The Foundation Review

Volume 11 | Issue 3

9-2019

Scaling Programs With Research Evidence and Effectiveness (SPREE)

Nan Maxwell
Mathematica

Scott Richman
Mathematica

Recommended Citation

Maxwell, N., & Richman, S. (2019). Scaling Programs With Research Evidence and Effectiveness (SPREE). *The Foundation Review, 11*(3). https://doi.org/10.9707/1944-5660.1481

doi: 10.9707/1944-5660.1481

Scaling Programs With Research Evidence and Effectiveness (SPREE)

Nan L. Maxwell, Ph.D., and Scott B. Richman, Ph.D., Mathematica

Keywords: *Implementation science, scaling, evidence, evaluation*

Introduction

Funders increasingly use evidence to select practices and programs that can best address individual and community needs. Evidence can also play a role in replicating the effects of these practices and programs, so that foundations can serve more people and increase their reach. To support effective scaling, funders need a comprehensive methodology for identifying effective interventions and assessing the readiness of the interventions and implementing organizations for scaling (Miller, Sorensen, Selzer, & Brigham, 2006; National Implementation Research Network [NIRN], 2018).

This article describes a process called SPREE — Scaling Programs with Research Evidence and Effectiveness — and provides insights into conditions under which foundations can apply it to help them and their grantees scale successfully. Implementing SPREE can assist foundations in two ways: (1) using evaluation research as a tool to determine which interventions are likely to produce desired outcomes, and (2) identifying those organizations ready to scale them. The insights and lessons discussed here are derived from the experiences of the Corporation for National and Community Service (CNCS), a federal grantmaking agency, in applying the process.

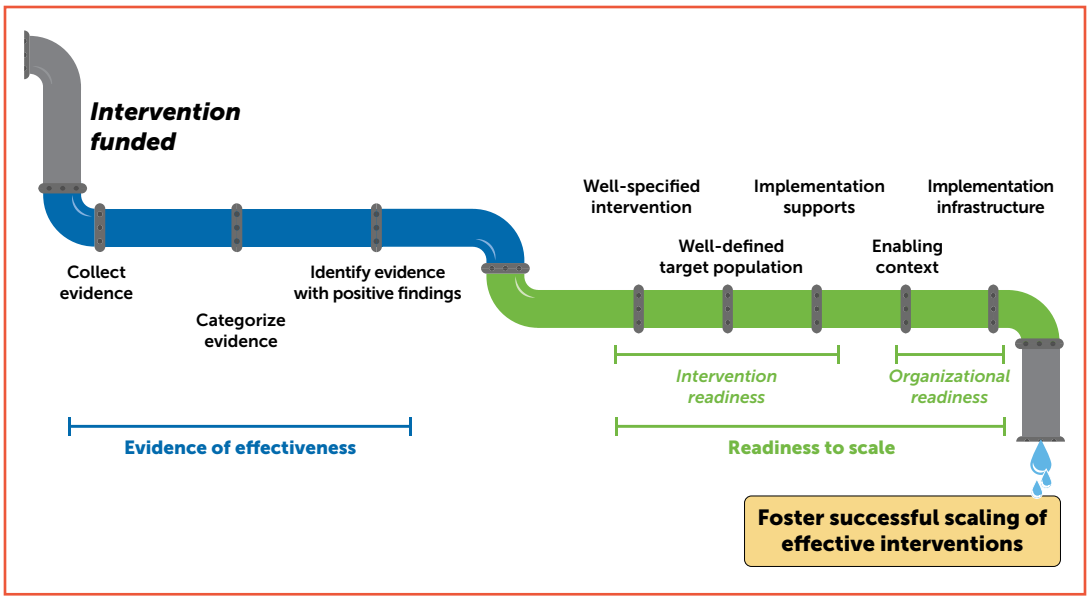
Key Points

- Foundations can serve more people by identifying and supporting effective interventions that are ready to be scaled. This article describes a process called SPREE — Scaling Programs with Research Evidence and Effectiveness — that can help funders and their grantees scale successfully. Implementing this process can assist foundations in using evaluation research as a tool to determine which interventions are likely to produce desired outcomes, and to identify which organizations are ready to scale them.
- The SPREE process is grounded in evaluation and implementation science frameworks and has been applied since 2016 by the Corporation for National and Community Service. This article explores how the agency's application of the process helps it ensure that the interventions it funds are likely to improve outcomes and extend its reach through successful scaling. In addition, the process generated discussions about using evidence and readiness to scale to guide funding decisions.
- While the SPREE process might work best when foundations and the grantees they fund have a culture of measurement, learning and evaluation, the process itself can be used to help them build or strengthen that culture. It can also help funders identify and provide the kind of support grantees need in demonstrating that an intervention is effective and in building the conditions needed to scale it successfully.

