

EVALUATING SOCIAL INNOVATION



Hallie Preskill and Tanya Beer

2012



EVALUATING SOCIAL INNOVATION

The front edge of the philanthropic sector has spent the last decade experimenting with innovative grantmaking in the hopes of triggering significant and sustainable change. But the sector's approach to evaluation is not keeping pace with these innovations. In many cases, traditional evaluation approaches fail to meet the fast-paced information needs of philanthropic decision makers and innovators in the midst of complex social change efforts. At worst, the application of traditional evaluation approaches to innovative change initiatives may even decrease the likelihood of success because they restrict implementers to pre-set plans that lose their relevance as the initiative unfolds.

In this paper, we explore ways that common evaluation approaches and practices constrain innovation and offer lessons about an emerging evaluation approach—developmental evaluation—which supports the adaptation that is so crucial to innovation. For what kinds of grantmaking strategies should funders consider using developmental evaluation? What organizational conditions are necessary for it to work? How can grantmakers grapple with the challenging questions that developmental evaluation raises about innovation, accountability, rigor, and adaptation? Drawing on the reflections and insights of foundation staff and evaluators who have experimented with developmental evaluation, we call on philanthropy to re-envision the role, purpose, and processes of evaluation so that social innovations have a better chance of success.

Introduction

In recent decades, the philanthropic sector's efforts to produce bigger and more lasting results have been challenged by increasingly complex and dynamic problems with no clear path to a solution. The sector's literature and conference agendas—increasingly populated with systems- and network-oriented perspectives on social change—signal a growing sense of disquiet that traditional programmatic grantmaking alone is not equal to the task of fixing the stubborn problems many funders seek to address. In light of this, foundations and nonprofits are experimenting with a variety of approaches such as systems building, policy advocacy, cross-sector collaboration, movement and network building, and collective impact to create, test, and/or scale promising social innovations.

A contented and stable world might have little need for innovation. Innovation becomes an imperative when problems are getting worse, when systems are not working, or when institutions reflect past rather than present problems. (Mulgan, 2007, p 9)

A parallel trend intended to improve grantmakers' effectiveness is the sector's attention to strategic philanthropy. Undoubtedly, the core principles of strategic philanthropy (e.g., clearly articulated goals, a plausible theory of change, well-aligned partners and grantees, attention to performance metrics, and evaluation to measure progress against desired outcomes) have led to sector-wide improvement in performance. However, we are witnessing that **the practical implementation of these principles can often work against social innovation**, resulting in "calcified" social change strategies as innovators become beholden to plans and metrics that don't evolve in response to the dynamic context. This can have the unintended consequence of acting as a drag on, if not completely snuffing out, innovation.

This tension between the practices of strategic philanthropy and social innovation does *not* mean that grantmakers who are creating, testing, and/or scaling social innovations should abandon the principles of strategic philanthropy.

On the contrary, these principles can guide when and how funders decide whether collaborative systems- and network-oriented strategies are the best bets to trigger the change they seek. If kept alive and adaptable, theories of change can help innovators surface and test their core assumptions, as well as help to align multiple players around a shared vision. Performance metrics, when interpreted with careful attention to context, can keep everyone focused on the work's progress. Nonetheless, **innovative grantmakers must ask themselves whether and how philanthropic practices temper or nurture the spirit of trial, error, and adaptation that is at the core of innovation.**

In addition to flexible long-term funding and better collaboration, we believe that **one of the most transformative changes that grantmakers can make in support of innovation is to think about and use evaluation differently.** Although the best thinking on strategic philanthropy has always promoted evaluation as an indispensable tool for learning and mid-course correction in any type of social change effort, many funders still struggle to find an approach to evaluation that is *specifically* designed for social innovation. Rather than applying an evaluation approach designed to measure program impact to their social innovation strategies, grantmakers need to broaden their portfolio of evaluations to include an approach that is tailored to the unique and complex characteristics of innovation.

Given the pace of change today, funders will need to get smarter more quickly, incorporating the best available data and knowledge about what is working and regularly adjusting what they do to add value amidst the dynamic circumstances we all face. (Fulton, Kasper, and Kibbe, 2010, p. 3)

What kind of evaluation can support adaptation and leave space for the unexpected? How can evaluation be designed to give innovators the information and data they need to discover new patterns and pathways, to rapidly test solutions and abandon the ones that fail, and to detect what's emerging in response to their efforts? Most importantly, what do foundations and evaluators have to learn and *unlearn* about how they design, use, and even think about evaluation in order to create the space for successful innovation? In this paper, we advocate that traditional formative and summative evaluation approaches are no match for new and innovative programs and initiatives that experiment with solutions to complex social problems. Instead, we believe that the future of evaluating social innovation lies in the use of developmental evaluation (DE). Through insights garnered from more than 70 articles and papers, interviews with 19 funders and evaluators who are testing a new approach to evaluation that supports social innovation, and three brief case studies of DE in action, we explore what it takes for foundations to put this kind of evaluation into practice.

What Does Social Innovation Really Look Like?

“**I**nnovation” has long been a buzzword in the philanthropic sector, applied to a wide variety of approaches and initiatives that sometimes fit one of the many definitions of innovation and sometimes do not. Our interviews demonstrated that funders do not always operate with a similar understanding of what characterizes an innovative social change approach. For our purposes, a social innovation is “*a novel solution to a social problem that is more effective, efficient, sustainable, or just than present solutions and for which the value created accrues primarily to society as a whole rather than private individuals.*”¹ Innovation can take the form of new programs, products, laws, institutions, ideas, relationships or patterns of interaction, and it is often a mix of many of these. But perhaps more importantly, the term also describes the *process* of generating, testing, and adapting these novel solutions, which is inherently exploratory and uncertain.

Compared to more traditional programmatic interventions, social innovation strategies often cross sectors, involve changing the dynamics, roles, and relationships between many players, and challenge conventional wisdom

¹ Center for Social Innovation, Stanford Graduate School of Business. For discussions on the various definitions of social innovations, see <http://csi.gsb.stanford.edu/social-innovation>.

about the nature of the problem and its solutions. For example, the development and proliferation of “mhealth” (the use of mobile communication technologies to support the delivery of health care in hard-to-reach places) involved dislodging some long-held assumptions about how health care must be delivered and the kind of physical infrastructure it requires. Another example is the J.W. McConnell Family Foundation’s YouthScape program, which tested strategies for involving excluded youth in the community development process, in part by encouraging traditionally adult-focused organizations to examine and re-design their values, structures, and processes.

Over time, some innovations may become well-formed stable programs that can be tested and replicated in new settings. Others trigger a shift in existing resources and players that are unique to the place and time but are not intended to take the shape of a distinct program or service. However, what these efforts all have in common is that they seek to address problems that are both complicated (with many moving parts) and complex (including an interdependence of variables, multiple factors interacting at once, iterative and nonlinear feedback loops, and rapid change in dynamic contexts). Moreover, the pathways to results and sometimes even the results themselves are unpredictable and emergent. When problems—or the systems from which those problems emerge—are complex, they behave in ways that make more straightforward programmatic solutions less effective. For example:

- When dealing with complex problems and solutions, the past does not necessarily predict the future. Although we can make educated guesses about what might happen, the number of different factors and influences at play in complex problems makes it less likely that repeating a set of steps will produce the same results that they did the first time.
- Small changes can create large and sometimes unanticipated effects. Because all of the interrelationships between parts and players in a system are difficult to untangle, it is impossible to know for sure how—or whether—one change will “ripple” through to other players or change overall dynamics.
- When many different independent individuals, organizations, and institutions affect a problem and its solution, it can be difficult to produce specific outcomes at a pre-determined time (e.g., by the end of a particular grant cycle when grantees or foundation staff are often expected to report results). Nor at the outset can innovators predict all possible outcomes that might occur (as grantees or staff are often asked to do in a grant application or funding proposal). Innovators simply do not have enough control over the whole scope of factors or players to orchestrate outcomes in the same way that an implementer of a well-tested, stable program intervention often can.

In short, people who test new solutions to complex problems do not have the luxury of a clear or proven path for achieving their vision. They may know generally where they want to end up, but they may not know the most efficient or effective way to get there, nor do they know exactly how long it will take to arrive. Those who are interested and willing to experiment with social innovations must be willing to take risks and accept missteps or failure. They must be willing to live with uncertainty and acknowledge that their plans, regardless of how well laid out, will likely shift as the circumstances around them change.

With uncertainty and unpredictability comes an even greater need for **strategic learning** as an innovation is conceptualized, designed, and implemented. As defined by the Center for Evaluation Innovation, strategic learning is the:

Use of data and insights from a variety of information-gathering approaches—including evaluation—to inform decision making about strategy. Strategic learning occurs when organizations or groups integrate data and evaluative thinking into their work and then adapt their strategies in response to what they learn. Strategic learning makes intelligence gathering and evaluation a part of a strategy’s development and implementation—embedding it so that it influences the process. (Coffman and Beer, 2011, p. 1)

A strategic learning approach to decision-making and action acknowledges the reciprocal relationship between strategy and evaluation: that what gets evaluated should be related to and informed by the organization’s strategies, and that evaluation should feed into the development and refinement of strategy. When conceived of and implemented in this way, the organization is better able to continuously learn, grow, adapt, and change in meaningful and effective ways.



