from the system—sometimes indicators will look worse before they look better. In many foundations, program staff are managing all of this shifting terrain while boards are looking for steady upward-trending progress on a small handful of narrow quantitative outcome indicators. Given this situation, how can staff select meaningful interim outcomes and indicators that give reasonable assurance that the system is headed in the right direction and that they are making a difference? When staff are working towards outcomes that will likely take years or decades to materialize, but are judged on their progress in 3-to-5 year cycles, what should they be measuring to understand progress?

Potential solutions to explore

- **Use and adapt existing systems change frameworks.**
  Adapting existing frameworks can help identify where a strategy should focus, anticipate how outcomes and indicators might evolve over time, and clarify the range of outcomes that could be of interest at different stages. Examples include the **Disrupting Systems Dynamics framework**, **ABLe Change Framework**, and the **Consolidated Framework for Implementation Research**. All can be carefully adapted for systems change in a range of content areas.

- **Consider using predictive methods to "pressure test" the value of interim outcomes.**
  Qualitative and quantitative methods that predict whether future changes might occur can be modified to help increase our confidence that the interim outcomes already achieved will, in fact, improve the chances that a future outcome will be achieved. For example, explore **limiting factor analysis, scenario planning,** and **agent-based modeling**.

- **Avoid creating rigid or overly detailed outcome frameworks at the outset.**
  While the discipline of thinking through interim outcomes in advance can produce stronger plans, over-doing it can be a waste of time in the face of a changing environment. Instead, consider beginning with a handful of outcomes and indicators that are closely linked to the initial stages of work and then build in a regular review of outcomes and indicators to revise and expand as the initiative evolves.

- **Avoid presenting and interpreting individual outcomes and indicators in isolation.**

Use *packages* of outcomes and indicators, always created, presented and interpreted together and in context, to acknowledge that systems change can't be understood with a few simple numbers. Use data as the starting point for discussion and exploring progress, given what else is happening in the larger system, rather than treating it as the conclusive end point that tells the whole story.

- **Embed interim outcomes in a research-based model or theory of change to illustrate their link to longer term outcomes.** Research-based theories can be used to illustrate how interim outcomes are at a critical mid-point in a long pathway of change. See, for example, John Mayne's "causal package" framework or "punctuated equilibrium" theory. ORS Impact's 10 Pathways to Policy Change offers research-based theories of how policy change occurs.

Although in a systems change effort interim outcomes are never definitive predictors of future success, they can be explored, tested and presented in ways that increase confidence that an initiative is on the right track. However, learning how to interpret packages of indirect outcomes requires a different mindset and skillset than most boards and staff are accustomed to. Re-designing how data are presented and discussed is an indispensable part of making this shift.

Dilemma #3: How do we set realistic boundaries around the aspects of the system that our strategy and our evaluation should focus on?

Complex systems often have "fuzzy" boundaries, meaning that many different factors that are outside the purview of an initiative could have an effect on its success (e.g., an initiative may be focused on changing how university faculty collaborate across disciplines but state budget cuts could lead to faculty shortages that re-direct faculty attention to teaching and research within their core discipline). Practically speaking, this means that the range of factors that affect how a system operates could be enormous. Setting the boundaries that determine which factors are too far afield to address with a strategy or monitor through an evaluation can easily lead staff to feel overwhelmed. How can staff know when they are addressing a fundamental root cause that may seem distant but is in fact a significant leverage point, versus when they have gone too far afield?

Potential Solutions to explore

In general, boundary setting is a matter of experimentation and judgement, as well as the simple reality of what a board of trustees will accept as a reasonable change lever. However, there are some approaches that can help.

- **Critical Systems Heuristics** – a framework of 12 "boundary" questions designed to outline and provoke systematic thinking about boundary judgements. They help strategists determine who should be in their frame of reference as they examine how the system works and where to intervene.

- **Attractor Mapping** – operating much like a heat map, attractor mapping helps to distinguish areas of intense activity (and thus opportunity or energy for change). Though this doesn't set definitive boundaries, it can support decision making about where to intervene and what to leave behind.