The SPREE Process

Program managers can make informed decisions by incorporating measurement, learning, and evaluation into their strategic planning. Developing an inventory of currently funded interventions, requiring grantees to demonstrate evidence of an intervention's effectiveness, and using evidence requirements to structure contracts and grants can ensure that a foundation's funding is directed toward interventions most

FIGURE 1 Scaling Programs With Research Evidence and Effectiveness – The SPREE Process



likely to achieve desired outcomes among specific target populations (Pew Charitable Trusts, 2016). Similarly, foundations can use research from implementation science about intervention and organizational readiness for scaling to expand their reach.

The two-part SPREE process aims to help foundations identify which of their funded interventions can be scaled successfully. The first part of the process helps foundations identify the interventions that are most likely to achieve desired outcomes; the second part helps them identify which of those effective interventions demonstrate a readiness for scaling and which organizations might be ready to scale them. (See Figure 1.)

Identifying Effective Interventions

The availability of rigorous research on the effectiveness of social programs has increased dramatically over the past decade. Most prominently, three federal research clearinghouses are providing information about interventions to help policymakers and program managers identify effective interventions: The Department of Education's What Works Clearinghouse[1] (WWC) reviews research to determine which education interventions are effective; the Labor Department's Clearinghouse for Labor Evaluation and Research[2] (CLEAR) reviews studies for their ability to establish a causal impact for an intervention; and the Department of Health and Human Services' Home Visiting Evidence of Effectiveness[3] (HomVEE) project reviews research on home-visiting models to identify effective interventions for pregnant women or families with children from birth to kindergarten. Still, foundations often make funding decisions without looking at the evidence of an intervention's effectiveness. The first part of the SPREE process includes three steps a funder can take to identify an intervention's effectiveness so that information can be used in decision-making. (See Figure 2.)

[1] https://ies.ed.gov/ncee/wwc

[2] https://clear.dol.gov

[3] https://homvee.acf.hhs.gov

FIGURE 2 Identifying Effective Interventions

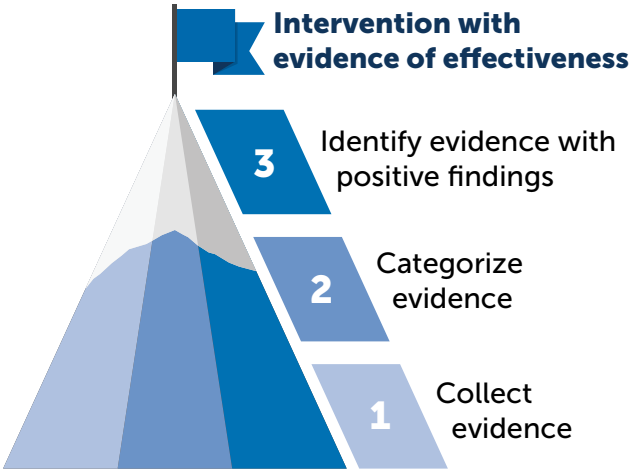


TABLE 1 Clearinghouse Standards Frequently Used to Identify Effective Interventions

Reporting on Methods	The study includes adequate information about the research design and statistical approach to gauge impacts.
Evaluator Independence	The evaluator was external to the grantee to ensure independence in findings.
Study Design	Research contains a comparison group, ideally with members assigned randomly. In addition, the study has:
	Low attrition: Few people in the treatment or comparison group who left the study.
	No reassignment: No people randomly assigned to comparison group switched to the treatment group and vice versa.
	Baseline equivalence: People in the treatment and comparison groups in the analytic sample did not differ at the start of the study.
	No confounding factors: The design precluded factors other than the intervention from producing outcomes.

1. *Collect evidence.* To identify effective interventions, foundations need to compile a comprehensive inventory of funded programs and the evaluation research for each one. This inventory should include a description of each program, its goals, the target population, the number of participants served, and the research providing evidence of the program's effectiveness.

2. *Categorize evidence.* Because the quality of the research may vary, foundations need to define standards to demonstrate that the effects estimated can be attributed solely to the intervention. (See Table 1.) For foundations that lack the staff to develop and apply such standards, research clearinghouses are a useful source. For example, a foundation funding a college and career intervention

FIGURE 3 Summary Ratings of Evidence

- A **high rating** indicates confidence that the intervention caused the desired outcomes.
- A **moderate rating** indicates some confidence that the intervention produced the outcomes, but that other contributing factors might have also intervened.
- A **low rating** indicates little confidence that the intervention produced desired outcomes, because other factors likely contributed.

that is structured like a career academy could use CLEAR and the WWC to find research on whether career academies have been shown to be effective. Because the clearinghouses provide summary ratings of evidence for an intervention's effectiveness, the foundation can compile summary ratings for interventions it funds to help it assess the level of confidence in the effectiveness of each intervention. (See Figure 3.)

