Evaluation Principles and Practices

A Guide to Evaluation at the William and Flora Hewlett Foundation

2ND EDITION, MARCH 2019

by Fay Twersky, Amy Arbreton, Prithi Trivedi

WILLIAM + FLORA
Hewlett Foundation

7 PRINCIPLES OF EVALUATION

1   We lead with <u>purpose</u>.

2   Evaluation is fundamentally <u>a learning process</u>.

3   Evaluation is an explicit and <u>key part of the strategy lifecycle</u>.

4   We <u>strategically choose</u> what to evaluate.

5   We choose methods that <u>maximize rigor without compromising relevance</u>.

6   We <u>share our findings</u> with <u>appropriate audiences and publicly</u>.

7   We <u>use the data</u>!

TABLE OF CONTENTS

Evaluation Principles and Practices
A Guide to Evaluation at the William and Flora Hewlett Foundation
2nd Edition, March 2019

Foreword

Evaluation is, and has long been, an important tool to help us learn and improve our strategies at the William and Flora Hewlett Foundation. In 2012, we harmonized our approach to evaluation across our programs, with input from staff and external advisors. This resulted in our first commonly defined set of Evaluation Principles and Practices,[1] designed to guide program staff on how to commission purposeful evaluations and use evaluation findings to inform decision making.

In 2017, after five years of implementing these principles and practices, we took a step back to systematically analyze the foundation's evaluation quality and spending patterns.[2] Based on our findings, which were largely positive but did surface areas for improvement, we thought it made sense to update our evaluation guidance. The updated guidance in this document reflects lessons from living these principles and insights from the broader field of evaluation. As such, this document is a revision of the original Principles and Practices paper and aims to supplement the principles—which remain the same—with the practical lessons that have been gained over the past five years.

Fay Twersky, Amy Arbreton, and Prithi Trivedi

Introduction

Evaluation is part of the fabric of the William and Flora Hewlett Foundation. It is referenced in our Guiding Principles[3]. Evaluation is also an explicit element of our strategy lifecycle (described in Outcome-Focused Philanthropy)[4] and is meant to be practiced with frequency, intensity, and skill across all programs.

The purpose of this document is to provide a refreshed description of our philosophy, purpose, and expectations specifically regarding evaluation at the Hewlett Foundation; to clarify staff roles; and to outline available support. Although the foundation's programs and initiatives differ widely, our evaluation principles and guidance for practice apply to all of them. The principles described in the original version of this paper are enduring, but this version of the paper includes additional details, as well as updated sections on practice—with new examples and refined guidance.

> **There are many ways that foundation staff learn and measure progress—evaluation, monitoring, and feedback. Evaluation is the focus of this paper.**
>
> **What is evaluation?**
>
> **Evaluation** is an independent, systematic investigation into how, why, to what extent, and for whom outcomes or goals are achieved. It can help the foundation answer key questions about strategies, substrategies, clusters of grants, and (occasionally) single grants.
>
> **What is monitoring?**
>
> Strategy **monitoring**—what we call "**tracking progress**"[5]—is a process for checking in on a regular basis to make sure we are on track toward strategy outcomes. Tracking progress and evaluation are complementary but different. Tracking progress helps keep track of and describe progress toward some longer-term change we want to see, whereas evaluation helps us understand why and how change is happening (or not). When tracking progress, staff identify the key signs and landmarks that will help them understand if they are on the right track or off course; we call these signs "implementation markers." In contrast, evaluation will often draw on progress tracking data but will typically include other methods and data sources to answer more strategic questions.
>
> **What is feedback?**
>
> Foundation staff also learn through **feedback** from the experience of the people who we hope will ultimately be positively touched by our work. Feedback[6] provides new information and insight that can be a valuable source for learning; while it complements evaluation and tracking progress, it merits focus as a distinct channel.

The following paper is organized into three substantive sections: (a) Principles, (b) Roles and Responsibilities, and (c) Practice Guide.

Our primary audience for this paper is our foundation staff, to promote an internal culture of inquiry and practical evaluation. In particular, **program staff—as those responsible for planning, budgeting, commissioning, and using third-party evaluations**—should carefully review the guidance and tips presented throughout the practice guide.

