

Thoughts on Assessing Performance and Building Effectiveness

Below are several features and ideas the SBP design team discussed that can improve the practice of assessment at OSF. This document is not meant to be an overarching policy proposal, but rather ideas for discussion. We thought it would be most useful to begin, however, by stating some shared understandings that should guide our efforts:

1. The social change processes we immerse ourselves in are complex, dynamic, and involve multiple actors. Linear descriptions of how change happens are at best rough approximations. This affects both our ability to predict exactly how our resources will be used over time (plans are our best approximations at a given time, but we continuously pay attention to what is emerging and adapt accordingly) as well as our ability to attribute a particular change to our intervention. In general our assessment efforts need to reflect that complexity and seek to capture adaptation to changes and emerging opportunities and determine how significant our contribution was in any given effort, rather than being based on a model that falsely promises predictability.
2. The ultimate purpose of assessment at OSF must be at least twofold: a) to ensure accountability and provide sufficient oversight of the funds with which we are entrusted; and b) to constantly learn from and thereby improve our work. Any assessment exercise will need to include both aspects. Senior management and unit leadership have the responsibility for establishing a learning environment across the organization and within each unit. As a term, "learning" carries the danger of being a catchall category used to justify a less thoughtful or less rigorous approach to assessment. But a real learning organization does not leave the concept in that abstract plane. Although everyone is learning all the time on an individual level, on an organizational level we need a deliberate effort to establish a culture and systems that reward critical reflection and sharing.
3. We recognize that field-building and unit initiative work is deeply interwoven, and our assessment of our work in each area will also be linked. The differing level of agency we project in each category, however, points to somewhat divergent assessment strategies. Field work will be more centered on the development of key grantees, partners, individual leaders and the advancement of the field in a broad sense. Thus assessment will hew towards the qualitative end of the spectrum, alongside some quantitative measures that track the growth of key partners and perhaps public opinions on field-related issues. When units drive or participate in initiatives with clear change objectives, assessment of progress can happen in a rather more straightforward (but by no means simple) manner.

For both field support and specific initiatives, we will necessarily focus less on whether our interventions have been carried out as planned and more on what has changed in the subject area, how we have contributed to it, and how we should react. All the monitoring and evaluation practices proposed below should be carried out in that spirit of flexibility.

4. Good monitoring, assessment and learning begin with good planning. To be able to assess progress and identify how our work can be more effective, we need to incorporate evaluative thinking in our strategic planning efforts. The new common template for strategy development implies a welcome rigor in this sense. As we monitor and assess our program strategies, we may well identify ways in which our strategy template can evolve in ways that improve our ability to assess the effectiveness of our strategies.

Clarifying an OSF approach to accountability and performance assessment

Accountability¹: When thinking about accountability, it may be useful to ask ourselves what we mean: Accountability to whom? for what? and how? Typically accountability can be upward, downward or mutual. For OSF, this might mean upward accountability to the Global Board and George Soros; downward accountability to partners, allies and stakeholders; and mutual accountability understood as a responsibility to the organization's mission and purpose, as well as to other colleagues in advancing the OSF Mission. Some of these levels may require financial and governance tools; all of them may involve considering "performance in achieving results, identifying effective strategies, and focusing on a mission that creates the greatest social value."²

For this purpose, we will need both tools (evaluations, performance assessments), and processes like spaces for reflection and analysis that can be applied back to program plans and strategies as well as more attention to assessment in our planning processes. In addition, a practice of transparency through, for example, allowing PhD or Masters' students to use internal documents in their research, would be another pathway to accountability. Both tools and processes are included in the recommendations below.

Performance assessment: It may be helpful to briefly mention some trends and current innovation in evaluation and performance assessment from which we might profitably draw.