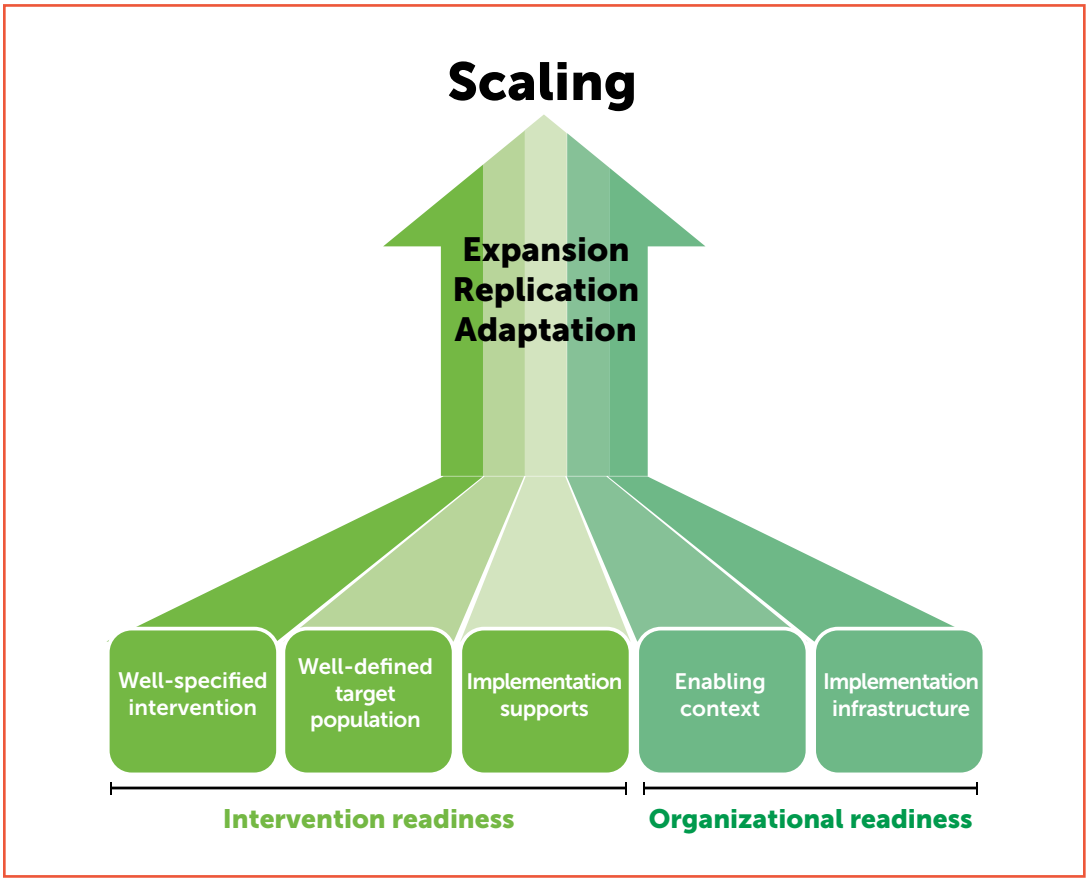
3. *Define evidence of positive outcomes.* For evidence to support confidence in an intervention's outcomes, it is not necessarily the case that all evaluations of the intervention show a positive impact or that expected benefits exceed costs. Each funder must define what evidence is adequate to consider an intervention effective. One evaluation showing a positive causal relationship on at least one outcome might be adequate evidence for one foundation, for example, while another might require that most evaluations show such an impact on the majority of outcomes examined or that one evaluation indicates that an intervention's benefits shown through causal evidence outweigh its costs.

Since not all interventions will have been researched for their effectiveness, foundations themselves may have to make those assessments. But standards that are too rigid might lead funders to discard potentially effective interventions that have not yet been able to establish such evidence. Accurate impact measurement can be difficult for some types of outcomes or in work with specific target populations.

An evaluation of a program that attempts to reduce drug use, for example, faces the often difficult challenge of locating people for whom the intervention did not work; as a result, the evaluation might overstate the program's effectiveness because the study could not fully administer post-intervention surveys among those participants. Or while randomly assigning participants into either a treatment group that receives the intervention or a comparison group that does not is the gold standard for evaluation research, circumstances might not allow for random assignment. Legislation might mandate that members of a certain group receive an intervention, thereby precluding their assignment to a group that does not receive it; or ethical concerns about withholding services from those who need them for the sake of research might prevent an organization from using random assignment. Insufficient resources might also be a barrier to evaluation. High-quality evaluation of an intervention can entail considerable costs that might rule out an evaluation altogether, or lead to less rigorous or poorly implemented research — adequate funding may not be available, for example, to train staff about specific evaluation tasks (Despard, 2016; Gondolf, 2015).

