Next Generation Evaluation: Embracing Complexity, Connectivity, and Change
A Learning Brief

Srik Gopalakrishnan, Hallie Preskill PhD, Shijie (SJ) Lu
Discovering better ways to solve social problems

FSG is a nonprofit consulting firm specializing in strategy, evaluation, and research.

Our international teams work across all sectors by partnering with corporations, foundations, school systems, nonprofits, and governments in every region of the globe. Our goal is to help companies and organizations—individually and collectively—achieve greater social change.

Working with many of the world’s leading corporations, nonprofit organizations, and charitable foundations, FSG has completed more than 400 consulting engagements around the world, produced dozens of research reports, published influential articles in *Harvard Business Review* and *Stanford Social Innovation Review*, and has been featured in *The New York Times*, *Wall Street Journal*, *Economist*, *Financial Times*, *BusinessWeek*, *Fast Company*, *Forbes*, and on *NPR*, amongst others.

Learn more about FSG at www.fsg.org.
Acknowledgements

FSG gratefully acknowledges the support of the interviewees we spoke with as part of this research project. Their invaluable contributions helped make this effort possible.

We would like to express our gratitude to the Skoll Foundation, The Colorado Trust, and The California Endowment for their generous support of the Next Generation Evaluation project. In addition, we would also like to thank Eric Nee and Regina Starr Ridley from Stanford Social Innovation Review for their partnership in planning and executing the Next Generation Evaluation conference.

Finally, this report would not have been possible without the support and insightful guidance of our colleagues Mark Kramer, John Kania, Eva Nico, and Alrie McNiff Daniels.

This report was first published August 30th 2013.

Authors

Srik Gopalakrishnan
Director, FSG
srik.gopalakrishnan@fsg.org

Hallie Preskill PhD
Managing Director, FSG
hallie.preskill@fsg.org

Shijie (SJ) Lu
Associate, FSG
shijie.lu@fsg.org

All statements and conclusions, unless specifically attributed to another source, are those of the authors and do not necessarily reflect those of the other organizations or references noted in this report.
**Foreword**

A core thread of FSG’s work has always been the advancement of ideas within the social sector. We do this by giving form and structure to emerging ideas, highlighting real-world practitioners who are on the cutting edge, and pushing social sector organizations beyond their comfort zone in adapting innovative practices. In addition, we seek to bring together social sector leaders to explore these ideas in greater depth.

The *Next Generation Evaluation* initiative is our attempt to engage the field in conversation around how evaluation and learning should evolve in order to continue to stay relevant, timely, and useful. As part of the initiative, we are publishing this Learning Brief, as well as convening a conference at Stanford University in partnership with *Stanford Social Innovation Review* on November 14, 2013. The conversation will extend beyond the conference in the form of an article or white paper, as well as in ongoing conversations through blogs and social media.

We welcome your comments, ideas, and reflections both before and after the *Next Generation Evaluation* conference. Please send us your thoughts via email to nextgeneval@fsg.org and/or contact us through Twitter @FSGTweets using #nextgeneval. We look forward to hearing from you.
Table of Contents

Introduction ..................................................................................................................... 2
Why Next Generation Evaluation? .................................................................................. 3
The Six Characteristics of Next Generation Evaluation ................................................. 5
Application in Practice: Three "Game-Changing" Approaches ................................. 8
Where Do We Go From Here? ....................................................................................... 13
References ..................................................................................................................... 14
List of Interviewees ....................................................................................................... 15
Introduction

Over the last several years, the public and philanthropic sectors have been developing and experimenting with new and creative approaches to solving social problems in response to the increasing recognition that complex problems require complex solutions. Some of these efforts may be described as truly experimental, in that they are using new untested practices or models, making serious attempts to affect systems change, and using innovative technologies to improve the social well-being of many.

At the same time, the evaluation field has been evolving in ways that are more pluralistic, inclusive, and increasingly multi-disciplinary. The interest in, and commitment to, strategic learning, evaluation capacity building, and timely and useful information, all bode well for ensuring that evaluation continues to be an indispensable tool for decision making and action. However, as the social sector continues to innovate, it will be ever more important for evaluation to evolve and grow in ways that respond to the increasingly complex, inter-connected, and rapidly changing world in which these initiatives live.

