BRIEFING PAPER

TOPIC 3

How Do We Approach Impact and Evaluation in the Context of Scale?

REFRAMING THE CONVERSATION:
A GEO BRIEFING PAPER SERIES ON GROWING SOCIAL IMPACT

Grantmakers of all kinds are shifting the way they think about scale, emphasizing not size or reach but impact. Growing impact doesn't necessarily require organizational growth or the wholesale replication of programs — it may instead require expanding an idea or innovation, technology or skill, advocacy or policy change. An aspect of scaling impact is about making effective use of what works in a variety of contexts.

For this reason, scaling often entails a mindshift — conceptualizing and working collaboratively through an entire "ecosystem" of key external resources, beneficiaries and environmental conditions.

What should be scaled, why, where and how? From a grantmaker's point of view, evaluation can play an important role not only in identifying what works and why — but also in assessing readiness for, planning and implementing the scaling process itself.

Yet there is often a distinct gap between what we know and what we need to know — between the best available and the best possible evidence — to inform the scaling process. Financial, logistical and other considerations often influence how and whether the learning objectives of an evaluation are met.

This gap has become increasingly apparent within philanthropy as a field of practice, as the ecosystem of evaluation resources, beneficiaries and context is still developing. There is wide variation in the practice, use and dissemination of evaluation in grantmaker and grantee organizations alike. Among grantmakers, for instance, GEO's 2008 triennial field survey found that only 50 percent of staffed foundations formally evaluate their work (with significant differences in practice across foundation size and type).[1] Among grantees, Innovation Network has found that evaluation is the second lowest organizational priority (above only research, which was ranked at the bottom).[2]

That said, grantmakers are increasingly awakening to the power of evaluation as a learning tool to strengthen programmatic and organizational effectiveness, collaboratively designing and creating mutual opportunities for field-wide learning with a variety of stakeholders.[3] This is critical not only for strengthening grantee results, but also for the process of growing what works beyond a single organization or intervention.

---

[1] Grantmakers for Effective Organizations, *Is Grantmaking Getting Smarter? A National Study of Philanthropic Practice* (Washington, D.C.: GEO, 2008), 11.

[2] Ehren Reed and Johanna Morariu, *State of Evaluation 2010: Evaluation Practice and Capacity in the Nonprofit Sector* (Washington, D.C.: Innovation Network, October 2010), 8.

[3] See, for instance, Peter York, *A Funder's Guide to Evaluation: Leveraging Evaluation to Improve Nonprofit Effectiveness* (Washington, D.C.: Fieldstone Alliance and Grantmakers for Effective Organizations, 2005), 103.

Philanthropy has a unique role to play in supporting the evidence base needed to inform any scaling process. Choosing from among the extensive set of analytic approaches and methods used by the evaluation field over many years of practice — including logic models, evaluability assessments, case studies, surveys, quasi-experimental designs, needs assessments and early implementation reviews — depends on what is being scaled, its implementation stage, how results will be used and the kinds of decisions an evaluation is meant to facilitate.[4]

Good evaluation data may appear in atypical or nontraditional ways — and still be sound and appropriate to the purpose at hand. For instance, while methods such as randomized control trials may be sound for proving causality, causality is not what is sought in every case.

Sometimes the question is: Why did this work? Or, for whom did this work — and not work? To get at answers to these questions, different approaches may be needed.

Figuring out which approaches to use and when can be informed by a solid understanding of *where you are in the scaling cycle*. Clarifying *what you seek to learn* — the questions you seek to answer — will drive the evaluation approach, level and type of evidence needed, and help you to avoid seeking more evidence than is needed to inform decisions along the way. And determining *how you intend to use evaluation findings* can play an important role in designing the evaluation — and, particularly important in the context of scale, in facilitating learning across organizational boundaries.

About This Paper Series

This briefing paper is the third topic in a series from GEO's *Scaling What Works* initiative, which will be released throughout 2011. Authored by Dara Major, the collection pulls together the best thinking, research and actionable approaches to scaling impact, as well as provides additional resources for grantmakers that would like to dive deeper into paper concepts and questions. Individual papers delve into topics such as understanding approaches to scale, developing grantee evaluative capacity, structuring and financing scale, building partnerships to grow impact, and more.