Such limits on the accurate assessment an intervention's effectiveness are not inconsistent with the SPREE process. The process does not dictate that only effective interventions be considered for their scaling potential; it merely highlights how scaling effective interventions enhances the

FIGURE 4 Conditions for Successful Scaling



probability that a funder will be able to improve lives of more people. Furthermore, foundations and other mission-driven organizations might embrace values other than participant outcomes when assessing which interventions to scale. Expanding diversity, inclusion, and equity; investing in new or innovative programs and practices; and supporting a particular practice or program (e.g., community service and volunteering) are all goals that funders might want to emphasize when deciding which interventions to scale.

Identifying Interventions and Organizations Ready for Scaling

Funding and implementing effective interventions increase the likelihood of improving participants' lives. Scaling takes implementation to the next step; the focus goes beyond executing an effective intervention to replicating the same effects for a greater number of people. The SPREE process was developed to assess readiness for three types of scaling. The first type is expansion, or extending an intervention to more people in the same target population and location, and requires increasing the capacity of an existing infrastructure. The second type is replication, or extending an intervention to the same target population but in a new location, and requires a new implementation infrastructure. The third type of scaling is adaptation — modifying an existing intervention to serve a new target population or to implement it in a new setting while adhering to the intervention's intentions.

The SPREE process identifies five conditions indicating that both an intervention and the organization implementing it are ready for successful scaling. Successful scaling means that the intervention is implemented with fidelity — as it was intended — after it is adapted to serve a larger number of people.[4] (See Figure 4.) Both fidelity and effectiveness often flounder during scaling as capacity increases and adjustments are made (Larson, Dearing, & Backer, 2017). Maintaining fidelity to the intervention model after scaling helps ensure the intervention will continue to generate its beneficial outcomes.

The first three conditions for successful scaling indicate whether the intervention has the features that will allow it to be implemented with fidelity after scaling:

- *A well-specified intervention* clearly identifies the core set of elements critical to achieving beneficial outcomes (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005). Each element must describe what it takes to produce the intended outcomes, including the intervention's content (e.g., activities or services); how it is to be delivered; how much of the intervention participants should receive; the requirements for staff delivering the intervention; and the setting in which the intervention will take place (Blase & Fixsen, 2013). These elements provide structure that ensures the intervention is delivered with fidelity and consistency; without those elements, it is less likely that the intervention will improve participants' outcomes to the extent expected given the intervention's success before it was scaled.

- *A clearly defined target population* ensures that the organization is offering the intervention to those for whom it was designed and shown to be effective. This definition must specify the characteristics necessary for people to participate in the intervention (Garg, 2016; McElroy & Ladner, 2014); if the organization intends to serve a new population, that definition should be adapted accordingly.

- *Implementation supports* must be in place. They include a monitoring team that ensures the intervention is implemented as intended, continuous quality-improvement processes, and pre-service and in-service staff training (Breitenstein et al., 2010).

Even if the intervention is ready for scaling, the organization must be able to support the scaling for it to be successful. This means that the organization must have an environment that is conducive to scaling and have supports in place to ensure the scaled intervention's success. The final two conditions indicate an organization is ready to scale an intervention:

- *An enabling context* must be present: The organization's leadership and culture must support innovation, learning, and improvement. This support is necessary for the creation of an environment hospitable to the implementation of effective interventions and the use of effective implementation supports for staff. Although an organization's enabling context develops in different ways, having successfully tackled challenges in the past is one way such a context can develop. The organization's structures, roles, and functions should facilitate, rather than hinder, service delivery and its ability to affect beneficial outcomes.

- *A solid implementation infrastructure* must exist. An organization's infrastructure must contain sufficient financial, human, and physical resources to support the intervention (Bernfeld, 2006; Fixsen, 2009; Klingner, Ahwee, Pilonieta, & Menendez, 2003) and its successful implementation (Mihalic & Irwin, 2003) after scaling. To effectively

[4] Other frameworks also provide guidance in scaling an intervention. Although many are, like SPREE, broadly focused in implementation science (e.g., Achieving the Dream, 2011; Barker, Reid, & Schall, 2016), they lack the simplicity that allows a funder to easily capture a readiness for scaling (e.g., Cooley, Ved, & Fehlenberg, 2012). Still other frameworks are more narrowly focused. For example, Meehan and Jonker's (2018) readiness-to-scale matrix uses a management perspective to focus on an organization's readiness without considering that of the intervention.