Consider for example, the work of the Carnegie Foundation for Advancement of Teaching. When the foundation wanted to tackle the complex issue of high failure rates in developmental mathematics among community college students, it used Networked Improvement Communities (NICs) that diligently engaged in “improvement research” (Berwick, 2008), a concept the foundation adapted from the field of quality improvement in healthcare. That strategy’s underlying premise was that organizations can find solutions to complex problems through networks, and that each network would design and test effective interventions while generating information about how, for whom, and under what organizational conditions those interventions work (Bryk, Gomez, and Grunow, 2011).

In its attempt to shift the paradigm of how improvement happens in education, the foundation also began to reframe the role of evaluation into a helpful tool for improvement in the social sector. Common measures for examining and talking about outcomes across the network were agreed upon and implemented. Predictive analytics were employed to support early identification of students at risk of falling behind, and to engage them in productive ways to prevent failure. Finally, real-time and rapid feedback processes were put in place to support the NICs’ “learning through doing” ethos. In other words, evaluative inquiry (Preskill and Torres, 1999) was built into the NICs as a way of doing business.

The Carnegie Foundation is one of a growing number of organizations that are beginning to define the ways that evaluation must evolve to stay relevant, timely, and useful in an emergent and dynamic environment. In this Learning Brief, we draw from literature and research, as well as more than a dozen interviews with foundation leaders, evaluation practitioners, and social sector thought leaders to start the conversation around what we at FSG have come to call Next Generation Evaluation. The conversation will continue at the Next Generation Evaluation conference at Stanford University on November 14, 2013, and beyond the conference through additional publications and blogs. We invite you to join us in this journey as we explore the next frontier of evaluation and learning.
Why Next Generation Evaluation?

A number of recent changes have created the need for the social sector to learn more effectively. The economic crisis has created an environment of greater social need and declining resources for vulnerable populations. At the same time, the sector has benefitted from an increase in attention, effort, and private resources that are invested in discovering better ways to solve social problems. New ideas have expanded the ways in which organizations think about improving society and have opened up additional opportunities for collaboration. Simultaneously, technological innovation has expanded the sector’s capacity for understanding and addressing social problems, and it has created more data that the social sector can leverage in its work.

We found three primary trends that are driving the need for evaluation to evolve:

1. **New Philanthropic Innovations**: There is an increasing realization that traditional philanthropic models have had limited success in curing chronic social ills. Even tested solutions are forced to experiment when faced with the challenge of scaling to new populations and geographies. In addition, several new and often untested approaches are coming to the fore, including social entrepreneurship, impact investing, social impact bonds, and others that do not lend themselves to traditional methods of evaluation. This growth of experimentation in the social sector demands that evaluation better capture learning in complex environments.

2. **Different Rules of Interaction**: Over the past few years, the pace of change in the sector has accelerated, increasing the number of solutions and approaches involving multiple interdependent actors. The environment has become more fluid and the solutions less predictable. Ideas such as collective impact have reinforced the need for organizations to work together across boundaries, share information, and build on the lessons each has learned. As rules of interaction between social sector organizations evolve, evaluation approaches and methods must adapt to provide relevant, credible, and useful feedback to social sector stakeholders.

3. **Proliferation of Digital Infrastructure**: Technology adoption is widespread and a “digital infrastructure” (Bernholz, 2013) is now emerging for philanthropy and the social sector. The explosion of social media and the use of handheld devices have rapidly reduced the length and duration of the feedback cycle between funders, nonprofits, and end beneficiaries. As more data are created and analyzed, evaluation must expand to allow social sector stakeholders to better understand the nature of social issues and maximize the use and effectiveness of data to solve social problems.
As we studied these trends, we realized that they constituted a fundamental change to the context in which evaluation operates. This led to us embark on an effort to understand how social sector organizations should think about and practice evaluation differently in light of these trends. Our research led to us to identify six characteristics that unite the social sector’s most promising efforts to expand its practice of evaluation, as well as three specific “game-changing” approaches that best embodied the potential of Next Generation Evaluation. This Learning Brief provides a summary of what we’ve learned thus far.
The Six Characteristics of Next Generation Evaluation

We identified six characteristics that exemplify how leading social sector organizations are thinking about Next Generation Evaluation (see Table 1). It is important to reiterate that we see these as complementary, rather than outright replacements of traditional evaluation approaches and methods.