To access the latest topic and learn more about *Scaling What Works* and how you can get involved, please visit: www.geofunders.org/scalingwhatworks.aspx.

---

[4] American Evaluation Association, *An Evaluation Roadmap for a More Effective Government* (Fairhaven, Mass.: AEA, September 2010).

EVALUATION IN THE SCALE-UP CYCLE

Evaluation can encompass any activity that informs learning and drives improvement. It may be used to assess results, strengthen implementation and communicate progress toward goals. In practice, good evaluation is a continuous process that begins with planning, is informed by data collection, leads to analysis and reflection, action and improvement — and then begins again.

PLANNING

ACTION & IMPROVEMENT

DATA COLLECTION

ANALYSIS & REFLECTION

Source: Innovation Network. "Building Capacity for Evaluation and Learning" (presented at Grantmakers for Effective Organizations, Washington, D.C., January 24, 2011).

This process is especially important in the context of scaling, a multistage cycle with different learning objectives and levels of evidence needed at each stage. Across this cycle, grantmakers want to learn more than just whether an intervention is a good idea — but how it works, where and under what conditions it works, and how it can be sustained.

In the spring 2010 issue of *The Evaluation Exchange*, Sarah-Kathryn McDonald of the University of Chicago describes the scaling process as the "discovery to implementation cycle," with five clear stages:[5]

▷ *Stage 1* — **Proof of Concept**: Involves determining whether an intervention is sufficiently promising to develop and scale. The goal is to produce evaluation data that can demonstrate which parts of the intervention can accommodate flexibility and which parts are not negotiable. In the context of advocacy or policy change, this may involve looking carefully at the efficacy of a campaign approach or the political context in which an organization is advocating.

▷ *Stage 2* — **Establish Efficacy**: Determines whether the intervention can achieve its intended results under ideal circumstances. At this stage, it is crucial that the intervention be implemented (and evaluated) with the features and in the context that are seen as optimal for success.

▷ *Stage 3* — **Demonstrate Effectiveness**: Aims to assess whether an intervention will achieve its objectives outside the ideal context measured in Stage 2. The objective is to establish whether an intervention "works" in a real-world setting, with all of its complications.

▷ *Stage 4* — **Scale-Up and Test Again**: Aims to demonstrate the intervention's impact when it is implemented among larger numbers of individuals across many contexts. This stage also examines the contextual factors that may influence the intervention's impact in different settings. In addition to identifying reasons for any observed discrepancies, these data can provide feedback to help refine the intervention or develop guidelines to ensure it operates as intended in particular contexts.