TABLE 2 CNCS Tiered Evidence Ratings

Rating	The Evidence
Strong	Supports causal conclusions that assess the intervention nationally, regionally, or at the state level.
Moderate	Supports causal conclusions but has limited generalizability beyond the study context.
Preliminary	Is based on an outcome study with no comparison group.
Pre-Preliminary	Has some data collection and data.

support the scaled intervention, the organization's infrastructure must enable it to supply the new staff necessary for scaling; support hiring, supervision, and staff development through a human resources management system; engage in continuous quality-assurance processes; and provide funding and other resources (e.g., materials, physical space). Of note, the infrastructure could include resources external to the organization: For example, if partners play a key role in implementation, their policies, priorities, systems, and so forth must also support successful scaling of the intervention.

CNCS: A Case Study in Applying the Process

The Corporation for National and Community Service is the nation's largest grantmaker for national service and volunteering. By funding programs such as AmeriCorps State and National, VISTA (Volunteers in Service to America), and Senior Corps, it enables thousands of Americans to effect change in their communities through interventions in economic opportunity, education, disaster services, environmental stewardship, healthy futures, organizational capacity building, and support for veterans and military families. The CNCS and its grantees also invest significant resources in evaluating the effectiveness of these interventions. Because of the diverse nature of its programs and in their expected outcomes — including impacts on increased literacy and education attainment, employment, career growth in volunteers, and conserving natural resources — the CNCS has applied the SPREE process since 2016 to determine how to identify effective interventions and decide which of those to scale.

Laying the Groundwork

Since its inception in 1990, the CNCS has assessed the programs it funds by holding grantees accountable to performance measures. Starting in 2010 with the launch of the Social Innovation Fund program, the agency began to more systematically organize and develop the evidence base for its programs. These efforts included (1) developing a tiered evidence-rating framework to assess the quality and strength of evidence underlying the impact of the interventions the agency supports, and (2) establishing tiered evaluation requirements for grantees. (See Table 2.) The CNCS contracted with independent, third-party evaluators to review and apply the appropriate evidence rating to documentation submitted by grantees addressing the effectiveness of proposed interventions. Through an iterative process, a body of evidence on the programs the agency supports emerged, and the agency conducted a number of meta-synthesis and meta-analysis studies to determine areas of strength, weakness, and growth concerning target outcomes. This evidence base would not

have emerged without grantee investments and efforts to capture data on their programs and the CNCS's efforts in compiling, categorizing, and making meaning of those data.

In developing guidelines for its grantees, the CNCS wanted to bring more uniformity, strategic learning, and a focused vision to evaluate the range of evidence frameworks and metrics for determining what constitutes an effective intervention applicable to its programs. In 2016, it began a multiyear effort to deepen its understanding of the interventions it supports and to build its knowledge base on scaling them. Its vision was to leverage its investments by ensuring that its most effective interventions could be scaled to engage more people and communities across the country.

Implementing the Process

The CNCS selected the SPREE process as the vehicle to further its thinking on using evidence in funding and scaling. Working with a contractor, it completed four main tasks. First, it compiled research from grantees the agency previously rated as having moderate or strong evidence (Richman, Maxwell, Streke, Needels, & Eddins, 2018). Next, it used standards set by federal research clearinghouses to develop its own standards to categorize research; these went beyond the agency's tiered evidence ratings. In its third task, the CNCS defined an effective intervention as having at least one study that showed a positive impact in research meeting these standards. Lastly, the CNCS applied the SPREE's scaling framework to determine whether an intervention and organization implementing it were ready for scaling (Needels, Selekman, Jones, Richman, & Maxwell, 2018).

The CNCS contractor applied the SPREE process by developing and applying a rubric to extract information about the research's ability to provide evidence that the intervention leads to participant outcomes and evidence of the intervention's and organization's scaling readiness. The rubric served two key purposes: to enable the contractor to systematically review the research, to determine what met the standards for effective intervention; and the scaling plan documents, to determine whether the intervention and organization met SPREE's five conditions for scaling readiness. It is important to note that the contractor applied the rubric to evidence and scaling documents that grantees had already developed and submitted to the agency based on the agency's existing reporting requirements; the SPREE process was applied to these documents after the fact. As a result, the CNCS case study provides an example of benefits the SPREE process might provide in the absence of an ideal set of information to feed into it.

Results

Applying the SPREE process helped the CNCS understand which of its funded interventions are likely to be effective, and which of those effective interventions and the organizations implementing them might be ready for scaling. The process accomplished the following:

First, it identified the primary reasons why an intervention did not meet the standards set for effectiveness: the evidence that could establish whether the program produced desired outcomes did not consistently provide favorable results and the evidence could not establish that the intervention produced the desired outcomes. The latter finding was not necessarily surprising given the variety of programs the agencies offered. Programs were subject to different requirements for producing evidence and had different expectations for outcomes, with some prioritizing community service and career growth among volunteers over participant outcomes.

The process also highlighted the need for more detailed and structured information from grantees about their readiness to scale an intervention. Because scaling documents were developed before the CNCS adopted the SPREE process, information provided was not always specific enough to assess readiness, the criteria for which are now clarified through the SPREE process.