Although this learning loop is important for all kinds of social change efforts, **it is an indispensable part of social innovation.** Without it, decision makers and implementers lack crucial information about what patterns and pathways are emerging that require adaptation. Because innovation happens amidst uncertainty, decision makers and implementers often feel their way forward, testing an approach, reflecting on what seems to be happening, abandoning what doesn’t seem to work, and focusing in on what seems to be taking hold. But few evaluation approaches are well-suited to support this kind of trial and error, and in fact, many work against it.

Why Don’t Traditional Approaches to Evaluation Work For Social Innovation?

Historically, two types of evaluation have been used to understand the processes, effects, influences, and impacts of programs and initiatives. *Formative* (process or implementation) evaluations typically focus on details about how a program model takes shape; their purpose is to improve, refine, and standardize the program. Formative evaluation typically assumes that a program is or will soon become a model with a set of key activities that, if implemented correctly and with high quality, will produce a predictable chain of outcomes. The same assumption of a stable program model underlies *summative* evaluations that strive to answer questions such as, “Did the program work?” or “Should the program be continued or expanded?”

Illustrated simply, consider a common practice at the start of an initiative: foundation staff, grantees, and/or evaluators are commonly asked to provide a logic model or theory of change that illustrates the set of activities and inputs that will plausibly lead to short-, intermediate-, and long-term outcomes. Formative evaluation would examine whether the planned activities in that model are well implemented, how they affect target populations, and how they should be revised to better achieve the short- and intermediate-term outcomes. When program

Assumptions and Principles of FORMATIVE and SUMMATIVE Evaluation
<ul style="list-style-type: none"> • The focus is primarily on model testing, with a clearly hypothesized chain of cause and effect. • It is important to measure success against predetermined goals. • The evaluator should be positioned as an external, independent, and objective observer. • Evaluations should be based on predictive logic models. • Evaluations should follow a fixed plan. • Evaluation’s purpose is to refine the program or model and then render definitive judgments of success or failure.

implementers are satisfied that the core set of activities is clearly defined and could be repeated faithfully, the program is ready for summative evaluation. A summative evaluation would then test whether that well-defined set of program activities produced the desired outcomes, make judgments about the program's effectiveness or efficiency, compare it to other programs or activities that might produce the same outcomes, and/or make conclusions about whether the model can be repeated elsewhere with the same effects.

However, as described above, **social innovation is a fundamentally different approach to change than implementing program models with a known set of elements or “ingredients.”** While the long-term goals of a social innovation might be well defined, the path to achieving them is less clear—little is known about what will work, where, under what conditions, how, and with whom. Instead, decision makers need to explore what activities will trigger change; and activities that successfully trigger a desired change may never work again. Further, once one change occurs, decision makers often need to take stock of the context before they decide which activities to try next. Formative and summative evaluation designs are typically not structured in a way that gives decision makers timely information or data that supports new developments where next steps are unknown. As Tim Brodhead, past-President and CEO of the J.W. McConnell Family Foundation explains,

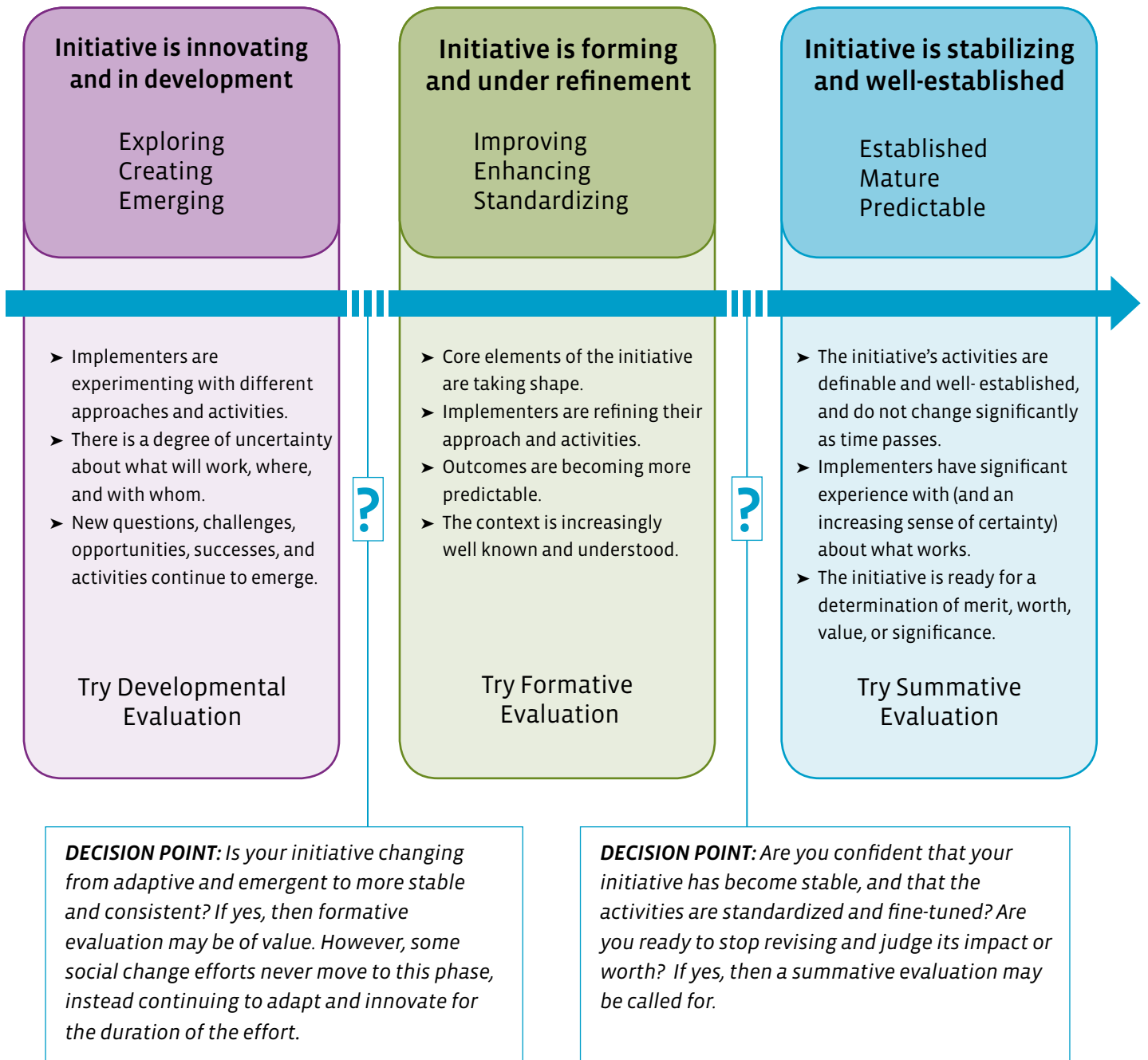
Of course, formative and summative evaluations have an important role. But many of the McConnell foundation-funded initiatives work in uncertain territory, developing and testing their strategies as they proceed; there are no blueprints for empowering youth, attacking poverty, or promoting innovative approaches to solve entrenched social problems. What is most useful for such efforts is not an ex post facto assessment of success or failure, but constant feedback from a critical, supportive observer. (Dozois, Langois, & Blanchet-Cohen, 2010)

Foundations who apply formative and summative evaluation approaches to their investments in innovation are not only missing an opportunity to obtain actionable data that increases their chance of success; they are also working at cross-purposes with their own social change investments. **When a formative or summative evaluation approach is applied to an innovation that is still unfolding, it can squelch the adaptation and creativity that is integral to success.** Decision makers and innovators (often grantees) who are evaluated on how well a set of planned activities is implemented—or whether those planned activities produce the predicted outcomes—have a strong incentive to stick to those plans no matter how the environment around them or the interests of other stakeholders change. Exploration and experimentation, and perhaps even the ability to envision alternative paths, are shut down. Finally, misinterpretation of evaluation findings as conclusive judgments of an initiative's impact when that initiative is still in a stage of development and exploration can cause funders to prematurely abandon complex efforts that may have the most promise for transformational change over the long haul.

One example in which a formative evaluation of a social innovation actually did a disservice to its constituents occurred at The Saturn School of Tomorrow in St. Paul, Minnesota, in the early 1990s. The school was considered an experimental approach for educating middle school students and was described as “high tech, high teach, and high touch.” It included a number of innovations for its time, including the following.