We share the paper broadly—not as a blueprint, but in a spirit of collegiality—and with an interest in contributing to others' efforts and continuing our collective dialogue about evaluation practice.

This document has been updated to reflect many lessons, including those from three significant bodies of work the foundation has undertaken since the original principles and practices guidance was developed and adopted in 2012.

OUTCOME-FOCUSED PHILANTHROPY

In 2016, the foundation formally adopted Outcome-Focused Philanthropy (OFP), an evolution of our approach to strategy;[7] **OFP incorporates evaluation into its guidance, outlining its importance at every stage of the strategy lifecycle: origination, implementation, refresh, and exit.** OFP reinforces both the importance of evaluation and the strong connections between strategy and evaluation. The OFP guidebook states:

> The usual practice in philanthropy has been to think about evaluation as something to do [only] at the refresh or exit stage. In fact, evaluation is relevant and important at every stage of the strategy lifecycle. **Done well, it clarifies assumptions, contextualizes evidence, and helps us learn and adapt as our work proceeds.** It is useful and important to integrate evaluation planning into the development of a new strategy from the outset. Building evaluation into the origination stage provides a proper "baseline" against which to measure subsequent developments, prepares staff to collect data in a useful and common format, lets grantees know what to expect and when, and sets us up to engage in ongoing evaluation in the implementation phase.

DIVERSITY, EQUITY, AND INCLUSION WORK

The foundation refined its Guiding Principles in 2016, bringing additional attention and explicit focus to issues of diversity, equity, and inclusion[8] (DEI)—both internally, at the foundation, and externally, in our grantmaking and related activities. It is now referenced specifically in the Hewlett Foundation's Guiding Principles: [9]

> We seek to promote the values and practice of diversity, equity, and inclusion in our workforce, our culture, and our grantmaking. When we speak of diversity and inclusion, we mean the whole range of attitudes, outlooks, and perceptions that matter to the people who work with us—whether coming from familiar sources of personal identity, like race, gender, or religion; from less common sources that are particular to our institution, like place in the foundation's hierarchy; or from sources that are idiosyncratic and individual in nature.

The foundation's work on diversity, equity, and inclusion has included a growing number of projects, workgroups, and activities—all charged with improving how we consider, acknowledge, and account for groups that have been marginalized or historically disadvantaged. How we practice evaluation at the foundation has been one area of focus through this process. There are three types of considerations we want to make to apply a DEI lens in evaluation: first, the specific *evaluation questions asked*, includ-

The foundation seeks to promote the values and practice of diversity, equity, and inclusion in its workforce, its culture, and its grantmaking.

ing who asks them and whether any address strategy-related issues of diversity, equity, and inclusion; second, *the process and lens* the evaluator uses to gather and interpret different types of data, from whom the data are collected, and whether and how different stakeholders are engaged in interpreting and using findings; and third, *who conducts the evaluations* and their competencies related to understanding the relevant contexts, communities, issues, and tools.

AN ASSESSMENT OF THE QUALITY OF OUR EVALUATIONS

In 2017, we took stock of our evaluation practice—and found both strong positive developments and areas for improvement.[10] On the positive side, the quality of the evaluations we commission has improved over time. They are more rigorous, more purpose-driven, more useful and used, and more widely shared. The analysis also uncovered three primary ways in which we can make our evaluations still more valuable.

> **First, our evaluations would benefit from more rigor,** including sharper evaluation questions and more sources of comparison—which would strengthen our ability to learn from our evaluations.

> **Second, we can and should do much more to** engage grantees when planning, implementing, and using evaluations.

> **Third, we need to take greater steps to ensure that we share what we are learning, publicly,** from every evaluation we commission— whether in full form, executive summary, or a presentation; this is part of our commitment as a foundation to increased openness and transparency,[11] as reflected in our foundation's guiding principles,[12] and allows more interested parties to learn from the evaluations we have commissioned.

The foundation is committed to openness, transparency, and learning.