Recent thinking in the sector of funders and nonprofits working towards social change goals has included the notions of *complexity*, *contribution over attribution*, and the *focus on improving effectiveness* mentioned above. These notions imply the need to pay attention to what is happening in our context and as a result of our work, for making sense of and responding strategically to emerging patterns or opportunities. Benchmarks can be used effectively to instill rigor and discipline in our programming and our monitoring, and as importantly, in understanding how and why we may change course in our work. Although such benchmarks or indicators may be useful in order to make visible the ways in which contexts are changing, they should not blind us to unexpected openings or changes that require an adjustment of strategy. It is important to be deliberate in our planning, but we acknowledge that divergence from benchmarks does not necessarily imply a failure; performance is not only about whether we hit benchmarks, but whether we are moving forward on more meaningful progress.

[Additional information about "developmental evaluation" (or "evaluation for social innovation") and "collective impact", two frameworks that we have found useful to draw from, can be found in Annex I]

Features of an OSF assessment approach

Energies at the senior management level, as well as within the SBP team, have not yet been particularly focused on developing policies for assessing our performance and effectiveness. Rather than proposing a framework for an entire system, we are recommending for experimentation in 2013 several features, which we set out below, with an eye towards developing clearer guidance for the Foundations' 2014 efforts. We recognize that these recommendations are an incomplete set of features that we see as minimum requirements.

For 2014 and beyond, we believe that OSF is well-positioned to define an OSF approach to assessment that is rigorous, that focuses on what is meaningful, that minimizes distorting disincentives, that deepens our understanding, and that increases our effectiveness. Unlike many foundations, OSF has a founder, a president, and board members who resist a formulaic approach and who are comfortable with

¹ Categories and ideas here were taken from an article by Alnoor Ebrahim at the Harvard Business School, called "The Many Faces of Nonprofit Accountability," Working paper 10-069, HBS, 2010.

² From Ebrahim working paper, *ibid*.

experimentation. OSF's institutional values of flexibility, innovation, and adaptation and George Soros' philosophy of "imperfect understanding" lend themselves to experimentation with how we can best assure appropriate accountability and meaningful assessment in ways that encourage greater understanding, effectiveness, and impact. Because of the OSF network's size and breadth, we have the opportunity to become a live laboratory that contributes to relevant and useful thinking on how to go about assessing performance on complex and dynamic social issues, where bringing about positive change does not involve following a recipe. How to assess performance in the fields in which OSF works is itself an emerging field. OSF can both draw from and contribute to the promising new ideas coming from this field, while developing a system (or systems) that work for us.

We believe that it would be unwise to box ourselves in too soon, as we may end up hewing to familiar, yet unsatisfying practices. We want for our tools and processes to serve us in furthering OSF's mission, rather than the other way around, and we want to avoid the pitfalls of evaluation practices that discourage innovation and big picture thinking. We thus suggest implementing some basic elements for assessment in 2013, and generating conditions for productive experimentation.

- **Ongoing, regular monitoring.** Each unit and sub-unit is responsible for carrying out robust, active, ongoing monitoring³ of the portfolios for which it is responsible. Good monitoring begins with a clear definition from the outset of what OSF's interests are in funding or otherwise engaging with a particular organization or initiative, what we hope to see from this engagement, and what signals of change we will watch closely.⁴ It will include site visits, phone calls and emails, and information picked up in various ways, in addition to formal reporting by grantees and the assessments we make of individual grantee performance. Portfolio reviews can be useful tools both to collect and share monitoring results. We place greater emphasis on monitoring than on evaluation, as monitoring provides specific information and evidence that serves as the basis for evaluative assessments. While evaluation offers an opinion about performance, monitoring provides the building blocks on which evaluative opinions are based.
- **Shared reflection and self-assessment.** At least twice a year, units (and/or sub-units) should conduct a staff-level review of progress towards objectives. Essentially, this is a shared reflection and analysis on what we have seen through our monitoring. Rather than to report only on the implementation of activities, these reviews can be used to understand whether we have identified the problem we seek to address correctly, to clarify what has changed and how in the environment, and to articulate how the units/subunits have adapted or should adapt to these changes. Units or sub-units should produce documents that can be shared prior to advisory board meetings and dovetail with any portfolio reviews that are conducted.
- **Selective tracking of bellwether indicators.** We have not (yet?) developed either the culture or systems for tracking indicators rigorously for all of the results we hope to achieve. However, any unit/sub-unit director can point to several changes that we believe need to happen if our work in a particular area is to be considered a success. We recommend that OSF begin by identifying and tracking progress on such bellwethers by picking a couple at each subunit's level and a couple at the overall unit level. Each unit and sub-unit should pick at least one field-building outcome and one initiative outcome that is appropriate for such tracking. Progresses against these outcomes