- There was one computer for every two students.
- The curriculum changed every nine weeks based on students' needs and interests.
- The primary instructional strategies were collaborative and project-based learning projects.
- The school used a differentiated staffing model that included four lead teachers, a number of regular teachers, and several paraprofessionals.
- Students took multi-grade classes.
- Learning in the community was supported by the school's location in a renovated downtown building.

What evaluation approach is the best fit for each life stage of an initiative?



The formative evaluation was designed and implemented over a three-year period, during which the evaluator spent about eight hours onsite every week observing, interviewing, reviewing documents, and conducting focus groups. It didn't take long, however, for the evaluator to realize that a formative evaluation was insufficient for this experimental, complex undertaking. While she could anticipate and even predict some of the outcomes the school was hoping to achieve, the real-time design and implementation of those strategies meant that staff members' needs for information on what was or was not working changed almost daily. As the *formative* evaluator, she was limited to gathering and analyzing data and reporting it back to the stakeholders through year-end reports. Just a few months into the school's first year, however, the evaluator could tell that things were not going well: staff weren't getting along, discipline problems were rampant, few school policies were in places, there were accounts of racism among the teaching staff,

and a local newspaper published multiple negative stories about the school. These issues contributed to parents' loss of faith in the school and many parents withdrew their children.

When the evaluator expressed interest in sharing key insights from her ongoing data collection efforts with the teaching staff as they emerged, a process that often falls outside the formative evaluator's role, the teachers insisted that she remain the "objective observer and documenter" of the school's development and progress. The formative evaluation approach with this educational social innovation had several negative, unanticipated outcomes: 1) the lessons generated by the evaluation were delivered too late to be useful, and 2) the findings that were shared publicly in the year-end reports were used prematurely by parents, the school board, and other community members who concluded early on that the school "was not working." This experience convinced the evaluator that there must be another way to evaluate social innovations—innovative and real-time initiatives for which there is no accepted model and/or that are in development.

What Kind of Evaluation Supports Learning and Adaptation for Social Innovation?

So what is the alternative? How can foundations and nonprofits remain committed to results and get the data they need without bringing innovation to a halt? An emerging approach called **Developmental Evaluation (DE)** is gaining traction among some funders who support collaborative, complex, evolving change processes. Originally conceptualized and described by evaluator Michael Quinn Patton, DE has an evolving definition:

*Developmental evaluation informs and supports innovative and adaptive **development** in complex dynamic environments. DE brings to innovation and adaptation the processes of asking evaluative questions, applying evaluation logic, and gathering and reporting evaluative data to support project, program, product, and/or organizational **development** with timely feedback. (Patton, 2011a)*

The DE evaluator works collaboratively with social innovators to conceptualize, design, and test new approaches in a long-term, on-going process of adaptation, intentional change, and development. (Patton, 2011b)

Data collected as a part of DE should be analyzed and interpreted *with* key stakeholders in a timely way to detect how the context is changing, affirm current practices or inform new activities, and guide programmatic and strategic questions regarding the innovation's progress and likelihood of success.

Developmental evaluation has five characteristics that distinguish it from other evaluation approaches. These include the focus of the evaluation, the intentionality of learning *throughout* the evaluation, the emergent and responsive nature of the evaluation design, the role and position of the evaluator, and the emphasis on using a systems lens for collecting and analyzing data, as well as for generating insights.

1. *DE's focus is on social innovations where there is no accepted model (and might never be) for solving the problem.* As a result, the evaluation explores how, why, and with what effects the project is designed and implemented, and how

TYPES OF QUESTIONS ANSWERED BY DE

- What is developing or emerging as the innovation takes shape?
- What variations in effects are we seeing?
- What do the initial results reveal about expected progress?
- What seems to be working and not working?
- What elements merit more attention or changes?
- How is the larger system or environment responding to the innovation?
- How should the innovation be adapted in response to changing circumstances?
- How can the project adapt to the context in ways that are within the project's control?

it is evolving, adapting, and responding to internal and external conditions.

In the end, we want to get to outcomes, but we can't get there if we don't have a clear theory of change. We first think and develop an idea of what could work, then we develop a model. You test it to see if it works in different setting, tweaking it as you go based on feedback, and then you scale it. When the objective of the project is to develop that model, this is when you can use DE. —John Cawley, The JW McConnell Family Foundation

2. Continuous Learning is intentionally embedded into the developmental evaluation process. It is a means of ensuring that all stakeholders not only take in information as the initiative unfolds, but also integrate the evaluation findings into their thinking and decision making processes, thus using their learning to refine, adapt, and change the innovation as needed. This involves providing the space, time, and resources for dialogue, reflection, questions, identifying and challenging values, beliefs and assumptions, and instituting feedback loops. It also involves the evaluation team's involvement in ongoing sense-making activities: interpreting, synthesizing, and generating insights and recommendations using multiple forms of written and verbal communications.

I would be open about what you can learn (as opposed to knowing what you want to learn). Is the desired outcome about learning a specific thing, or are the desired outcomes being able to learn any number of things that you don't even know you're going to learn until you get there? —Susan Patterson, Knight Foundation

3. An emergent and adaptive evaluation design ensures that the evaluation has purpose and that it can respond in nimble ways to emerging issues and questions. While there will often be a set of negotiated and planned evaluation questions and data collection activities at the beginning of a developmental evaluation, this plan will likely evolve and change as the innovation is designed and implemented. This might involve adding or eliminating certain data collection activities, expanding or adjusting the pool of data sources, going deeper into or rethinking some of the key evaluation questions, and/or developing new ways for communicating and reporting evaluation findings for various internal and external stakeholders.

There is a connotation of exploration—the undertaking is complex enough and so contextually driven that we have to continue to have mechanisms to take a good look to see what's changing and what factors are at play, and to continuously modify our methods and measures to make sure we're attuned to what's happening on the ground and in the system. —Meg Long, OMG Center for Collaborative Learning

4. The role of the developmental evaluator is a strategic learning partner and facilitator, which reflects a different role for most evaluators and their clients. Developmental evaluation requires that evaluators not only have basic professional evaluator competencies, but also experiences and strengths in:

- Facilitating group learning processes
- Identifying themes, synthesizing information, and generating insights in real-time
- Listening for what is being said as well as what is not being said
- Observing interpersonal and power dynamics
- Interpreting conversations and situations from different angles
- Zooming in to see the details of what is happening and zooming out to see big picture trends
- Building strong client relationships based on good rapport and trust
- Asking the right questions at the right time
- Facilitating courageous conversations and providing constructive feedback
- Being comfortable with ambiguity and change
- Using multiple forms of communications
- Being tactful and diplomatic

CASE EXAMPLE: Bill & Melinda Gates Foundation's Community Partnerships

Provided by Meg Long, OMG Center

Context: Traditionally, education reform stakeholders have worked in silos to improve postsecondary outcomes from their own vantage point, resulting in a fragmented system that helps some, but fails many—particularly low-income young adults in underserved communities. The Bill & Melinda Gates Foundation's Community Partnerships portfolio posits that increasing postsecondary completion rates for low income young adults requires the activation and coordination of a number of diverse stakeholders, including the K-12 educational system, higher education, the business community, political, civic, and community leaders, and social service providers. The goal of the portfolio's evaluation is to understand what it really takes for a community to coalesce around a postsecondary completion goal. How does a community build a multi-tiered partnership among individuals and organizations that may never have worked together? And then, how do you get partners to change the way they do business to more effectively and efficiently support college success for low-income young adults? Answering these questions will help the Foundation structure its innovations, policy, and practice changes to increase the likelihood of successful implementation on the ground.

Evaluation Approach: While a Theory of Change articulated expectations within the Community Partnerships sites, that theory was very underdeveloped. There were four "buckets of work": communities would develop partnerships, they would use data to inform their strategy, they would seek to build commitment among stakeholders, and they would tackle policy and practice change. That was essentially the extent of the "theory" behind this investment. As one of the program officers noted, you could drive an eighteen-wheeler through the holes in this theory of change, indicating just how large the gaps were in our knowledge. There was nothing to test yet; it was entirely up to the selected communities, armed with deep knowledge about their local context, to make sense of these four "buckets" and to shape the work as they saw fit. A developmental evaluation approach allowed the OMG Center evaluation team to make sense of the process alongside the grantees and the technical assistance providers (The National League of Cities and MDC). This helped a) define what the work actually looks like, b) explore how it differs under different conditions, and c) understand the factors that facilitated or impeded the sites' progress. These factors will be critical for the Foundation to consider when pursuing future strategies that require local implementation.

Evaluation Methods: Developmental evaluation requires near-constant contact. The evaluation team connects directly with the grantees and their partners through interviews and site visits every three to four months. We speak with the technical assistance providers and the foundation program officer every two weeks and review documents and data from sites on a rolling basis. In most cases, we have access to document sharing websites that grantees have set up to support the work of their partnership, becoming defacto partners. We structure interviews to build off of previous conversations; this technique requires significant preparation from the evaluation team but results in a running narrative that documents in detail how the work is unfolding. Following every major data collection point, we share a rapid feedback memo with the site, the technical assistants, and the foundation team containing our observations and questions for consideration. We share a product nearly every eight weeks, and we pair most products with a debriefing call or a reflection meeting. We also have an annual Theory of Change refresh meeting that allows the evaluation, foundation, and technical assistance partners to fill in the Theory of Change as we learn more.