5   The model McDonald describes was developed by the Data Research and Development Center, a research and technical center funded by the National Science Foundation as part of the U.S. Interagency Education Research Initiative. McDonald summarizes the model in the spring 2010 issue of *The Evaluation Exchange* and unpacks it more fully in: Sarah-Kathryn McDonald. "Scale-up as a Framework for Intervention, Program, and Policy Evaluation Research" in *Handbook of Education Policy Research,* eds. G. Sykes, N. Schneider and D.N. Plank (New York: Rutledge, 2006), 4 and 191-208.
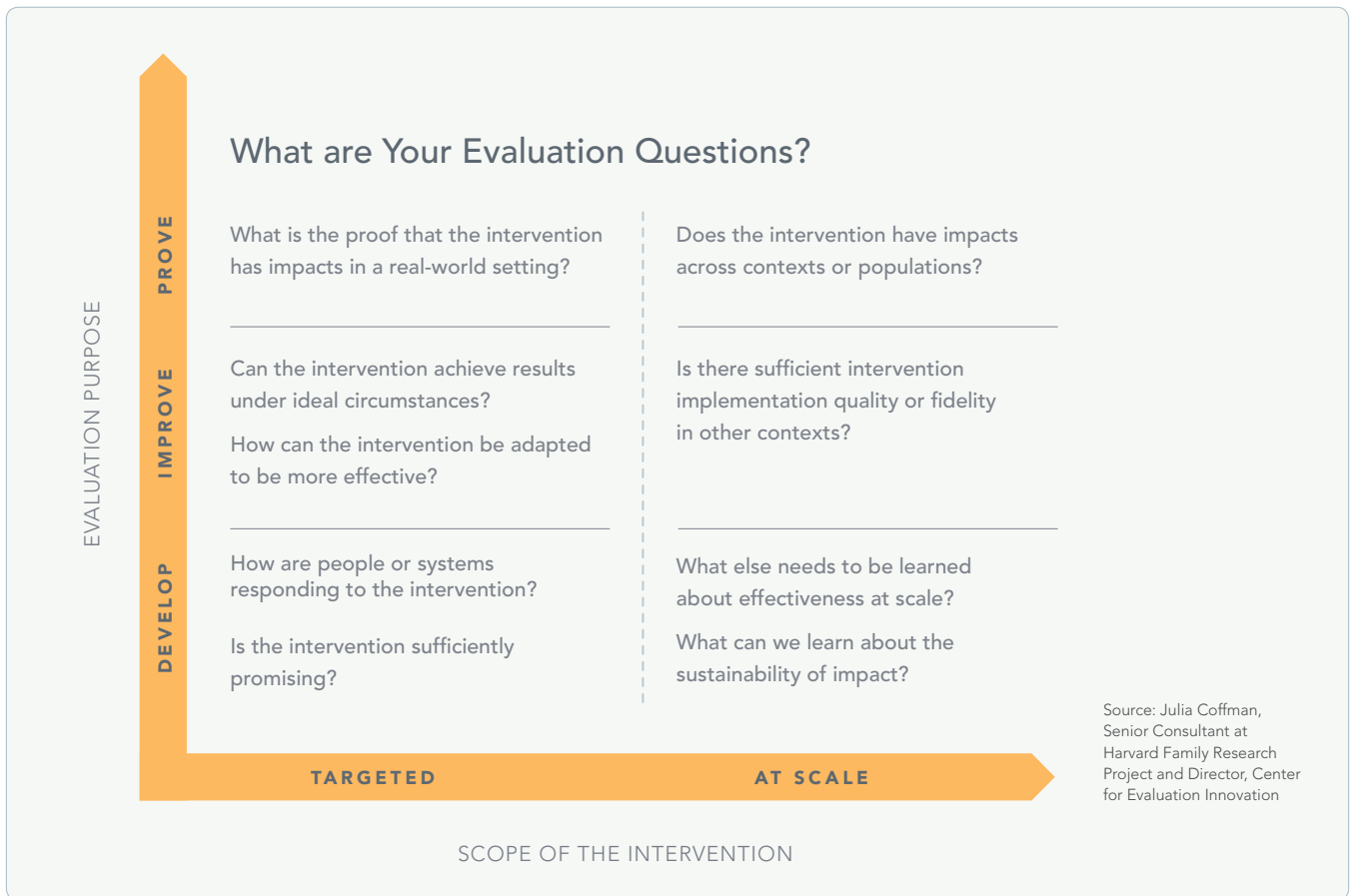
▷ *Stage 5* — **Postmarketing Research**: Explores the following questions: (a) If an intervention demonstrated to work at scale is then more widely adopted, what else do we need to learn about its effectiveness in additional contexts at larger scales? (b) What can and should we seek to learn about the sustainability of its impact and relevance once a market has been saturated and the intervention becomes the status quo?[6]

Evaluation can play a clear role at each of these iterative stages, from designing and developing an intervention or concept to understanding whether and how it worked. Knowing where you are in the scale-up cycle will better position you to identify, at any stage, your evaluation questions, evidence needed and how you intend to use evaluation findings.

IDENTIFYING YOUR EVALUATION QUESTIONS: WHAT DO YOU WANT TO LEARN?

Good evaluations begin with good evaluation questions. Sometimes these questions seem fairly simple — why, what, how? — but in practice they are far from straightforward and often overlooked. Yet these questions form the basis of a deeper inquiry. Do you want to prove a conclusive result? Explain why results were achieved? Why they were consistently achieved across varying contexts? Empower stakeholders? Sustain impact? Inform resource allocation? Where you are in the scale-up cycle affects the evaluation purpose and questions it seeks to answer.

What are Your Evaluation Questions?

EVALUATION PURPOSE

PROVE

What is the proof that the intervention has impacts in a real-world setting?

Does the intervention have impacts across contexts or populations?

IMPROVE

Can the intervention achieve results under ideal circumstances?