- **Pre-mortems** – a predictive (and fun) exercise to identify the exogenous factors to which the initiative or the system as a whole is most susceptible. Premortems can help point out during the planning phase what factors need to be addressed, and can flag which factors to monitor because they could cause failure if they do not move in a particular direction.

Regardless of the boundary setting approach, it is good practice to include more explicit descriptions of what falls outside the boundary in a strategy description. Oftentimes, boundaries are only implicit. The practice of naming it more explicitly is a good way to reveal underlying assumptions about drivers and opportunities. Some foundations set too tight a boundary around their *evaluation* by focusing exclusively on metrics that are directly tied to a specific intervention. This can cause staff and evaluators to misunderstand why something isn't working or what's driving progress.

Dilemma #4: How can we embed continuous, intentional learning into our strategy and evaluation work?

Complex systems change requires tools for smart sensing and navigation, as well as a habit among program staff of regularly naming assumptions, testing them in action, and adapting according to what is learned. However, staff managing initiatives with lots of moving parts struggle to find the time and space to systematically reflect. They also often find that data that is collected to satisfy grant monitoring needs provides little real insight into what's happening. How can a foundation improve how they capture, process, and interpret data that is actually useful for decision making? How can foundations build the informal attitudes, norms, habits of working, and the know-how that is required to engage in genuine inquiry. We all learn constantly, but learning in teams and learning systematically in a way that accelerates progress is much more of a challenge.

Potential Solutions

A considerable amount of thinking and work has been done on learning in philanthropy and learning in systems change efforts in particular. While learning individually can be an intuitive process, learning in teams requires more intentional structure and process to making thinking and assumptions visible, draw meaningful insights that can be shared and applied in different contexts, and pass along knowledge as staff change over time.

- **Build organizational incentives for learning, and diagnose and address disincentives**. Formal incentives (e.g., building learning expectations in program staff job descriptions and performance reviews) and Informal incentives (e.g., verbal recognition and reinforcement from leadership and trustees for staff who are more candid about challenges) can go a long way to building learning habits. According to

program staff time is the biggest barrier to learning, so consider reviewing workloads to determine whether staff can realistically pause for systematic reflection.

- **Re-cast existing meetings, processes and structures to foreground more intentional learning rather than treating learning as an add-on task.** Team or staff meetings that include task updates can be retooled around strategic questions, or thoughtfully structured to include outside/alternative perspectives to avoid groupthink. Consider using learning tools: Facilitating Intentional Group Learning: A practical Guide to 21 Learning Activities by FSG, Before and After Action Reviews, Emergent Learning Tables, or methods for reducing *Cognitive Traps in Philanthropic Decision Making.*

- **Treat learning as an indispensable part of the strategy—and as a critical outcome—from the outset of the strategy design process.** Articulate during the strategy design process how the team and grantees will use data and evaluation to learn over time and to adapt other components of the strategy and/or to influence other actors. When learning is viewed as a critical outcome of the work for which staff are accountable, they are more likely to commit serious time and energy to it.

- ✓ **Consider developmental evaluation where appropriate.** Developmental evaluation is an evaluation approach explicitly intended to support innovation, where people are working under conditions of high complexity and the way forward is uncertain. The evaluator's role is to answering real time questions with rigorous data about what's unfolding so that the team can make decisions about next steps.

- ✓ **Hone the skill of asking the right questions.** Organizations whose learning returns the most value are those who pose meaningful questions that, when answered, make a real difference in their ability to achieve the results they want.

Rather than creating large-scale learning systems and foundation-wide processes to improve learning all at once, try setting up a small voluntary working group of staff who will experiment with different learning methods for a period of time to see if it increases the depth of reflection and provides them with more strategy direction. It can also be helpful to focus on getting 2-3 concrete "wins" on high-return learning, as people will carve out the time and resources for more intentional learning if they believe it will actually make a difference in their results. Having an ongoing relationship with a learning-focused evaluator through the course of complex initiative could help the staff ask and answer better strategy questions and result in higher quality data and findings for the board.