Key Learnings: The Community Partnerships portfolio was conceptualized as a systems change effort to realign stakeholder policies and practices with the ultimate goal of increased completion. The DE has allowed us to capture and synthesize an unprecedented level of nuance about how change happens in a given community—who needs to drive the agenda, who needs to support it, how they can get onboard, and what structures are needed to support the effort. DE has also helped us unearth the habitual and cultural practices and beliefs, as well as the informal systems, that exert enormous influence on how formal systemic players (e.g., school districts, higher education institutions, municipal leaders) operate. These informal systems could have been easily overlooked in a more traditional formative evaluation with a more structured framework of analysis. Such a result would have been very unfortunate, since a key lesson from the Community Partnerships portfolio is the importance of being aware of and managing both formal and informal systems in order to truly see an uptake of the postsecondary completion agenda.

If one doesn't understand the dynamics of developmental situations, it's very hard to communicate and operationalize developmental evaluation. It rests on embracing the uncertain, ambiguous, and emergent nature of the world and how we try to change it not only intellectually, but emotionally. In my experience, most of the challenges in DE aren't really about the techniques of DE...they are challenges with working adaptively in a messy world.—Mark Cabaj, Here to There, and Tamarack: An Institute for Community Engagement

5. *The developmental evaluator brings a **complex systems orientation to the evaluation**.* The evaluation team is attuned to the complexity of social innovation and recognizes the interdependencies of variables, the importance of feedback loops, the location and amount of energy within the system, and the existing and developing boundaries, relationships, and power dynamics. The DE evaluator is also committed to identifying both expected and unanticipated outcomes of the initiative. A systems orientation acknowledges that change does not occur along a linear pathway and that even the best logic models and plans cannot fully predict the effects an innovation may have. The developmental evaluator is less interested in *isolating* the individual factors and effects of the innovation than she might be in a summative evaluation. Instead, she is particularly interested in the ways in which the various program elements, components, and actors interact with one another, as well as understanding how particular activities and interactions trigger changes in other dynamic parts of the system.

You don't just look at the target of your gaze. When you have a systems thinking lens, you also step back and look at the environment, and at how things might have happened. Borrowing a phrase from the TV series "The Wire," with "soft eyes" you step back to look at the situation in different ways. What are the boundaries here, what are the inter-relationships, and what are the different perspectives at play?—Margaret Hargreaves, Mathematica

While some formative and summative evaluations may also be characterized by a few of these principles, DE always includes all five. Without them, the evaluation cannot close the feedback loops that allow decision makers and implementers to adapt as the initiative unfolds.

What Kinds of Grantmaking Strategies Are A Good Fit for Developmental Evaluation?

Funders should consider using developmental evaluation rather than formative or summative evaluation (or worse yet, no evaluation at all) for grantmaking strategies that embody the characteristics of innovation: initiatives **that attempt to address or solve a complex problem with an adaptive solution**. Such efforts may have clear goals and a vision of a long term outcome, but the path to achieving these outcomes is not linear and the outcomes are not well known, agreed upon, or clear. In reality, the work is a bit of a mystery...

*Malcolm Gladwell, in "What the Dog Saw," has a great story where he makes a distinction between a puzzle and mystery. It's a nice metaphor for talking about the difference between simple, complicated, and complex. In a puzzle, if you do your homework and get more data, you will solve the puzzle. More data is better. In a mystery, it's not so much data as **sense-making** that's really critical. A lot of DE situations aren't puzzles, they're mysteries—like how to solve homelessness. There are multiple perspectives, there's no one truth, data are shifting. The burden of proof may or may not be high, but the data even at its best are somewhat ambiguous, so you really need great sense-making skills. —Mark Cabaj, Here to There, and Tamarack: An Institute for Community Engagement*

There are several common grantmaking strategies that are better suited for developmental evaluation than formative or summative evaluation (Patton, 2010):

For initiatives that develop in an ongoing way and will never settle into a fixed model, such as *advocacy and policy efforts* or *comprehensive community initiatives*.

The way these kinds of initiatives unfold depends deeply on the context within which they occur, and many independent factors and actors affect progress. As a result, they will not stabilize into a fixed model and must continually adapt to the changing environment. Developmental evaluation in these contexts provides an ongoing stream of information and insights that help designers, implementers, and funders know how relationships and structures are developing, what is working well, what challenges are emerging, what unanticipated consequences are surfacing, and what level of progress emerges on desired outcomes.

When exploring the creation of a new model, the basic ingredients of which are not yet known.

Even when the eventual goal is the testing, replication, or scaling of a proven model, DE is the first step in understanding what the right core ingredients might be. Similar to beta testing, the developmental evaluation would provide useful information for developing the intervention before it is fully rolled out, refined, or expanded, thus increasing confidence in its potential effects, influence, and impact.

When replicating a program in a new site, context, or population that requires adaptation. A developmental evaluation for these purposes would pay particular attention to how the guiding principles of a program are being incorporated into work practices, the systems and structures that support or impede implementation, the relationships that are developed or challenged, the ways in which the principles are adapted to fit into new cultural and social contexts, and the unexpected as well as anticipated outcomes.

For systems change initiatives that seek to trigger change across multiple systems or in a particularly complex system. When a social change effort requires major disruptions to a current system, developmental evaluation focuses on the phenomenon known as the “butterfly effect,” where seemingly small events can lead to more significant changes to the larger system. DE evaluators track and collect information on emergent patterns, relationships, feedback loops, and energy as the initiative is designed and implemented.

Understanding how the various systems interact, adapt, and change in response to the environment is critical to supporting wider scale implementation in other contexts and locations.

When decision makers need to rapidly adapt an initiative in response to a sudden crisis or change. Where the organization is interested in responding to and evaluating the effects of an unforeseen change or crisis, there

Questions for Determining if DE Fits Your Initiative

- What is the nature of the problem we’re attempting to solve? Is it truly a complex problem?
- What is the system we are trying to affect and how complex is it? (E.g., are there many organizations, actors, activities in this system?)
- To what extent is our intervention a complex or adaptive solution to this problem (e.g., exploratory and flexible, and dependent on the moves or actions of other players)?
- Is our intervention based on a model that is already developed? If yes, do we know what sequence of activities are expected to happen?
- To what extent can we predict most of the short, interim, and long-term outcomes of our intervention? If we cannot predict these, why not?
- Do we need data and feedback as we work to be able to decide next steps?

I’ve used DE inside traditional evaluations where there is a model—for example, in home visiting models, which are very specific. Evaluations of these programs can be very traditional. There is a focus on the fidelity of implementation and on the achievement of particular outcomes. However, there can be a developmental side to those evaluations as well, looking at how those grantees are building infrastructure to support those particular implementation models. We want to know, “What kinds of infrastructure are needed to implement, sustain, or scale up those models?” The answers to those questions are much less well known. —Margaret Hargreaves, *Mathematica*

is usually no time to develop a model for action. Given new circumstances and uncertainty about what to do, decision makers must develop approaches in real time and need data quickly. Because developmental evaluation by definition is adaptive and responsive to arising information needs, DE evaluators are able to nimbly and quickly hone in on what information is needed by whom, when, and in what formats.

A social change effort might include a combination of the innovative components above and components that look more like a stable program that is ready for formative or summative evaluation. In these cases, some aspects of the initiative might be best served through a developmental evaluation, while other components could be evaluated using a formative or summative approach.

Deciding whether developmental evaluation is the right approach for a social change initiative requires a candid assessment of the extent to which the initiative is truly *in development*. Even in cases where an initiative's *concept* is innovative (i.e., the problem is complex and there is no model for addressing it), the culture, structure, and processes of the funder(s) and implementers may not *allow for development*. In other words, developmental evaluation works best for the kinds of initiatives described above *as well as* those in which foundations and implementers are open to and flexible enough for adaptation. While this might sound simple enough (what organization doesn't want to think of itself as open and flexible?), we found that some of the most common practices in the fields of evaluation and philanthropy—from foundation contracting and evaluator business models to the traditional relationship between the evaluator and program implementers—are fundamentally designed around the assumption that adaptation and exploration will be kept to a minimum. For developmental evaluation to work well, foundations and evaluators both need to proactively examine how their processes and habits of interaction can support the adaptation that is necessary for social innovation.