How can the intervention be adapted to be more effective?

Is there sufficient intervention implementation quality or fidelity in other contexts?

DEVELOP

How are people or systems responding to the intervention?

Is the intervention sufficiently promising?

What else needs to be learned about effectiveness at scale?

What can we learn about the sustainability of impact?

TARGETED   **AT SCALE**

SCOPE OF THE INTERVENTION

Source: Julia Coffman, Senior Consultant at Harvard Family Research Project and Director, Center for Evaluation Innovation

---

[6] Sarah-Kathyrn McDonald, "Developmental Stages for Evaluating Scale," *The Evaluation Exchange* 15, no. 1 (Spring 2010): 14.

For many grantmakers, strategy — described by Patricia Patrizi as "where an organization is headed and how it intends to get there" — is emerging as a focus of evaluation inquiry, which has implications for the scale-up cycle as well. In practice, however, as the Evaluation Roundtable, a learning network of evaluators in philanthropy, found in its recent survey of major philanthropic foundation evaluators, there seems to be "a general absence of strategy evaluation" and "little is invested in ongoing learning about strategy as it evolves."[7]

"We think it's important to find ways to assess our overall strategy beyond the efforts of individual grantees. Part of our work focuses on diffusing ideas, attracting additional philanthropic capital to support these efforts and encouraging others to adopt them."

Mayur Patel, John S. and James L. Knight Foundation

For instance, Mayur Patel notes, "One of the goals of a matching-fund contest we support, the Knight Community Information Challenge, is to increase the participation of community and place-based foundations in addressing the information needs of their communities. We're investing in the capacity of the foundations to embark on this work and supporting this through shared learning, research, convening and technical assistance."

Formal theories of change and logic models are useful tools for articulating related questions in the context of scale. As Diana Scearce of Monitor Institute has observed, "We like to think about direct and indirect impact: Your theory of change and intended action posits a causal link. But scaling, over time, is about broader impact, and that tends to be indirect."

There is no one-size-fits-all type of evidence; identifying your evaluation question is an important step toward identifying an evaluation approach, type and level of evidence needed.

GETTING CLEAR ABOUT THE TYPE AND LEVEL OF EVIDENCE NEEDED

With evaluation questions in clear view, specifying evaluation approaches is a much less ambiguous task. Three main approaches to evaluation — developmental, formative and summative — can each help to address evaluation questions in the scale-up cycle. (See chart on next page.)
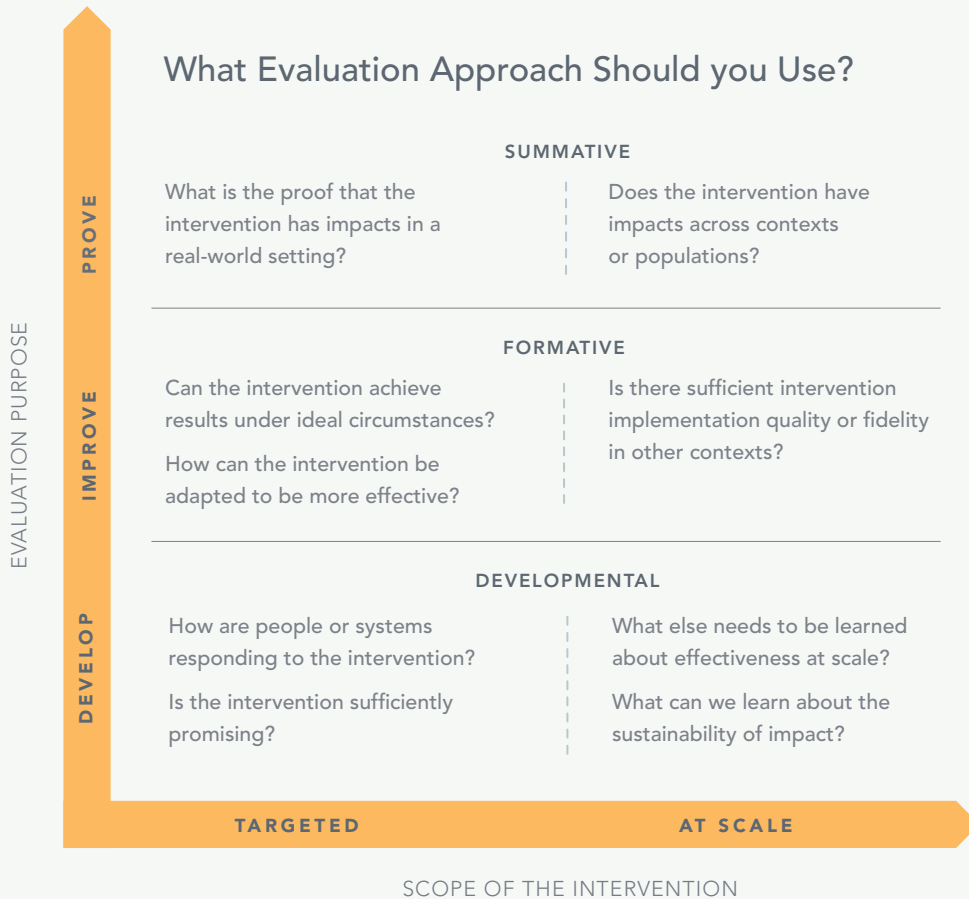
To make scaling efforts possible and successful, grantmakers and grantees alike must carefully consider whether an investment to develop strong evidence is needed and then be prepared to support it as an essential component of growing impact. The key is matching what is feasible to implement with what will meet the learning objectives of the evaluation. Sometimes the best available evidence is preferable to the best possible evidence — especially when considering resource constraints such as timing, budget and organizational capacity.