<table>
<thead>
<tr>
<th>Not Just…</th>
<th>But Also…</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus on Individual Programs and Projects</td>
<td>Focus on Whole Systems</td>
</tr>
<tr>
<td>Fixed Evaluation Plan with Interim and Year-end Reports</td>
<td>Shorter Cycles, More Real-Time Feedback Using Alternative Formats</td>
</tr>
<tr>
<td>Traditional Data Collection Methods</td>
<td>Newer, Innovative, Often Digital, Data Collection</td>
</tr>
<tr>
<td>One Foundation, One Grantee, One Evaluation</td>
<td>Shared Responsibility for Data Collection and Learning Across Multiple Organizations</td>
</tr>
<tr>
<td>Traditional Data Reporting Techniques</td>
<td>Use of Sophisticated Data Visualization and Infographics</td>
</tr>
<tr>
<td>Evaluator Collecting Data</td>
<td>Everyone Collecting and Using Data as Part of Ongoing Practice</td>
</tr>
</tbody>
</table>

1. **Evaluation of whole systems, not just individual programs and projects**

There is growing interest in moving beyond evaluating singular programs and grants toward more systematic evaluation. Programmatic evaluations are certainly useful in some situations, but as our understanding grows about the complex environment in which most programs work, a wide-angled view of the system and its component parts becomes increasingly essential.

*Focusing on systems makes you adopt a more long-term time horizon, but also invites you to think about sustainability in a much more meaningful way, especially if your goal is to change the way systems function, as opposed to just producing the particular outcomes you are after.* – Tanya Beer, Center for Evaluation Innovation
2. **Shorter cycles and more real-time feedback using alternative formats, not just a fixed plan with end-of-year reports**

Social sector organizations are well-placed to move toward shorter and more frequent feedback cycles, especially when evaluation focuses on learning (as opposed to accountability). Shorter feedback cycles help facilitate swifter learning when experimenting with social innovations, a key component of many current change efforts.

*Traditional evaluators say that ‘if it can’t be written, it doesn’t count,’ but I disagree. The opportunity for learning is often lost because of an ‘inappropriate’ method of sharing. Sometimes doing a formal report misses the point of providing timely and useful learning.* – John Cawley, J.W. McConnell Family Foundation

3. **Newer, innovative, often digital, data collection, not just traditional data collection methods**

Increasingly, organizations are leveraging innovative, often digital, data collection methods in addition to traditional evaluation methods such as surveys and interviews. Data is increasingly being collected through the use of cellphones, text messaging, social media, web analytics, and sensors as digital technology becomes more of a norm in society.

*Consider a community that had lots of disasters and in between they had resilience training. When a disaster hits, we can use anonymized cellphone data to analyze whether people are calling, moving, or spending in the same way…. We can compare the level of disturbance before and after this program.* – Robert Kirkpatrick, UN Global Pulse

4. **Shared responsibility for learning across multiple organizations, not just one foundation tasking one grantee with an evaluation**

While foundations have traditionally funded single organizations to undertake well-defined evaluations, an increasing number of funders have been supporting initiatives where responsibility for data collection and learning is distributed across multiple organizations. Often, this takes place within collaborations and networks that allow for continuous feedback and iteration.

*We have been using a common framework with communities and walking them through a structured continuous improvement process for how to define a shared problem or goal, measures for helping to understand that goal, and then analyzing data to build a plan for continuous improvement.* – Geoff Zimmerman, Strive Partnership
5. **Use of sophisticated data visualization and infographics, not just traditional data reporting techniques**

Highly sophisticated data visualization techniques, mapping tools, and infographics that have emerged in recent years can help us more effectively understand the complexity of social phenomena and design solutions. Big advancements in geographical information systems, for example, allow us to map social issues across geographies in powerful ways that help us see trends and gaps.

*We’ve seen a huge development in visual data—use of maps and animation—this is what the world expects now. We are using mapping and network analysis to understand how systems are evolving and changing in real time.* – Michael Quinn Patton, Independent Evaluator

6. **Everyone collecting and using data as part of ongoing practice, not just the evaluator collecting data**

We see formal as well as informal networks and communities that are using tools for data collection and analysis in organic and emergent ways as part of ongoing practice. In addition, the new digital infrastructure allows diverse groups of people in multiple locations to generate, collect, and use data—often involving no evaluators at all.