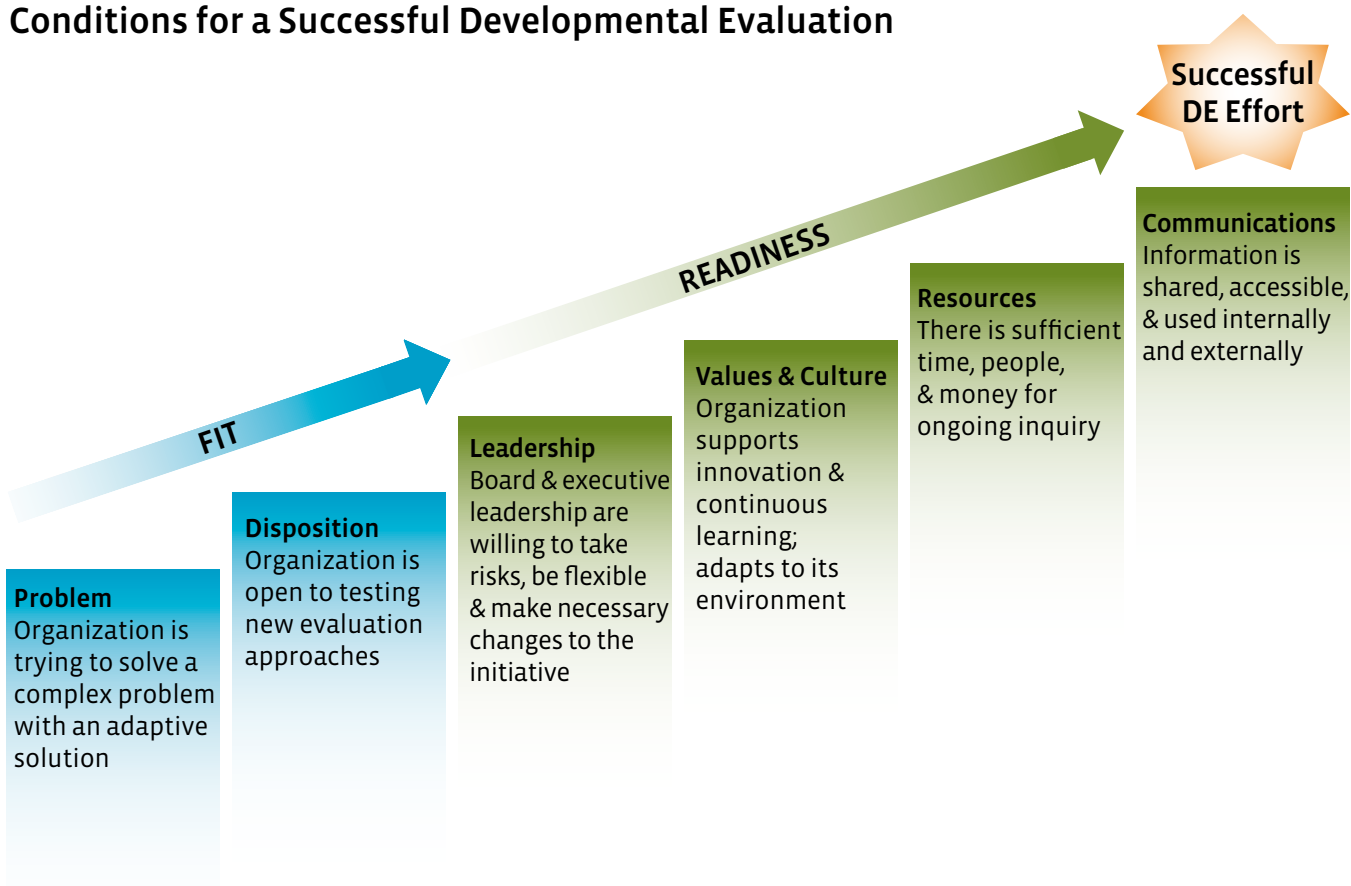
You can give the best feedback and evaluative thinking in the world, but it doesn't matter if evaluation users are not really interested in or capable of innovation or adaptation. Alternatively, even if you have all the other conditions in place—evaluation users with the true ability or authority to adapt the intervention, open timelines, engaged funders, and so on—they must be ready, willing, and able to submit their experiment and their work to evaluative thinking and rigorous sense-making.
—Mark Cabaj, *Here to There, and Tamarack: An Institute for Community Engagement*

How Can Funders Know If They Are Ready for Developmental Evaluation?

Developmental evaluation requires a genuine and conscious commitment to learning and change. It is grounded in a humbleness and ability to say, “We don't quite know what's going to happen here, but we believe that our intervention (initiative, project) can contribute to making a difference in solving this complex social problem. And, because we can't predict all possible outcomes or how this effort will affect the people and systems with whom and within which we're operating, we need an evaluation approach that provides us with real time questions and data so we can learn and adjust our strategy and activities along the way.”

Successful developmental evaluations require a set of conditions that ensure that DE is the right fit for what is being evaluated. In addition to first determining whether the problem being addressed and its solution fit the criteria described in the section above, the foundation, nonprofit, or team of program implementers must be ready to effectively engage in and support developmental evaluation.

Conditions for a Successful Developmental Evaluation



Adapted from *DE 201: A practitioner's guide to developmental evaluation* by Dozois, Langlois, and Blanchet-Cohen

CONDITION 1: The organization's leadership is willing to take risks, be flexible, and make necessary changes to the initiative.

Developmental evaluation often involves testing the innovation's ability to achieve a particular strategy. Thus, it is imperative that leaders not only exhibit a high level of willingness and ability to engage in the evaluation whenever necessary, but also that they use evaluation findings to make changes in the initiative's implementation and/or strategy. Leaders need to a) communicate, in multiple ways, the importance of the developmental evaluation; b) be willing to provide adequate time, personnel, and financial resources; and, perhaps most of all, c) be committed to using evaluation information (learnings), to make course corrections, improvements, and changes to the initiative and its related strategy.

To determine whether leadership is ready for DE, consider:

- To what extent does leadership value and use evaluation findings for programmatic and strategic decision making?
- To what extent are leaders comfortable with making changes in the initiative or the strategy as new learnings emerge?
- To what extent are leaders able to balance emergent and ongoing evaluation findings typical of developmental evaluation with the accountability demands made by the Board and external stakeholders?

DE needs to be seen as something that's tied closely to strategy. It needs to have strong ownership from the President or chief program officer, and has to be part of a program officer's core responsibilities, rather than something that's simply delegated to consultants. The team has to be involved in the design of the evaluation and think carefully through the questions.... It has to be part and parcel of their job. —Mayur Patel, Knight Foundation

CONDITION 2: The organization's values and culture support innovation, continuous learning, and adapting to its environment.

Developmental evaluation works best when an organization has an authentic commitment to learning and change. A learning organization culture supports risk taking, information sharing, trust and respect for individual and cultural differences, collaboration, asking questions, embracing and learning from failure or mistakes, and using data for decision making. Developmental evaluation can play a key role in supporting individual, group, and organizational learning under the right conditions.

The following questions may be helpful in diagnosing an organization's culture:

- To what extent is the organization a learning organization? Where and how does it learn best?
- To what extent is there a culture of risk taking within the organization? What supports and/or inhibits staff and leaders from taking risks?
- What is the level of trust within and among the staff and leadership? What supports and/or inhibits this level of trust?
- How well does the organization adapt to change and unanticipated events or consequences?

They need to have the actual desire to learn. Everyone says they want to, but that's not always the case. Learning is change, and people are often afraid of change. —Elizabeth Dozois, *Principal, Word on the Street Ltd.*

People should not shy away from failure. The developmental evaluator would suggest corrections so we could see what's going wrong. There is no point in doing something new if you are not open to the possibility that it might not work; DE can help you learn if it's working or not. —Michelle Clarke, *Burns Memorial Fund*

CONDITION 3: The organization provides sufficient time, people, and financial resources for ongoing inquiry.

Evaluation of any kind is an investment in the organization's ability to make sound operational, programmatic, and strategic decisions. Valuing evaluation means providing the necessary time, personnel, and financial resources to support evaluation activities and the use of findings. While evaluation costs vary greatly depending on how much information is needed and from how many sources, the depth of information required, and the time required to collect, analyze, and synthesize the data and findings, developmental evaluation *may* require more resources than some other types of evaluation. This is usually because evaluators spend more time in strategy discussions and facilitating the application of findings to decisions. The organization needs to provide sufficient resources so that the DE can be implemented in ways that provide the most meaningful and useful information.

I think you can have a budget and a plan of what you expect to happen—when the reporting is happening, the number of meetings needed—but you have to be flexible as time goes on because things might change. For example, you might change the reporting, the methodology, or how much facilitation you need. —Michelle Clarke, *Burns Memorial Fund*

When considering whether the organization is willing to invest the necessary resources in developmental evaluation, ask:

- What is the organization's history in resourcing evaluations?
- To what extent is the organization prepared to allocate sufficient time, personnel, and financial resources to a developmental evaluation?
- How flexible would the organization be to changing allocations of time, personnel, and funds during the course of the developmental evaluation?