The Hewlett Foundation's Seven Principles of Evaluation Practice

We aspire to have the following principles guide our evaluation practice:

---

7 PRINCIPLES OF EVALUATION

1   We lead with <u>purpose</u>.

2   Evaluation is fundamentally <u>a learning process</u>.

3   Evaluation is an explicit and <u>key part of the strategy lifecycle</u>.

4   We <u>strategically choose</u> what to evaluate.

5   We choose methods that <u>maximize rigor without compromising relevance</u>.

6   We <u>share our findings</u> with appropriate audiences and publicly.

7   We <u>use the data</u>!

---

1. WE LEAD WITH PURPOSE.

We design evaluation with actions and decisions in mind. We ask, "How and when will we (and others) use the information that comes from this evaluation?" By anticipating our information needs, we are more likely to design and commission evaluations that will be useful and used. It is all too common for evaluations to be commissioned without a clear purpose and then to be shelved without generating useful insights.

2. EVALUATION IS FUNDAMENTALLY A LEARNING PROCESS.

As we engage in evaluation planning, implementation, and use of results, we actively learn and adapt. Evaluative thinking and planning inform strategy and target-setting. They help clarify evidence and assumptions that undergird the approach to our work. Establishing evaluation questions helps us make visible and refine our thinking about how, why, to what extent, for whom, and when outcomes or goals are expected to be achieved. As we implement our strategies, we use evaluation as a key vehicle for learning, bringing new insights to our work and to the work of others. A key part of learning from evaluation is taking time to reflect on findings and to <u>ask, "Now what?"</u>[13]

3. **WE TREAT EVALUATION AS AN EXPLICIT AND KEY PART OF THE STRATEGY LIFECYCLE.**

Building evaluative thinking into strategy does two things: (a) it helps articulate key assumptions and logical (or illogical) connections in a theory of change; and (b) it encourages us to establish a starting point for evaluation questions and propose a way to answer those questions in a practical, meaningful sequence, with actions and decisions in mind, throughout the strategy lifecycle. In practice, this can take many forms, such as a planned series of periodic evaluations of substrategies or clusters of grantees, a developmental evaluation that is ongoing, a formative or summative evaluation planned for a key juncture, or a summary of lessons learned at a planned exit.

4. **WE CANNOT EVALUATE EVERYTHING, SO WE CHOOSE STRATEGICALLY.**

Several criteria guide decisions about where to put our evaluation-related time and dollars, including urgency to consider course corrections or future funding decisions; the opportunity for learning; the potential for strategic or reputational risk; and size of investment as a proxy for importance. Within these parameters, every strategy, or a key part of every strategy, will have an evaluation underway within three years of origination or refresh. Planning for an evaluation within such a timeframe ensures that we do not go too long without getting an external, third-party perspective on how well (or not) the work is going and whether any unexpected issues have arisen.

5. **WE CHOOSE METHODS OF MEASUREMENT THAT ALLOW US TO MAXIMIZE RIGOR WITHOUT COMPROMISING RELEVANCE.**

We match methods to questions and do not choose one approach, or privilege one method over others. We select evaluation designs that use multiple methods and data sources, when possible, in order to strengthen our evaluation design and reduce bias. All evaluations clearly articulate methods used and the limitations of those methods. Evaluations include comparative reference points or methods to help us assess how well we are doing compared with our own and others' expectations—over time, and across types of interventions, organizations, populations, or regions.

6. **WE SHARE WHAT WE ARE LEARNING WITH APPROPRIATE AUDIENCES AND SHARE PUBLICLY, SO THAT OTHERS CAN LEARN AS WELL.**

As we plan evaluations, we consider and identify audiences for the findings. We communicate early with our grantees and co-funders about our intention to evaluate and our plans to share findings, and we involve them as appropriate in issues of planning, design, and interpretation. We presumptively share the results of our evaluations so that others may learn from our successes and failures. We will make principled exceptions on a case-by-case basis about whether to share a full report, executive summary, or presentation from an evaluation, with care given to issues of confidentiality and not wanting to cause harm.

7. **WE USE THE DATA!**

Not using our findings is a missed opportunity for learning, improvement, and course correction—and a waste of time, energy, and resources. It is imperative that we take time to reflect on the evaluation results; generate implications for our strategies, grantees, policy, and/or practice; and adapt as appropriate. We recognize the value in combining the insights from evaluation results with our own experiences. We support our grantees in doing the same.