*We are fundamentally changing the distinctions between information generators and information users. By cultivating the social practice of use, we are striving for a state where we are all improvers.* – Anthony Bryk, Carnegie Foundation for Advancement in Teaching
Application in Practice: Three "Game-Changing" Approaches

Game-changer (noun): a newly introduced element or factor that changes an existing situation or activity in a significant way (Merriam-Webster)

Even if we accept the six characteristics of Next Generation Evaluation, the challenge still remains to translate them into practice. Fortunately, three approaches have emerged in recent years that embrace the Next Generation Evaluation characteristics and have the potential to change how we view evaluation in a significant way. They are Developmental Evaluation, Shared Measurement, and Big Data.

1. Developmental Evaluation

Developmental Evaluation, which emerged as a new evaluation approach conceptualized and described by Michael Quinn Patton, offers evaluators, nonprofit organizations, and funders a significantly different way of engaging in evaluation practice (Patton, 2011). Instead of operating against a fixed evaluation plan along with a linear and predictive logic model, Developmental Evaluation offers a much more real-time, learning-oriented, feedback-based, and insight-driven approach to program design and early implementation of a program or initiative.

Developmental Evaluation has five characteristics that distinguish it from formative and summative evaluation approaches (Patton, 2011; Preskill and Beer, 2012). 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. Given these characteristics, Developmental Evaluation touches on almost all of the characteristics of Next Generation Evaluation.

A variety of organizations currently employ Developmental Evaluation, including leading foundations such as the Bill & Melinda Gates Foundation, the J. W. McConnell Family Foundation, the John S. and James L. Knight Foundation, and the Grand Rapids Community Foundation. An emerging body of research has focused on ways to effectively implement Developmental Evaluations, including the use of appropriate tools and processes for data collection, group facilitation, and results reporting (Langlois, Blanchet-Cohen, and Beer, 2013).
2. Shared Measurement

When groups of organizations come together to co-determine outcomes and indicators, to share their data, and to learn from each other, they engage in Shared Measurement. Coupled with a technology platform, shared measurement systems have the potential to build common understandings, reduce redundancies, identify gaps, and increase transparency, accountability, and alignment (Kramer, Parkhurst, and Vaidyanathan, 2009).

While shared measurement systems do not replace program level or strategic organization level evaluations, they can provide quantitative and qualitative data on key outcomes and indicators that help track progress, facilitate shared learning, and form the basis of and foundation for deeper inquiry. High quality shared measurement systems demonstrate several characteristics of Next Generation Evaluation, most notably shared responsibility for learning across organizations.

Organizations that have implemented Shared Measurement include the Strive Partnership, the E3 Alliance, the Cultural Data Project, the Portland Metro STEM partnership, the 100Kin10 network, and the Arizona Community Foundation in partnership with the Alliance of Arizona Nonprofits (“Project SAM”). While these initiatives range across sectors, each one has dedicated the time and resources necessary to build the strong infrastructure that is needed for shared measurement, including technology systems, human resources for data entry and analysis, and processes for network-wide learning and growth.

Spotlight: Using Developmental Evaluation to Support College Access and Success

The Challenge Scholars program was created to address the challenge of college access and success in Grand Rapids, Michigan. The program is a long-term partnership between the Grand Rapids Community Foundation and Grand Rapids Public Schools designed to help students successfully prepare for college. Many college scholarship programs focus on providing students with financial assistance, neglecting the more complex challenges around college readiness. Challenge Scholars is unique in that it focuses on providing wrap-around supports and improving academic preparation starting in 6th grade.

Given the level of experimentation, the dynamic nature of the partnership, and the constantly shifting environment around the initiative (for example, the school district is undergoing a restructuring effort), the implementers of the Challenge Scholars program chose a Developmental Evaluation approach. They work closely with the evaluators to ensure real-time learning to help them shape the program as it evolves. The evaluators, on their part, implement an adaptive and flexible design, and co-create synthesis and meaning with program implementers.