³ "Monitoring is the ongoing collection and analysis of information on progress toward results, changes in context, strategies and implementation." Cheyanne Church and Mark M. Rogers, *Designing for Results: Integrating Monitoring and Evaluation in Conflict Transformation Programs*, Search for Common Ground, 2006.

⁴ Information that signals change can be called indicators. "Indicators are inevitable approximations. They are not the same as the desired change, but only an indicator of that change. They are imperfect and vary in validity and reliability." Michael Quinn Patton, *Utilization-Focused Evaluation*, Third Edition, SAGE Publications, 1996, p. 159.

could be checked on an annual basis and discussed with program staff, and could be a central part of the above mentioned annual review of progress towards objectives. The review of these selective indicators (progress, setbacks, implications of a changing context etc.) could also serve as a useful basis for new strategy development every two years.

- **Portfolio reviews**⁵. Rather than focusing on individual investments, units/sub-units can use portfolio reviews to look at a set of grants or interventions which the unit/sub-units believe advance a particular set of objectives. Such reviews can be used among staff in order to share learning, with boards to understand how particular objectives are being pursued programmatically, by senior management to provide insight into program implementation and management.
- **Mix of peer and external reviews in evaluation.** Occasionally, and no less frequently than once every four years, a deeper and more distanced assessment of a unit's performance and effectiveness should be required. We would recommend the judicious use of external evaluators and a more frequent practice of cross-unit peer evaluation (perhaps with board members participating as available/appropriate). This would build horizontal linkages and spark new program ideas across the network while ensuring that the reviewer has deep knowledge of the OSF context. We would need to ensure that staff were both guided effectively in conducting such peer reviews and relieved of some of their regular responsibilities temporarily so as to free up sufficient time.
- **The bigger picture**
All the different OSF units work to advance open society values, and since ultimately this is what we are all concerned about, OSF may want to think about ways by which we can assess whether we are advancing (or defending) open society values. Acknowledging the complexities of measuring impact, the SBP design team would suggest not being prescriptive at this stage but encouraging units to innovate by using a variety of approaches and being prepared to share their experience within the network.
This could include:
 - Developmental evaluation, to look at emerging patterns rather than pre-set benchmarks.
 - Storytelling, to show how a unit is advancing norms, contributing to change systems or creating changes in people's lives.
 - Polling (e.g. piggybacking on existing) or media tracking (e.g using tools like [Media Cloud](#)).
 - Best & Worst Grants: Some grant making units have found it useful to hold on an annual basis a best & worst grant competition where staff present to each other what they see as their best and worst grant of the prior year. This is a safe space and no one is judged on the outcomes they present, only on our ability to learn from both successes and trials. Grant making is a complex process that requires taking risks that sometimes do not pay off. This is a chance to learn and improve our practice.
 - If certain programs or shared frameworks experiment with the methodology of collective impact (see Annex 1), involved stakeholders together would define indicators or benchmarks of progress that can be monitored.
- **Assessment of shared strategic frameworks.** While shared frameworks will certainly benefit from the techniques suggested above, their cross-cutting nature implies a few different levels of assessment. Team and contributing unit leaders should ensure they are assessing:
 - What overall progress has been made towards the goals identified by the shared framework;

⁵ A portfolio can be defined as a set of grants that correlates directly to a particular thematic priority, geographic context or strategic method. The purpose of the portfolio review is to provide a level of detail about grant-making programs that allows a useful examination of strategy implementation. The desirable level of detail will depend upon the purpose of the review and the audience.