CASE EXAMPLE: The YouthScape Initiative

Provided by John Cawley, The J.W. McConnell Family Foundation

Context: YouthScape was designed as part of the Foundation's inclusion strategy and focused on young people who feel disconnected from school and other mainstream organizations led by adults. The objective was to test the assumption that engaging young people in local decision-making develops and draws upon their knowledge, skills, and attitudes. In doing so, the program serves to enrich democracy in our communities. The YouthScape initiative brought together partners from different sectors at multiple sites across Canada, many of whom had different perspectives on both the rationale and strategies for engaging marginalized youth. The design and implementation of the initiative were inspired by and contributed to the development of local youth leaders, yet the real challenge was to get youth-friendly values and practices "into the water supply" of social service agencies and municipal governments. Most youth engagement experiments are either well-meaning, adult-led youth programs or youth-led projects kept at arm's length from the core business of community organizations. YouthScape challenged mainstream partner organizations to embrace the centrality of youth participation and decision-making in their governance and operations while delivering community outcomes.

Evaluation: Approach: As the Foundation shifted its funding to complex, long-term initiatives that are not so much pre-planned as emergent, the inadequacy of conventional evaluation methods became evident. We need to know, in real time, how a cluster of complementary prototypes are doing, not whether we have arrived at a pre-determined spot, on budget, and at a specified time. Developmental evaluation overturns many of the assumptions of more traditional approaches; it is embedded rather than detached,

continuous rather than episodic, and—most importantly—its goal is learning, not accountability. It provides useful information to the do-er rather than cover for the giver.

Evaluation Methods: Each of the five YouthScape communities had a part-time developmental evaluator who participated in key events, gathering feedback before and after from a range of stakeholders: young people, staff, managers, volunteers, and community partners. The data was fed back to community organizations immediately. On a monthly basis, the five community-based developmental evaluators held a conference call under the mentorship of the national developmental evaluator. This provided an opportunity to aggregate data nationally and for coaching. The national developmental evaluator formally communicated key national trends (emerging opportunities, points of tension) to the managing agency and to the funder on a quarterly basis, and more often informally.

Key Learnings: During YouthScape's first 18 months, the developmental evaluators were able to illuminate tensions and misunderstandings that, had they gone unattended, might have undermined the entire initiative. Community organizations, the national managing agency, and the Foundation learned of design flaws, subterranean grumbling, hidden strengths, and promising breakthroughs as a result of the developmental evaluators' ability to provide real-time feedback to the initiative implementers. Developmental evaluation allowed us to modify program designs, provide training, convene partners, and create spaces for airing concerns in ways that we could never have anticipated. In short, their contributions to the success of the initiative were significant.

CONDITION 4: The organization is committed to open communications and ensuring that information is accessible and used internally and externally.

Developmental evaluation is all about learning and using what is being learned to a) better understand contexts and situations, b) make informed decisions, and c) take action when needed. In many ways, the success of DE is grounded in the frequency, format, accessibility, and transparency of communications throughout the evaluation. To decide whether the organization's communications practices will support DE, consider the following questions.

- What internal communications systems and supports already exist within the organization, and is there a culture of information sharing within the organization?
- How comfortable is the organization with receiving evaluation information at different times and in various formats throughout an evaluation (compared with traditional interim and/or final written reports)?
- How are lessons learned currently shared within the organization? With grantees? With the field?

The success of any evaluation, including developmental evaluations, often depends on matching the right evaluation approach to the questions and information needs at hand, as well as the readiness and willingness of the foundation and its grantees to engage in and support the various evaluation activities and learnings that it produces. While we learned from our interviews that few organizations embarking on DE have *all* of these conditions in place, it is important to determine which are missing and to create a plan for managing the gaps or building the organization's readiness and commitment to sustain developmental evaluation processes.

You can't wait until the end of a multiyear project to find out how it's going, but at the same time, you can't do it every month. You have to have enough time to let something actually happen. With KCIC, we have enough staggered cohorts to get some interim information along the way and I think that's about right. It takes a while for these projects to get going. You can ask questions too soon and not really learn anything, or wait too long and it's no longer useful. —Susan Patterson, Knight Foundation

How Can Foundations Be Ready For The Tough Questions?

Because developmental evaluation differs from other evaluation approaches in several ways, including purpose, position and role of the evaluator, and emphasis on real-time learning in the midst of uncertainty, our interviewees noted that a set of tough questions inevitably arises for both foundations and evaluators using DE. Intentionally discussing these questions up front—and revisiting them throughout the evaluation—can ultimately strengthen and enhance the overall evaluation effort.

What about accountability for impact?

Accountability in philanthropy is normally defined in one of two ways: 1) did grantees do what they said they were going to do, and 2) did the grantees or the funding strategy produce the intended outcomes? However, both of these conceptualizations of accountability can be a mismatch for innovative social change strategies, particularly early in the innovation process when the pathway to outcomes, and often the outcomes themselves, are emergent and uncertain. As a result, poorly-timed demands for traditional accountability information (whether in the form of grantee reporting against the programmatic outcomes predicted in their grant proposals or through summative evaluations designed to judge a program's merit, worth, or value) can send program staff and grantees mixed messages about whether their funder truly expects them to innovate, adapt, and test new approaches that may or may not work.

Rather than holding themselves accountable at the wrong time for meeting pre-determined results, or for accurately charting a pathway to specific outcomes up front, social innovators (including funders) are acting accountably when they *pay careful attention to what is emerging as they work and adapt accordingly*. With this kind of re-framing, accountability and learning in complex social change initiatives go hand in glove. Innovators are accountable *to the learning*. Boards and leadership should revisit this framing of accountability in innovation at the start of a strategy, but also throughout its life. If and when a strategy transitions from innovation and adaptation into a more stable, predictable phase, a corresponding transition in evaluation approach will help answer more traditional accountability questions.

The biggest challenge is that when people think about evaluation, they think outcomes. We've been hard wired to automatically think outcomes when we hear the word "evaluation." So you have to set expectations on the front end.... It's important to be focused on learning and improvement, but as an initiative develops, people will invariably ask "Where are the outcomes? What happened to people's lives?" That's always going to be an impulse and actually a very good one at that. I hope the emphasis on DE would never take us away from this. —Mayur Patel, Knight Foundation

CASE EXAMPLE: The Knight Foundation's Community Information Challenge

Provided by Mayur Patel, John S. & James L. Knight Foundation

Context: The John S. and James L. Knight Foundation's five-year, \$24 million initiative—the Knight Community Information Challenge (KCIC)—was designed to respond to the rapid disruptions in journalism and the decline of local, community news and information. Through more than 76 KCIC funded projects, foundations are creating and supporting new news models, online information hubs, citizen journalism activities, youth media outlets, advocacy campaigns, and civic engagement activities. The KCIC reflects a social innovation in that it aims to catalyze efforts among community and place-based foundations to creatively use media and technology to keep their communities informed in ways they have not done in the past. The initiative represents a new approach to strengthening community information by encouraging community and place-based foundations to focus on local media and news as key elements of their existing programmatic strategies.

Evaluation Approach: The Developmental Evaluation approach was chosen because it provides: a) continuous feedback to Knight about how the initiative is being implemented by its many grantees and its progress on moving the field, b) information about how grantees are adapting and changing as a result of their projects, and c) insights and recommendations for how Knight could think about and implement any changes in their current strategy, as well as implications for the next phase of strategy development. Some of the guiding evaluation questions include:

- *Motivators/Barriers:* What internal or external factors motivate community and place-based foundations to become engaged in addressing community information needs? What internal or external barriers prevent foundations from engaging?
- *Strategic Alignment:* To what extent are foundations' strategies aligned with community information? How are foundations integrating ("overlying") information and media on top of their program-level strategies?
- *Learning:* What are grantees learning about launching, implementing, and evaluating their information projects? What lessons can grantees share with each other and the field?
- *Impact:* In what ways are KCIC projects changing the behaviors or attitudes of community members?
- *Information Ecosystems:* What community-level conditions (e.g., organizations, infrastructure, and activities) foster a healthy information ecosystem? What role might community or place-based foundations play

to provide structure and support for information needs within the community?

Evaluation Methods: The KCIC developmental evaluation uses a collaborative, utilization-focused, and organizational learning approach and includes commonly used evaluation methods such as interviews, surveys, focus groups, observation, and document review. Evaluators pay particular attention to the ways in which grantees develop their projects and affect their local news and information ecosystems. In addition, they determine if and how the initiative is "moving the field" in the adoption of news and information as core to their organizations. As new and emerging questions and issues arise, the evaluators adjust the evaluation plan in timely and responsive ways to capture information as needed by the Knight team. After each data collection activity, evaluators analyze and synthesize the data, and then share the results with Knight in a variety of formats, including email memos, PowerPoint decks, and phone working sessions. Knight's commitment to building the field has led the evaluators to reformat many of the evaluations' findings for external audiences in the form of briefs, visually appealing reports, blog posts, toolkits, and webinars.