Roles and Responsibilities

Many people are involved in evaluations. **Programs have the primary responsibility for evaluations. They are responsible for building evaluations into their strategy, identifying what and when they will evaluate, and commissioning, managing, and using findings from the evaluations.** The evaluation officer in the Effective Philanthropy Group is available to support program staff. Grants management, legal, communications, and finance staff also have roles in the process.

PROGRAM STAFF

Program officers or directors play the leading role in planning for when and what aspects of their strategies to evaluate, for commissioning evaluations, and for managing them from planning and implementing to using and sharing. **Commissioners are the primary point people for coordinating evaluations, and they must carefully review evaluation reports, executive summaries, and presentations before public release. They are responsible for checking for sensitive information that may not be appropriate for public sharing and making sure that the evaluator has accurately described, for example, advocacy work with an appropriate disclaimer** (**see Appendix B for examples**)**.** The evaluation commissioner is responsible for managing adherence to the foundation's "Evaluation Checklist: Components for a Successful Evaluation."

Most commonly, program associates create a system for accessing data internal to the foundation (such as background documents and grant reports) to share with the evaluator. This is an important role and can be time-consuming at key junctures of the evaluation process. The program associates also typically assist with review and selection of evaluators, support contract development and monitoring of the contract, and help organize and participate in workshops (e.g., with grantees to discuss findings, and with evaluation advisory groups).

Some programs teams have selected one team member to manage their evaluations. This person coordinates with other members of a strategy team and acts as the point person with evaluation consultants. (For example, the program director for the Madison Initiative manages evaluations for that team, which consists of the director, two program officers, and a program associate. A Global Development and Population Program Officer manages the formative evaluation of the Transparency, Participation, and Accountability strategy, working with a team of four program officers and three program associates.) Other programs have individual staff manage their own substrategy or grant cluster evaluations.

GRANTEES

Grantees are involved as participants and often as intended audience and users of evaluation findings. Grantees should be informed about plans for evaluation as early as possible and should be included—to the extent reasonable and feasible—throughout the planning, implementation, and use (and sharing) phases of an evaluation. (See Appendix C for guidance on engaging grantees.)

THE EFFECTIVE PHILANTHROPY GROUP

As the foundation's approach to evaluation has become more deliberate and systematic, the foundation's leadership has come to appreciate the value and timeliness of expert support for evaluation. Therefore, as part of its Effective Philanthropy Group (EPG), in 2013 the foundation created a central support function for programs' evaluation efforts.

EPG staff act as the Hewlett Foundation's in-house experts on philanthropic practice, combining institutional knowledge and technical assistance to help program staff be more effective grantmakers. In the case of the evaluation officer's role, this includes all things related to evaluation, as well as coordinating effectively as needed with the other internal EPG officers—Strategy, Organizational Learning, and Organizational Effectiveness—on any overlapping areas of learning, assessment, and training.

Programs are not technically required to use the evaluation officer's support; however, it is a central form of support the foundation makes available to any program staff seeking evaluation assistance. The evaluation officer's role currently includes the following:

Internal Consulting to Program Staff
The bulk of the evaluation officer's time is spent on internal consulting, to support program staff as they work on their evaluations. Like the other EPG staff, the evaluation officer provides support in three main ways:

- *Resource Provider.* Maintain updated, practical resources to share with programs as examples and templates for each step in the evaluation process: e.g., an "Evaluation Checklist" one-pager and examples of evaluation planning tools, requests for proposals (RFPs), evaluation designs, and evaluation products.

- *Ad-Hoc Advisor.* On a periodic and as-requested basis, step in and support program staff: e.g., in framing evaluation priorities, questions, sequencing, and methods; in development of RFPs and review of proposals; and in supporting internal and external sharing of results—coordinating with relevant program, legal, and communications staff as well as grantees and other external partners.

- *Guide-by-the-Side.* When needed, provide deep ongoing support to program staff throughout an evaluation.