As Sarah-Kathryn McDonald points out, "[A]pproaching evaluation as a logic game designed to provide the most robust evidence may be a wonderful training exercise for philosophers and methodologists. It may, however, prove disappointing when the success of the evaluation effort is gauged not on the quality of evidence it produces, but by the extent to which it affects subsequent decisions and courses of action."[8]

---

[7] See Patricia A. Patrizi and Michael Quinn Patton, "Evaluating Strategy," *New Directions for Evaluation*, 128 (Winter, 2010): 9-10; and Patricia A. Patrizi, "Strategy Evaluation: Emerging Processes and Methods," *New Directions for Evaluation*, 128 (Winter, 2010): 87-102.

[8] Sarah-Kathryn McDonald, "Scale-up as a Framework for Intervention, Program, and Policy Evaluation Research," 193.

What Evaluation Approach Should you Use?

EVALUATION PURPOSE

PROVE

SUMMATIVE

What is the proof that the intervention has impacts in a real-world setting?

Does the intervention have impacts across contexts or populations?

IMPROVE

FORMATIVE

Can the intervention achieve results under ideal circumstances?

How can the intervention be adapted to be more effective?

Is there sufficient intervention implementation quality or fidelity in other contexts?

DEVELOP

DEVELOPMENTAL

How are people or systems responding to the intervention?

Is the intervention sufficiently promising?

What else needs to be learned about effectiveness at scale?

What can we learn about the sustainability of impact?

TARGETED   **AT SCALE**

SCOPE OF THE INTERVENTION

Developmental evaluation positions the evaluation as a part of a program's design and development process. The evaluator collects information and provides informal feedback to members of the design or planning team to help them refine their approach before it is implemented.

Formative evaluation is conducted during implementation, for the purpose of improvement. It can help to strengthen understanding of the implementation process itself and helps to pave the way for a later, summative evaluation.

Summative evaluation is implemented for the purpose of determining the merit, worth or value of an intervention for the purpose of concluding whether that effort should be continued, discontinued and/or scaled up.

Source: Julia Coffman, Senior Consultant at Harvard Family Research Project and Director, Center for Evaluation Innovation

USING EVALUATION FINDINGS

Ultimately, what benefits will be realized through the evaluation findings and by whom, including community, grantee, grantmaker or policymaker?

Each stage in the scaling process will suggest a different standard of use. For instance, findings on how an innovation described in research actually worked in a limited setting can be used to inform the scaling process itself. Grantmakers could also support grantees' efforts to base their work on best available evidence that currently exists. Designing evaluations and sharing findings with key stakeholders is essential in ongoing scaling processes.