More details can be found at www.grfoundation.org/challengescholars
3. **Big Data**

The explosion of social media and the development of an increasingly robust digital infrastructure have created a staggering amount of data. To express this new phenomenon, experts have coined the term “Big Data”, which includes everything from sensors used to gather climate information and posts to social media sites, digital pictures and videos, purchase transaction records, cell phone use, and more. Increasingly, Big Data is characterized by the “3 Vs”—sheer **volume**, greater **variety**, and increased **velocity**.

Big Data brings to life several *Next Generation Evaluation* characteristics, including shorter feedback cycles and multiple people generating data, often just as a way of doing business. The number of social sector organizations that successfully use Big Data is small but growing. For instance, organizations such as UN Global Pulse and Ushahidi leverage large amounts of data generated through cellphone use, text messages, and social media posts to understand the immediate impact of disasters on economic development. The Justice Mapping Center has used GPS data to map the residential addresses of every inmate in various U.S. prison systems to identify what it calls “million-dollar blocks”—areas that spend more than $1 million annually to incarcerate the residents of a single census block, in order to focus social interventions in these areas.

---

1 The term “Big Data” was popularized by a May 2011 report from the McKinsey Global Institute (www.mckinsey.com/mgi) titled “Big Data: The Next Frontier for Innovation, Competition, and Productivity.” While most of its applications thus far have been in the private sector, there has been an increased focus on its use in the social sector as evidenced by topical conversations on www.marketsforgood.org and www.skollworldforum.org.
Big Data’s ramifications for the social sector remain unclear, though it has the potential to challenge some long held assumptions in the evaluation field. For example: would the availability of large amounts of data and the analytic engines that crunch the data make traditional evaluation designs and methods obsolete? Who needs sampling techniques when you can pretty much capture characteristics of whole populations? Why convene a focus group when you can just analyze Twitter feeds? Moreover, why hire an evaluator at all when a well-structured algorithm can draw the same conclusions? We should keep in mind, however, that the phenomenon of Big Data is still very new and we need to address several issues related to privacy, accuracy, reliability, and use. Creating some agreements around use and preventing abuse will be critical to incorporating Big Data into evaluation practices.

**Spotlight: Leveraging Big Data for Global Development**

Global Pulse is an innovation initiative launched by the Executive Office of the United Nations Secretary-General, in response to the need for more timely information to track and monitor the impacts of global and local socio-economic crises. The Global Pulse initiative is exploring how new, digital data sources and real-time analytics technologies can help policymakers understand human well-being and emerging vulnerabilities in real-time, in order to understand the impact of shocks and disasters and how to better protect populations.

Global Pulse has partnered with private sector companies in several areas including telecom, banking, social media, and online search to find actionable information that can help protect vulnerable populations. A collaborative research project between Global Pulse and the SAS Institute, for example, analyzed employment through the use of social media. The analysis revealed the insight that increases in the volume of employment-related social media conversations appeared three months before official increases in unemployment (UN Global Pulse, 2012).

More details can be found at [www.unglobalpulse.org](http://www.unglobalpulse.org)

**Summary**

While each of these approaches is unique, together they represent the six characteristics of *Next Generation Evaluation* in action and embody a profound change in how we conceive of learning and evaluation in the social sector. Some pioneering organizations are beginning to embrace a combination of all three approaches and see benefits through enhanced learning and program effectiveness. Table 2 provides an overall summary of the three game-changing approaches.
Table 2: Summary of the Three “Game-Changing” Approaches