The SPREE process also fostered conversations about the desire to incorporate evidence in decision-making and scaling. (See Figure 5.) It spurred discussion on how best to use evidence as a basis for funding intervention scaling and

FIGURE 5 Promoting Discussion

The SPREE Process Stimulated Discussion About:

- research standards an agency should embrace;
- assistance an agency can provide to help grantees provide evidence of their intervention's effectiveness;
- assistance an agency can provide to ready grantees to scale an intervention;
- how an agency might reconcile differences between an intervention research found to be effective and a greatly modified version of that intervention a grantee proposes for scaling; and
- how the agency might retain its ability to fund innovative programs while also stressing the need to show an intervention to be effective.

support grantees in documenting the potential effectiveness of their interventions. Such support might include, for example, helping grantees understand what constitutes evidence of an intervention's effectiveness and what it takes to be ready to scale an intervention.

The knowledge and discussions resulting from applying the SPREE process helped the CNCS identify the following imperatives:

1. Build an agency consensus about appropriate standards for research evaluations and what constitutes readiness to scale an intervention;

2. Modify application and reporting requirements to ensure applicants and grantees fully understand the reasons for an intervention's effectiveness, provide evidence of the outcomes, and clearly demonstrate their readiness for scaling; and

3. Support grantees in their efforts to build capacity in evaluating and scaling interventions.

Insights

In addition to helping foundations ensure that the interventions they fund are likely to improve outcomes and reach more people through successful scaling, the SPREE process can generate much-needed discussions about using evidence and readiness to scale to guide funding decisions. The CNCS's application of SPREE highlighted these benefits as well as three conditions that could maximize its use.

A Learning Culture

A funder is best positioned to build research evidence and use it to make decisions if it has a culture of measurement, learning, and evaluation. Such a culture requires foundation leadership, management, and staff to develop a common understanding about the value of measurement and evaluation in decision-making and to agree on what constitutes evidence of an effective intervention (Austin & Claassen, 2008a).

Such a culture also strengthens grantees. Although some grantees might have an established culture of learning that includes measurement and evaluation, others might require a cultural change to accommodate a foundation's evidence-based decision-making. For those grantees, foundations would be wise to demonstrate the value of measurement and evaluation over time, rather than mandating their use in the short term (Walker & Soule, 2017). Grantees might need time to see that a high-quality evaluation that examines inputs, processes, outputs, and impacts can provide them with both formative feedback that informs successful implementation and summative findings about

the intervention's effectiveness. Together, that knowledge can be a powerful tool for improving intervention design when grantees use the results to examine the values and assumptions underlying a program. It is therefore important that foundations provide grantees with funding or other support to help them understand how evaluations can be used for improvement, and not as a "thumbs up/thumbs down" decision about whether to continue an intervention (Austin & Claassen, 2008b).

Support for Conducting Evaluations

In addition to a culture that values measurement, learning, and evaluation, a grantee might need additional evaluation-related supports to provide evidence of an intervention's effectiveness.

Funding for evaluation research is one such support. High-quality evaluations require financial resources. The CNCS found that evaluations that can provide evidence of effectiveness tend to cost 15% to 20% of a grant's budget for a small-scale evaluation and 25% or more for a large-scale evaluation (Zandniapour & Vicinanza, 2013). Evaluation costs include implementation expenses as well as the fees for experts in research design and implementation and for those who can distinguish two key types of research: evaluation research, which seeks to improve a program or intervention, and basic research, which seeks to test a hypothesis. Offering technical assistance to grantees, such as teaching them how to work with an evaluator to provide rigorous evidence, is another effective form of support.

Grantees should also know what constitutes a high-quality evaluation. Meaningful information can ensure a common understanding of the value of intervention evidence. Foundations can help grantees and their third-party evaluators improve the quality of evidence that shows the effectiveness of their interventions by using guidance materials developed by research clearinghouses. Such materials might be especially useful if used in conjunction with discussions about the challenges grantees may face in conducting rigorous evaluations of impact.

Funders can also help grantees in selecting an appropriate evaluator. Sometimes grantees do not understand that the greater objectivity of third-party evaluators leaves their studies — as opposed to those conducted by staff — in a better position to provide stronger evidence of an intervention's effectiveness. Foundations can help grantees see how a third-party evaluation complements the measurement, evaluation, and learning that their internal staff undertake every day. For example, during the evaluation design phase, grantees will work with evaluators on three key tasks: First, they will clarify the intervention's theory of action so that evaluators understand the indicators of inputs, processes, outputs, and outcomes that are important to track. Second, they will ensure the evaluation addresses all elements in the theory of action. Finally, they will create processes to translate information from the evaluation into organizational learning and improvement. During the implementation phase, grantees will work with evaluators to make sure the tasks are carried out as planned.