In philanthropy, however, the approach to evaluation is often fragmented, funding processes are overly complex or create unnecessary restrictions and grantmakers and grantees rarely design and use evaluation to jointly grow impact. For instance, GEO has found that grantmakers conduct evaluations primarily for internal audiences (88 percent for staff and 78 percent for their board); only 31 percent viewed grantee organizations as a primary beneficiary of results, and just 10 percent viewed "other grantmakers" as a primary intended audience.[9] The Center for Effective Philanthropy has found that grantees on average do not perceive grantmaker reporting or evaluation practices to be very helpful and approximately half have not even discussed their reports or evaluations with foundation staff or external evaluators.[10]

Grantmakers who do work closely with grantees to design and implement evaluation, however, put their findings to use in meaningful ways. To inform a multi-site scaling process, Jackie Williams Kaye, director of research and evaluation at Wellspring Advisors, formerly of Atlantic Philanthropies, described working in close partnership with Higher Achievement — a Washington, D.C.-based nonprofit that develops academic skills, behaviors and attitudes in academically motivated and underserved

middle school children — to prototype the scaling process while an impact evaluation was under way. "Higher Achievement approached Atlantic Philanthropies with a growth plan to move quickly from their local base to multiple sites around the country. We encouraged them to consider how much scale should happen when, given where they were currently in an evaluation process. Some funders might say: don't attempt scale before conducting a full impact evaluation. But we felt that two things are essential: knowledge that you're producing the benefits you want to produce and experience in implementing in a different context. So we supported their move to an additional site to help grow their evidence base." As Higher Achievement's CEO Richard Tagle notes, "Through this process, we're synthesizing rigorous data about what we want to scale with rigorous data on the scaling process itself. It's a powerful, multifaceted learning process."

Grantmakers could also work closely with other grantmakers to design and implement evaluation for the purpose of building knowledge about not just the discrete results of a single action, but the underlying mechanisms that make it effective in a variety of settings. Hallie Preskill, executive director of FSG's Strategic Learning and Evaluation Center notes, "Our point of view is that the best way to make evaluation relevant, credible and widely useful is to make it collaborative by involving a range of stakeholders. We begin new client engagements by consulting with a group of people and then creating a systems map. Let's figure out the context of this work, what's the issue you are trying to affect, who are the actors, who else is working in this sphere and then see yourselves in this system. I'm intrigued by the growing potential to scale learning among groups of grantmakers in this way."

Grantmakers seeking to facilitate evaluative learning *across* organizations should also give careful consideration to how evaluation findings will be used and make smart use of intermediaries to help guide collaborative knowledge sharing. Over time, this would increase the scope and quality of evaluative data available to the field.[11]

---

[9] Grantmakers for Effective Organizations and the Council on Foundations, *Evaluation in Philanthropy: Perspectives from the Field* (Washington, D.C.: 2009), 23.
[10] Ellie Buteau and Timothy Chu, "Grantees Report Back: Helpful Reporting and Evaluation Processes," Center for Effective Philanthropy, January 2011, 1-2, http://www.effectivephilanthropy.org/assets/pdfs/Data%20in%20Action//CEP_DatainAction_GranteesReportBack.pdf.
[11] For perspectives on improving the flow of nonprofit data, see Maisie O'Flanagan et al., "The Nonprofit Marketplace: Bridging the Information Gap in Philanthropy" (report for William and Flora Hewlett Foundation, Menlo Park, Calif., 2008); and Cinthia Schuman Ottinger, "A Dearth of Data Hides Nonprofit World's Impact," *Chronicle of Philanthropy*, February 7, 2010.

The potential for achieving this level of evaluation design and use ultimately resides among what Peter Senge of the MIT Sloan School of Management called "communities of commitment," noting that, "Without commitment, the hard work required will never be done. People will just keep asking for 'examples of learning organizations' rather than seeking what they can do to build such organizations." Yet, "Without communities of people genuinely committed, there is no real chance of going forward."[12]

In short, a more expansive view of evaluative learning within as well as beyond individual organizations may lead to more grantmakers working across organizational boundaries to harness complementary capabilities for improved use of evaluation. Learning organizations, as Marilyn Darling of Signet Research and Consulting has noted, are places of curiosity and humility: "People in these organizations talk a lot about their thinking. They are honest about acknowledging and learning from disappointing results; conscious about testing out theories of change. Leaders make learning a priority. They maintain a clear focus and seek deep alignment."[13]

Going forward, notes Mary Vallier-Kaplan, vice president and COO of the Endowment for Health, "Scale in the nonprofit sector is never exclusively about the grantee, nor even a program — sometimes the grantmaker plays a central or catalytic role in the scale-up cycle. If grantmakers really want to grow impact, well beyond the capacity of a single foundation, we have to recognize that we are all part of a larger system and can create new value by putting our wealth of resources to work across that system."