- What the unit in question's contribution was to implementing the shared framework;
- To the extent the framework is taking multiple vectors towards the same ultimate goals, how these differing theories of change have played out, and whether there is any need to reconsider their mix and number.
- Monitoring shifts in the context and articulating any corresponding shifts in strategy adopted by the team, along with the team's reasoning in doing so.

Looking Ahead

- **Remaining challenges.** We acknowledge that we have not addressed all outstanding questions with this skeletal set of proposals. Grappling with these challenges is not new for many units facing these challenges in relative isolation from each other. In many cases, particular units have developed attempted solutions to these challenges from which the overall network can learn:
 - *Systems that are planning-dominant and reporting-absent.* We fully support the ambition to make our planning documents living guides to our work that not only outline strategies but provide benchmarks for understanding our progress. More attention to monitoring and reporting will set us on that path.
 - *Limited agreement on language and concepts when speaking about planning, assessment or learning.* The unification of unit strategy processes and the goal of issuing a glossary and instructional webinars is a great step in the right direction.
 - *A need to connect our assessment of grantees' performance and our own self-assessment.* One useful practice is the Latin America Program's issuance of a "program grant letter" that follows the official one and outlines a) how the grant fits into LAP's current strategy and b) the grounds on which grantee performance will be assessed.
 - *Incorporating monitoring into current grant-making practice.* Some programs have a template for trip reports that helps staff crystalize and make accessible monitoring information.
 - *Staff time and capacity.* Given the multipolarity of our network and the various layers of our new strategy process, any new assessment or learning systems could add greatly to the reporting burden units face. The burden is greater at present because of the many processes that are being undertaken during the transition. Any new processes should thus dovetail with current practice as much as possible. We also bear a responsibility to build up staff capacity on learning, monitoring, and evaluation.
- **2013 and beyond:**
 - In 2013, Program units or shared framework teams are encouraged to adopt some of the approaches proposed above and to experiment with adapting various approaches to their practice. They are expected to capture associated learning and to share their experiences across the network. However, we anticipate that it may be quite difficult for this kind of innovation to take place at the unit or team level without significant support from a central unit or encouragement and resources to seek necessary assistance. Past experience with working groups in this area has shown that they do not work well in the absence of a clear charge from OSF's leaders. We have a great opportunity to build learning systems and an enabling environment that serve us and our collective mission (rather than the reverse). But that requires senior management to make this a priority and create the context in which such a system and culture will thrive.
 - In addition, in order to develop clearer guidance for the Foundations' 2014 efforts, OSF's leadership could bring together some leading thinkers in this field to engage on a set of questions OSF faces in developing a set of practices that are useful for different types of

units, but which provide the Global Board, senior management, advisory boards, and staff throughout the network with information that is accessible and relatable to each other. This would not be a consulting company engagement, but rather asking specific individuals to respond to a set of questions and engage in conversation with an internal working group.

Annex I

Two frameworks that may be useful for us to draw from are “developmental evaluation” and “collective impact.” Both of these have elements that could be significantly useful to us, if we adapt it to our own purposes. Following are some excerpts from recent articles:

Evaluating Social Innovation, Hallie Preskill and Tanya Beer, FSG and the Center for Evaluation Innovation, 2012:

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.”

.....

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.

.....

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.

Embracing Emergence: How Collective Impact Addresses Complexity, John Kania and Mark Kramer, Stanford Social Innovation Review, 2013

The solutions we have come to expect in the social sector often involve discrete programs that address a social problem through a carefully worked out theory of change, relying on incremental resources from funders, and ideally supported by an evaluation that attributes to the program the impact achieved. Once proven, these solutions can scale up by spreading to other organizations.

The problem is that such predetermined solutions rarely work under conditions of complexity—conditions that apply to most major social problems—when the unpredictable interactions of multiple players determine the outcomes. And even when successful interventions are found, adoption spreads very gradually, if it spreads at all.

Collective impact works differently. The process and results of collective impact are emergent rather than predetermined, the necessary resources and innovations often already exist but have not yet been recognized, learning is continuous, and adoption happens simultaneously among many different organizations.