<table>
<thead>
<tr>
<th>Game Changing Approach</th>
<th>Developmental Evaluation</th>
<th>Shared Measurement</th>
<th>Big Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What is it?</strong></td>
<td>A type of evaluation that informs and supports innovative and adaptive development of initiatives in complex, dynamic environments.</td>
<td>A system for reporting the performance, measuring the outcomes, and coordinating the efforts of multiple stakeholders.</td>
<td>The use of datasets that are beyond the ability of traditional database tools to store, manage, or analyze. Big Data is characterized by increased volume, variety, and velocity.</td>
</tr>
<tr>
<td><strong>Why it matters?</strong></td>
<td>Supports the evaluation of innovations where a proven theory-of-change does not exist, allowing for nimble and timely improvements.</td>
<td>Promotes ongoing learning among multiple stakeholders in timely and cost-effective ways.</td>
<td>Allows us to broaden our data set and find connections and correlations.</td>
</tr>
<tr>
<td><strong>What are its benefits?</strong></td>
<td>Enables continuous learning with shortened feedback loops, ensuring that the program or innovation is effective.</td>
<td>Aids more intentional data usage across organizations and provides a basis for engaging policymakers.</td>
<td>Improves the quality of understanding, accountability, and learning about social sector issues.</td>
</tr>
<tr>
<td><strong>What are its downsides?</strong></td>
<td>Potential concerns with objectivity and confidentiality; needs an evaluator with cross-disciplinary skills.</td>
<td>Takes time to build trust-based relationships among participants, as well as resources to support infrastructure.</td>
<td>Provides correlation not causation; entails concerns with privacy and accuracy; requires significant investment in infrastructure and skills to be used effectively.</td>
</tr>
<tr>
<td><strong>What are the challenges with implementing it?</strong></td>
<td>Potentially higher cost; social sector currently focuses on evaluation for accountability (as opposed to learning).</td>
<td>Aligning agendas and resources across multiple stakeholders.</td>
<td>Resource scarcity, digital divide, unique skillset needed for analysis, and capacity to use more broadly.</td>
</tr>
<tr>
<td><strong>Key takeaway</strong></td>
<td>Developmental Evaluation represents an essential and alternative way for evaluating innovations for solving complex social problems.</td>
<td>Shared Measurement can create transparency, accountability, and alignment between stakeholders working on similar issues.</td>
<td>Big Data potentially offers opportunities for making sense of large data sets in understanding social change.</td>
</tr>
</tbody>
</table>
Where Do We Go From Here?

This Learning Brief is merely meant to *start* the conversation about the ways in which evaluation should evolve, expand, and be enlightened if it is to continue to serve as a valuable tool. We have highlighted key trends in the social sector, the six characteristics of *Next Generation Evaluation*, and three emerging game-changing approaches in the field. As we look toward the next steps in the *Next Generation Evaluation* initiative, we invite you to reflect on the following questions:

1. In what ways have you seen the trends we described influence your work?
2. What do you think the implications are for how we think about evaluation? What explicit or implicit assumptions about evaluation are being challenged?
3. To what extent are the characteristics of *Next Generation Evaluation* evident in your work?
4. Which of the game-changing approaches do you find intriguing? Which ones do we need more clarity around as a field?
5. Do you see other game-changing approaches to evaluation emerging?

We welcome your comments and ideas, both before and after the *Next Generation Evaluation* conference. Please send us your thoughts via email to nextgeneval@fsg.org and/or contact us through Twitter @FSGTweets using #nextgeneval. We look forward to hearing from you.
References


List of Interviewees

Tanya Beer, Associate Director, Center for Evaluation Innovation
Anthony Bryk, President, Carnegie Foundation for the Advancement of Teaching
John Cawley, Director of Programs and Operations, J.W. McConnell Family Foundation
Joshua Greenberg, Program Director, Digital Information Technology Program, Alfred P. Sloan Foundation
Alicia Grunow, Senior Managing Partner, Design, Development, and Improvement Research, Carnegie Foundation for the Advancement of Teaching
Ben Hecht, President and CEO, Living Cities
Beth Kanter, Author and Consultant
Robert Kirkpatrick, Director, UN Global Pulse
Paul LeMahieu, Senior Vice President, Programs and Administration, Carnegie Foundation for the Advancement of Teaching
Darin McKeever, Deputy Director, Charitable Sector Support, The Bill & Melinda Gates Foundation
Michael Quinn Patton, Independent Evaluator
Marny Sumrall, Executive Director, YouthTruth
Fay Twersky, Director, Effective Philanthropy Group, The William and Flora Hewlett Foundation
Jennifer Vorse Wilka, Manager, YouthTruth
Geoff Zimmerman, Director of Continuous Improvement, Strive Partnership
FSG

FSG is a nonprofit consulting firm specializing in strategy, evaluation, and research, founded in 2000 as Foundation Strategy Group. Today, FSG works across sectors in every region of the world—partnering with foundations, corporations, nonprofits, and governments—to develop more effective solutions to the world’s most challenging issues.