WANT TO BE A GRANTMAKER THAT SUCCESSFULLY INTEGRATES EVALUATION INTO YOUR EFFORTS TO GROW IMPACT?

From a grantmaker's point of view, evaluation can play an important role not only in identifying what works and why — but also in assessing readiness for, planning and implementing the scaling process itself. Consider the following approaches:

1. ***Determine where in the scale-up cycle you are*** — The stages in this cycle include: demonstrating proof of concept; establishing efficacy; demonstrating effectiveness; scaling up and testing again; and post-marketing research.

2. ***Articulate your evaluation questions*** — What do you want to learn? Good evaluations begin with good evaluation questions. Sometimes these questions seem fairly simple — why, what, how? — but in practice they are far from straightforward and often overlooked.

3. ***Get clear about the type and level of evidence you need*** — Choosing from among an extensive set of analytic approaches and methods depends on what is being scaled, its implementation stage, how results will be used, available resources and the kinds of decisions an evaluation is meant to facilitate.

4. ***Consider how you will use evaluation findings*** — This may include using findings to support a scaling process or to build knowledge about not just the discrete results of a single action but the underlying mechanisms that make it effective in a variety of settings. Carefully consider whether your goal is to facilitate learning across grantmaking organizations and build that into your evaluation design up-front.

---

[12] Fred Kofman and Peter M. Senge, "Communities of Commitment: The Heart of Learning Organizations," *Organizational Dynamics*, 22, no. 2 (1993), 5-23.
[13] Marilyn J. Darling, *A Compass in the Woods: Learning through Grantmaking to Improve Impact* (Boston, Mass.: Signet Research and Consulting LLC, 2010), 7. Signet Research and Consulting has created a learning practices self-assessment tool, available at http://www.emergentlearning.com.

About *Scaling What Works*

Launched in 2010, *Scaling What Works* is a multiyear learning initiative of Grantmakers for Effective Organizations, a thought leader for promoting grantee-centric philanthropic practices that lead to more effective results. With the support of a coalition of 22 funders, GEO aims to expand the number of grantmakers and public sector funders that are working together to broaden the impact of high-performing nonprofits. Through *Scaling What Works,* GEO will offer training, networking opportunities and a host of tools and resources, such as this paper series, to better equip grantmakers to help the nonprofit organizations they support to plan, adapt and grow their impact in creating sustainable benefits for people, their communities and our planet.

For more about GEO and *Scaling What Works*, please visit our Web site at www.geofunders.org.

GEO THANKS THE FOLLOWING INDIVIDUALS FOR THEIR FEEDBACK ON THIS PAPER:

Ellie Buteau, Center for Effective Philanthropy

. . .

Julia Coffman, Center for Evaluation Innovation

. . .

Jackie Williams Kaye, Wellspring Advisors

. . .

Mayur Patel, John S. and James L. Knight Foundation

. . .

Hallie Preskill, Strategic Learning and Evaluation Center, FSG

. . .

Lisbeth B. Schorr, Center for the Study of Social Policy, Harvard University

. . .

Chris Tebben, Grantmakers for Education

. . .

ADDITIONAL RESOURCES ON IMPACT AND EVALUATION:

Kania, John and Mark Kramer. "Collective Impact." *Stanford Social Innovation Review* 9, no. 1 (Winter 2011): 36-41.

. . .

Mark Kramer et al. *From Insight to Action: New Directions in Foundation Evaluation.* Boston: FSG, June 2007.

. . .

Mark Kramer et al. *Breakthroughs in Shared Measurement and Social Impact.* Boston: FSG, July 2009.

. . .

Rosemary Hattrup McNelis, "Building Formal Knowledge Bases: Understanding Evaluation Use in the Foundation Community," *American Journal of Evaluation*, 17, no. 1 (Feb. 1996): 1 and 19-41.

. . .