Key Learnings: While the developmental evaluation provided ongoing learning for Knight, which enabled it to make real time changes to the initiative's implementation, the evaluation results and key insights are now helping inform the next iteration of Knight's strategy. The evaluation has provided feedback on the extent to which the current initiative has penetrated the community foundation field, as well as on possible opportunities for Knight to move beyond project-based support to larger multi-year partnerships with foundations that are highly committed to supporting local news and information. Specifically, the developmental evaluation:

- Helped Knight see opportunities to connect to local leadership
- Influenced the way Knight thought about selection processes
- Helped Knight see where foundations have focused their efforts and why.

The assessment has also shed light on the factors influencing foundations' engagement with news and information, as well as the types of support structures Knight Foundation could put in place to deepen its engagement.

Is DE rigorous?

Questions about rigor in evaluation most often focus exclusively on the evaluation’s design and methods. Yet rigor is found not only in design or methods. High quality evaluation requires rigorous evaluative and critical thinking, as well as critically examining assumptions—practices that are always integrated into developmental evaluation.

When focusing more narrowly on rigor as a feature of data collection *methods*, evaluators must adequately test data collection instruments, sample sufficient numbers, and triangulate data using multiple methods and sources. With most formative and summative evaluations, the data collection and analysis schedule is determined primarily by the evaluation design and methodological rigor required to establish certainty about the intervention’s effects. Yet, while methodological rigor provides greater confidence in the accuracy of findings, the trade-off is that data sometimes arrive too late to inform a tactical or strategic decision, or to illuminate a particular window of opportunity.

Professional evaluators spend too much time on methodological rigor and not enough time asking what people want to know that will help them do a better job. —John Cawley, J.W. McConnell Family Foundation

For evaluation data and feedback to be useful for decision making *while a strategy is unfolding*, they must be available to program staff at the right moment. Rarely do strategic decisions or emergent opportunities follow the predictability of the semi-annual reporting schedule dictated in many evaluation engagements. While important decision-making moments are sometimes planned from the outset (e.g., board meeting schedules when key funding decisions are made, scheduled program reviews, election cycles, partnership convenings), many are unexpected. For example, a key player or institution may enter or leave the scene, an unexpected economic or political crisis may occur, or a large new funding opportunity may appear, causing a significant shift in the dynamics of the system that requires a decision within weeks or even days.

In an innovative change strategy, it is impossible to anticipate all of these questions ahead of time. As a result, evaluators conducting developmental evaluations need to be nimble and responsive as they collect and analyze data quickly, while still maintaining a high level of quality. The evaluator and the program team must ask all along the way: *How much data is enough? How certain do we need to be before we take the next step? Which changes or dynamics do we need to be most certain about and which can be left to instinct or hunch? What is truly urgent for us to know and do now?* Evaluators, whose professional standards of practice call on them to be clear about the limitations of an evaluation and its findings, can include the program team in decisions about methodological rigor versus timeliness so that the group “owns” the trade-offs. And perhaps most importantly, developmental evaluators should ensure that rigor is guarded in other aspects of the evaluation, including rigorous questioning, analysis, and sensemaking.

Where is the boundary between the evaluator and program decision makers?

Evaluators are sometimes positioned as objective and somewhat distant observers, collecting and analyzing data and offering findings to others who decide whether and how to act in response. However, this role does not effectively serve developmental evaluations where evaluators must be “in the mix,” bringing observations and data to key meetings, asking tough questions, and drawing insights out of the group to increase

Trust at all levels is critical—the “do no harm” concept really applies here. Everyone has to know and trust that it’s not the intent of the evaluator to highlight the holes in the work and underscore what is not working. In developmental evaluation, you need to create space for honest conversations with the funders, the grantees, and the intermediaries to share your observations and help fine tune and improve the strategy as the work unfolds. Maintaining a rigorous methodology will help you maintain objectivity. —Meg Long, OMG Center for Collaborative Learning

understanding about *how well the strategy is working*, in real time. An evaluator who is too removed from the desires and hopes of those making strategic decisions will miss valuable conversations that can only be had through close connection to—or even integration within—the program team.

An evaluator who assumes this integrated role serves as a “critical friend,” coach, or technical assistance provider, which can sometimes generate confusion among grantees or tensions with other intermediaries or program staff. This tension is exacerbated when evaluators ask questions that may highlight either interpersonal problems or conflicts between funders and grantees. Balancing this tension, of course, requires significant trust between parties. Evaluators and program staff must build in (and budget for) sufficient time to create and manage this relationship. The program team must value and support the role of evaluator as critical friend, while the evaluator has to maintain integrity, professional evaluation standards, and a critical perspective.

Does DE work for everyone?

Many believe there are two types of individuals: those who briefly reflect on data and take action primarily based on intuition and experience, and those who mine data and explore interpretations through a process of dialogue, reflection, and questioning assumptions before taking action. For those who are more reflection-oriented, taking action on uncertain or imperfect information—a common occurrence in social innovation—can feel careless and unnecessarily risky. The developmental evaluator has to balance time spent on reflection and dialogue with time spent acting, and be prepared to deal with tensions between these two kinds of learning and decision-making styles.

The whole idea of reflective learning is frustrating to some people. A lot of people love the conversations, while others say, “Good God, do we have to do this again and are we ever going to get anywhere?” So it’s not one or the other learning style, but that’s certainly the way it gets labeled.
—Marc Langlois, *Skipping Stones Consulting*

While DE is specifically designed to provide data and findings in a timely manner, it can also cause a sense of information overload, particularly for action-oriented people who are unaccustomed to a robust feedback loop. Evaluators can balance the tension between reflection and action by helping the program team identify and prioritize evaluation questions and data collection activities that are pivotal to supporting action steps.

How do we budget for DE?

The practice of developmental evaluation is sufficiently different from most formative and summative evaluations that funders and evaluators often grapple with how to set appropriate evaluation budgets. Budgeting for DE includes anticipating the costs of the evaluator’s regular participation in strategy and/or working team discussions, as well as allotting time for facilitating frequent group reflection and learning sessions about the evaluation’s findings.

One solution to the budgeting dilemma is to develop an evaluation plan that serves as a guide and base scope of work for the evaluation team, but also to build in a retainer that can cover emergent and unexpected evaluation costs.

One of the biggest challenges in budgeting for a developmental evaluation is that learnings that emerge from the findings and resulting conversations often create new questions and information needs. As a result, it is entirely possible that additional data collection activities may be requested, or that the range of data sources be expanded, all of which may test the boundaries of the contracted scope of work. Because DE does not rely on a fixed evaluation design with a set number and type of data collection activities, it is important that the budget also be flexible to adapt to changes along the way. —Meg Long, *OMG Center for Collaborative Learning*

This amount and how it gets used can be negotiated and agreed upon by the funder and the evaluators as needs arise. Throughout the developmental evaluation, the team should take time to revisit and amend the evaluation plan and budget. Mayur Patel of the Knight Foundation offers the following advice to funders about developmental evaluation budgeting:

My impulse with all our evaluation work is to try to get everything nailed down as much as possible before we start. In many cases, that's the wrong impulse for a DE. You have to have so much more trust in the capacity of the evaluation team. We try to build in a minimum amount of work or set of activities that will have to be done for us to be satisfied and then create phases and "stage gates" where we say, at this juncture, we know we're going to have to make several decisions about the direction of the work.... We think there might be a ceiling to the amount, and we'll decide how to allocate that once we know more about how the work is unfolding.

Who gets to decide when and how to adapt?

When multiple organizations are involved in an innovative strategy, the question inevitably arises: *who decides what kind of adaptation is necessary, and who is expected to adapt?* Some developmental evaluations are designed to support grantees' innovation and adaptation, while others are designed primarily for the funder's use. This can be a particularly tough balance to strike when a funder commissions and manages the evaluation but expects the grantee(s) or a larger group of partners to adapt in response to the findings. Some grantees have reported frustration when DE is designed to inform the *funder's* strategy (including whether or not the grantee will continue to be part of the strategy), yet the grantee bears the burden of providing the bulk of the evaluation data. Alternately, when grantees have limited power over or are not included in the process of making sense of the data and their implications for action, grantees can experience "whiplash," as funders may expect them to transform their strategies too often or too early in response to developmental evaluation data.

Because DE is intended to support innovation and experimentation, those who are expected to revise their strategy in response to the developmental evaluation must play an active role in the evaluation, including helping to define what questions data collection should answer, participating in data interpretation, and deciding how to adapt in response to the data. Even when the DE is intended to inform the funder's strategy, inviting grantees and other partners to help make sense of the data and explore its implications can enrich the funder's strategy. At the very least, funders and evaluators should be transparent about how "sensemaking" will occur and how they will make decisions about adaptation. Additionally, the funder and evaluator can actively work to create safe opportunities and structures within which grantees can make sense of data on their own.