Even when grantees do realize the benefits of a third-party evaluation, they might not have the staff with sufficient expertise to select an appropriate evaluator. Because not all interventions are at a stage where their effectiveness can be accurately determined, the characteristics an evaluator requires will vary. Foundations can help grantees identify ideal characteristics after assessing the intervention's readiness for an impact evaluation and the grantees' current investment in measurement and learning. The foundation can then help grantees select an evaluator with those characteristics.

Capacity to Scale Successfully

The SPREE process was designed to counteract the struggles that often occur during scaling and diminish the effectiveness of an intervention. When seeing an opportunity to serve additional participants, grantees might not consider the need to step back and ensure they are prepared to maintain the intervention's effectiveness as they extend their reach. Foundations can help grantees both see the need to build capacity for scaling and gain that capacity. Requiring

grantees to assess whether they are ready for scaling before funding an effort can help them see how a priori preparation can smooth the transition to implementing an expanded, replicated, or adapted version of the intervention. Once that assessment is complete, foundations can provide funding to develop the infrastructure to support successful scaling. Examples of such funding include developing implementation manuals for an intervention, purchasing training materials, and acquiring equipment to build staff capacity to implement an intervention after scaling. By using the SPREE process, foundations can work with grantees to increase their capacity to scale an effective intervention and, by doing so, expand their own reach and improve more lives.

Conclusion

As foundations look to enhance their decision-making processes, a strategic use of research evidence can help them make more efficient funding decisions. The SPREE process can provide a systematic way to identify interventions that are likely to improve desired outcomes for their participants. The SPREE process can also help foundations identify whether these interventions and the organizations implementing them are ready to successfully scale the intervention. By adopting such a process, foundations can expand their reach and address needs for more people and communities.

Engaging in the SPREE process also can build or further develop a culture of measurement, learning, and evaluation in both the foundation and among the grantees it funds. As exemplified by the experiences of the CNCS, applying the process can stimulate internal conversations within foundations. These conversations can guide foundations in learning how to best use evidence in decision-making, identifying ways to support grantees that need to build evidence for their intervention's effectiveness, and recognizing situations in which grantees require additional resources to support their scaling and sustain their intervention's effectiveness.

The SPREE process was designed to counteract the struggles that often occur during scaling and diminish the effectiveness of an intervention. When seeing an opportunity to serve additional participants, grantees might not consider the need to step back and ensure they are prepared to maintain the intervention's effectiveness as they extend their reach. Foundations can help grantees both see the need to build capacity for scaling and gain that capacity.

Acknowledgments

This research was inspired by work performed by Mathematica under contract to the Corporation of National and Community Service, contract number GS10F0050L/CNSHQ16F0049. The views expressed are those of the authors and should not be attributed to Mathematica or the CNCS. The work has benefited greatly from discussions with and comments from Anthony Nerino, Lily Zandniapour, Mary Hyde, and Roshni Menon at CNCS, and Diane Paulsell and Josh Haimson at Mathematica. Mary Anne Anderson, Karen Needels, Rebekah Selekman, Andrei Streke, Katie Eddins, and Christopher Jones provided outstanding collaboration on research feeding into this work.

References

Achieving the Dream. (2011). *Scaling community college interventions*. Silver Spring, MD: Author. Retrieved from https://www.publicagenda.org/files/CuttingEdge2.pdf

Austin, M. J., & Claassen, J. (2008a). Impact of organizational change on organizational culture. *Journal of Evidence-Based Social Work, 5*(1–2), 321–359. https://doi.org/10.1300/J394v05n01_12

Austin, M. J., & Claassen, J. (2008b). Implementing evidence-based practice in human service organizations. *Journal of Evidence-Based Social Work, 5*(1–2), 271–293. htpps://doi.org/10.1300/J394v05n01_10

Barker, P., Reid, A., & Schall, M. W. (2016). A framework for scaling up health interventions: Lessons from large-scale improvement initiatives in Africa. *Implementation Science, 11*(12), 1–11. https://doi.org/10.1186/s13012-016-0374-x

Bernfeld, G. A. (2006). The struggle for treatment integrity in a 'dis-integrated' service delivery system. *The Behavior Analyst Today, 7*(2), 188–205. https://doi.org/10.1037/h0100086

Blase, K., & Fixsen, D. (2013). *Core intervention components: Identifying and operationalizing what makes programs work*. Washington, DC: U.S. Department of Health and Human Services, Office of Human Services Policy.

Breitenstein, S. M., Gross, D., Garvey, C., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation fidelity in community-based interventions. *Research in Nursing & Health, 33*(2), 164–173. https://doi/org/10.1002/nur.20373

Cooley, L., Ved, R., & Fehlenberg, K. (2012). *Scaling up — From vision to large-scale change: Tools and techniques for practitioners*. Washington, DC: Management Systems International. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&uact=8&ved=2ahUKEwj_z6mT-ZPgAhVGhuAKHTRRCogQFjACegQICBAC&url=https%3A%2F%2Fwww.aph.gov.au%2FDocumentStore.ashx%3Fid%3D23f40c5c-1c1a-473e-abed-a72b4fd6a9dc%26subId%3D252724&usg=AOvVaw2v23ip7om3EDoRAPXrY9i7

Despard, M. (2016). Challenges in implementing evidence-based practices and programs in nonprofit human services organizations. *Journal of Evidence-Informed Social Work, 13*(6), 502–522. https://doi.org/10.1080/23761407.2015.1086719

Fixsen, A. A. M. (2009). *Defining scaling up across disciplines: An annotated bibliography*. Retrieved from https://pdfs.semanticscholar.org/57fc/57337970444b7bfc536adee7d0791a36471b.pdf

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature* (FMHI Publication #231). Tampa: University of South Florida, Louis de la Parte Florida Mental Health Institute. Retrieved from https://nirn.fpg.unc.edu/resources/implementation-research-synthesis-literature

Garg, R. (2016). Methodology for research I. *Indian Journal of Anaesthesia, 60*(9), 640–645. https://doi.org/10.4103/0019-5049.190619

Gondolf, E. W. (2015). The evidence-based practice movement: Contributions, controversies, and recommendations. In R. A. Scott, S. M. Kosslyn, & N. Pinkerton (Eds.), *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource*. Hoboken, NJ: Wiley. Available from https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118900772.etrds0335

Klingner, J., Ahwee, S., Pilonieta, P., & Menendez, R. (2003). Barriers and facilitators in scaling up research-based practices. *Exceptional Children, 69*(4), 411–429.

Larson, S. R., Dearing, J. W., & Backer, T. E. (2017). *Strategies to scale up social programs: Pathways, partnerships, and fidelity*. New York, NY: Wallace Foundation. Available at https://www.wallacefoundation.org/knowledge-center/pages/how-to-scale-up-social-programs-that-work.aspx

McElroy, L. M., & Ladner, D. P. (2014). Defining the study cohort: Inclusion and exclusion criteria. In T. Pawlik & J. Sosa (Eds.), *Success in academic surgery: Clinical trials* (pp. 131–139). London, UK: Springer.

Meehan, W. F., & Jonker, K. S. (2018, March 30). Earning the right to scale. *Stanford Social Innovation Review*. Retrieved from https://ssir.org/articles/entry/earning_the_right_to_scale

Mihalic, S., & Irwin, K. (2003). Blueprints for violence prevention: From research to real-world settings — Factors influencing the successful replication of model programs. *Youth Violence and Juvenile Justice, 1*(1), 1–23.

Miller, W. R., Sorensen, J. L., Selzer, J. A., & Brigham, G. S. (2006). Disseminating evidence-based practices in substance abuse treatment: A review with suggestions. *Journal of Substance Abuse Treatment, 31*, 25–39.

National Implementation Research Network. (2018). *The hexagon: An exploration tool*. Chapel Hill, NC: Author. Retrieved from https://nirn.fpg.unc.edu/resources/hexagon-exploration-tool

Needels, K., Selekman, R., Jones, C., Richman, S., & Maxwell, N. (2018). *Planned scaling activities of CNCS-funded organizations: Benchmark findings*. Chicago, IL: Mathematica Policy Research.

Pew Charitable Trusts. (2016, July). *A guide to evidence-based budget development: How to use research to inform program funding decisions*. Philadelphia, PA: Author. Retrieved from https://www.pewtrusts.org/en/research-and-analysis/issue-briefs/2016/07/a-guide-to-evidence-based-budget-development

Richman, S., Maxwell, N., Streke, A., Needels, K., & Eddins, K. (2018). *Evidence of effectiveness in CNCS-funded interventions: Benchmark findings*. Chicago, IL: Mathematica Policy Research.

Walker, B., & Soule, S. A. (2017, June 20). Changing company culture requires a movement, not a mandate. *Harvard Business Review*. Retrieved from https://hbr.org/2017/06/changing-company-culture-requires-a-movement-not-a-mandate

Zandniapour, L., & Vicinanza, N. (2013, October). *Budgeting for rigorous evaluation: Insights from the Social Innovation Fund*. Washington, DC: Corporation for the National and Community Service. Retrieved from https://www.nationalservice.gov/sites/default/files/documents/Budgeting_for_Evaluation.pdf

Nan Maxwell, Ph.D., *is a senior researcher at Mathematica. Correspondence concerning this article should be addressed to Nan Maxwell, Mathematica, 505 14th Street, Suite 800, Oakland, CA 94612 (email: nmaxwell@mathematica-mpr.com).*

Scott Richman, Ph.D., *is a survey researcher at Mathematica.*