With developmental evaluation, the data often suggests that we change course. But sometimes we need to push the funder to keep from changing too much and becoming rudderless so that the grantees have unfair expectations on them or are forced to do something totally different than what they started. And we have to negotiate: if the funder changes his or her theory of change, how do we expect grantees to change theirs? Is it directive, so that the grantees are expected to re-align to meet the funder's changes, or is there some interaction both ways between the foundation's theory of change and the sites' theories of change?
—Kate Sandel, *Strategic Data Project*

Developmental evaluators and programmatic decision makers can productively manage these tough issues by paying close attention to when and where they appear and then addressing them head-on as a team. In most cases, doing so will provide the team with a natural opportunity to explore questions that are important not only to the smooth functioning of the evaluation, but also to the overall strategy. Revisiting whether the team's understanding of accountability (accountability for outcomes or accountability for learning) still applies as the innovation evolves

can help the team identify whether their effort is stabilizing into an intervention that could be expected to produce predictable outcomes. Weighing the trade-offs between methodological rigor and timeliness of information can help team members pinpoint when they should be making key decisions and which decision points they feel least confident about. Attending to the question of *who adapts* and who has decision-making power about *how to adapt* can help the team explore and clarify the role each player has in an innovative effort. In short, these tough questions are not a reason to shy away from using developmental evaluation to support innovation. Instead, they make both the evaluation and the social change effort better.

It's Time To Evaluate Differently

Developmental evaluation contributes to learning about the effectiveness, relevance, and clarity of an organization's strategy. It helps manage uncertainty in complex and changing environments, giving foundations and nonprofits greater confidence to experiment with solutions where none are known.

As this Foundation shifted its funding to complex, long-term initiatives that are not as pre-planned as emergent, the inadequacy of the usual evaluation methods became evident. We needed a compass, not a roadmap. We needed to know we were on the right track, not that we had arrived at a pre-determined spot, on budget, and at the specified time. —Tim Brodhead, past-President and CEO, The J.W. McConnell Family Foundation, quoted in, Dozois, Langlois, & Blanchet-Cohen, 2010, p. 6

The time is well past ripe for adding developmental evaluation to the philanthropic sector's portfolio of evaluation approaches. If we are serious about finding and using innovative ideas and practices to help solve complex, deeply rooted, and pervasive social problems, then we must have access to high quality, timely, and useful information from the beginning of an innovation's design throughout its evolution. We cannot afford to wait six months, a year, or longer to know if we had the right strategy, employed the most effective tactics, made the right investments, engaged the right stakeholders, or followed a path to achieving our desired outcomes. The benefits of engaging in and supporting developmental evaluation are fundamentally grounded in an organization's interest in and commitment to learning and using data to inform strategic decision-making. The insights and clarity that DE provides—about the influence and effects of a social innovation—may lead to a greater likelihood that the intervention's goals will be achieved.

References

- Coffman, J. and Beer, T. C. (2011). *Evaluation to support strategic learning: Principles and practices*. Center for Evaluation innovation.
- Dozois, E., Langlois, M., & Blanchet-Cohen, N. (July 2010). *DE 201: A practitioner's guide to developmental evaluation*. The J.W. McConnell Family Foundation and the International Institute for Child Rights and Development. Montreal, Canada.
- Fulton, K., Kasper, G., & Kibbe, B. (2010). *What's next for philanthropy?* Monitor Institute.
- Gamble, J.A.A. (2006). *A developmental evaluation primer*. The J.W. McConnell Family Foundation. Montreal, Canada.
- Morariu, J., & Brennan, K. (2009) Effective advocacy evaluation. *The Foundation Review*, 1 (3), 105.
- Mulgan, G. (2007) Social innovation: What it is, why it matters, and how it can be accelerated. Oxford: Skoll Centre for Social Entrepreneurship, SAID Business School, Oxford.
- Patton, M.Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York: Guilford Press.
- Patton, M.Q. (December 2011). *How to evaluate interventions in complex dynamic environments: Developmental evaluation*. <http://mymande.org/index.php?q=content/how-evaluate-interventions-complex-dynamic-environments>
- Patton, M. Q. (2012). Developmental evaluation workshop, the evaluators' institute, 2012, Washington, DC.

Overview of the Research Study and Acknowledgements

Purpose:

The purpose of this study was to:

- Understand the extent to which funders who are experimenting with innovative social change strategies are using an adaptive evaluation approach.
- Make explicit how adaptive evaluation can contribute to internal evaluation approaches and systems.
- Articulate a compelling case that adaptive evaluation is *imperative* for innovative, systems-oriented social change efforts and is an indispensable means of strategic learning.
- Leave readers with an understanding of the organizational requirements and benefits of adopting an adaptive evaluation approach.

Our hope is that the study's results and white paper will support strategic learning for all funders—not just those with large endowments and/or evaluation staff. Developmental evaluation is about thinking and practicing evaluation differently, regardless of size, assets, or focus. To that end, foundations of all sizes would benefit from including developmental evaluations into their portfolio of evaluation approaches.

Framing Questions:

- What environmental conditions have created the need for strategic and catalytic philanthropy, and what is the parallel need for evaluation?
- What kinds of evaluation approaches are used by funders who are experimenting with innovative, systems-oriented grantmaking approaches? To what extent are they satisfied with these approaches? What are they learning and not learning from these evaluations?
- How is developmental evaluation different from other evaluation approaches? To what extent does adaptive evaluation require the evaluator to have a different set of knowledge, skills, and attitudes/beliefs?
- What does it look like to do developmental evaluation? What conditions are necessary for engaging in developmental evaluations?
- What is the role of developmental evaluation within the foundation's portfolio of evaluation approaches?
- How must funders re-think the role of evaluation—and transform their approach to evaluation—in order to support strategic learning?
- What benefits exist for funders and nonprofits to using a developmental evaluation approach when evaluating change in complex social systems? Why should funders invest in this type of evaluation?

Data Collection Methods:

- Literature Review: We reviewed more than 70 articles, papers, books, reports, and blog posts on the topics of developmental evaluation, strategic learning, learning and evaluation in philanthropy, and complex adaptive systems.
- Phone Interviews: We conducted one- to two-hour interviews with the following foundation staff members, grantees, and evaluation consultants to better understand their experiences and perspectives on developmental evaluation. We are very grateful to these individuals, who gave freely of their time and greatly informed the thinking that is reflected in this paper. In alphabetical order, they are:
 1. Natasha Blanchet-Cohen, Senior Associate, International Institute for Child Rights and Development and Assistant Professor, Concordia University

2. Mark Cabaj, President, Here to There, and Associate, Tamarack: An Institute for Community Engagement
3. John Cawley, Program Manager, J.W. McConnell Family Foundation
4. Michelle Clarke, Executive Director, Burns Memorial Fund
5. Elizabeth Dozois, Principal, Word on the Street Ltd.
6. Kendall Guthrie, Lead Senior Program Officer, Measurement Learning and Evaluation Education, Postsecondary Success, U.S. Programs, The Bill & Melinda Gates Foundation
7. Marc Langois, Skipping Stones Consulting
8. Meg Long, Deputy Director, OMG Center for Collaborative Learning
9. Margaret Hargreaves, Senior Health Researcher, Mathematica
10. Liz Joyner, Executive Director, The Village Square, Inc.
11. Ellen Martin, Senior Consultant, FSG
12. Kate McKegg, Director, The Knowledge Institute (a member of the Kinnect Group)
13. Sara Meyer, Portfolio Manager, Education, Postsecondary Success, The Bill & Melinda Gates Foundation
14. Jenn Miller, Senior Policy and Research Analyst, Ontario Trillium Foundation
15. Mayur Patel, Vice-President Strategy and Assessment, Knight Foundation
16. Susan Patterson, Program Director, Knight Foundation
17. Michael Q. Patton, Founder and Director, Utilization-Focused Evaluation
18. Kate Sandel, Fellow, Strategic Data Project
19. Dan Wilson, Manager, Policy, Research and Evaluation, Ontario Trillium Foundation

About the Authors

HALLIE PRESKILL, PhD, is a Managing Director with FSG, where she oversees the firm's Strategic Learning and Evaluation practice. FSG is a nonprofit consulting firm specializing in strategy, evaluation, and research, with international teams that work across all sectors by partnering with corporations, foundations, school systems, nonprofits, and governments in every region of the globe. FSG's goal is to help all types of organizations—individually and collectively—achieve greater social change. Learn more about Hallie and FSG at www.fsg.org.

TANYA BEER is Associate Director of the Center for Evaluation Innovation in Washington, D.C. The Center's mission is to build the field of evaluation in areas that are challenging to measure and where traditional program evaluation approaches are not always a good fit, such as advocacy, systems change and communications. The Center's work is anchored in a strong belief in strategic learning, or using evaluation to help organizations learn and adapt quickly to increase their likelihood of success. For more about the Tanya and the Center, see www.evaluationinnovation.org.