If program staff are unsure what type of support is needed, they can ask the evaluation officer to review options.

Institutional Knowledge of Evaluation Principles and Practices
The evaluation officer is responsible for keeping the principles and practices up to date; serving as the internal source for all questions (internal and external) related to evaluation practice at the foundation; tracking evaluation quality and spending (with assistance from the finance department), and orienting and training staff in the foundation's principles and practices guidance. As a centralized function in a foundation with decentralized program staff, the evaluation officer also plays a role in sharing relevant lessons across programs, so all program staff can benefit from promising practice and lessons learned.

Sharing Evaluations Externally
The evaluation officer considers how to position the foundation as a good citizen and leader by staying current with, and contributing to, the philanthropic evaluation field. This is done in close coordination with the communications staff.

OTHER DEPARTMENTS

The communications department helps program staff to consider with whom, what, how, and when to share evaluation findings. Particularly if informing the field is a key objective of sharing evaluation findings, communications staff work with programs and consultants to plan for the best dissemination approach.

Program staff and communications staff are required to discuss proposed evaluations that raise specific legal concerns, such as lobbying or election-related work, with their legal team partner. For example, staff should speak with their legal team partner if evaluation questions refer to legislation, ballot initiatives, or campaigns and elections, or if evaluators plan to interview U.S. or foreign legislators. The legal department should be contacted before requests for proposals are issued or evaluation work begins to ensure that the evaluation is compliant with the laws on lobbying and electioneering. In addition, the legal department will review contracts for all types of evaluation in accordance with the foundation's normal contract review process.

The Grants management staff is often asked to provide data from the grants management system. This might include grant reports or other information relevant to a particular strategy. The Hewlett Foundation will typically provide a consultant access to grant materials (proposals, reports, previous reviews, etc.), contact information for grantees, and our own internal strategy materials.

The finance department reviews spending on evaluations and works closely with the evaluation officer to track evaluation spending as a proportion of each program's grant spending, to determine whether it is in line with our benchmark of 1.5 to 2 percent of the total grantmaking budget.[14]

EVALUATION CONSULTANTS

By definition, our evaluations include third-party evaluators. We expect the evaluators with whom we contract to follow the American Evaluation Association Guiding Principles for Evaluators.[15] We rely on the evaluators to gather data, ensure appropriate confidentiality, analyze and interpret findings, and prepare and present findings to the different audiences identified.

Typically, we expect an evaluator to draw on a variety of quantitative and qualitative data collection techniques—going beyond, for example, review of grant reports. The data collection strategy should, in most cases, be geared toward capturing multiple and diverse perspectives and use varied sources, to allow for triangulation across sources.

We expect the evaluator will: be dispassionate in reporting to us—i.e., we want honest, accurate, and useful information, and we do not expect or want flattery; be transparent about time commitment and expectations for evaluation participants (program staff, grantees, other participants); collect all data with appropriate confidentiality; as needed, share/interact with other evaluator(s) working with the same key partners; synthesize data and provide findings in formats and ways that encourage feedback on interpretation; and support discussion and learning.

In reporting, evaluators should address the evaluation questions and report on key successes and challenges. Evaluation reports should explain the evaluation's limitations and include data collection protocols and tools in appendices. In reports, the evaluators should clearly distinguish findings, conclusions, and (if they are requested to provide them) a prioritized set of recommendations.

Editorial guidance for evaluation reports to be shared publicly is included in Appendix B. This editorial guidance is shared with evaluators with every evaluation contract.

Practice Guide: Planning, Implementation, and Use

This practice guide follows the three stages of evaluation: (a) planning, (b) implementation, and (c) practical use of the evaluation findings.

```
······  PLANNING  ········  IMPLEMENTING  ········  USING  ······>
```

The seven evaluation principles apply across these three stages of an evaluation—and are visible at various points across the following practice guide.

Included in the practice guide are case examples illustrating successes, challenges, and lessons learned.

Planning

Planning is perhaps the most important and complex part of evaluation. We plan for evaluation at two levels:

- *Big picture:* As part of strategy, we use evaluative thinking to make explicit the assumptions that undergird our theories of change, consider when and how evaluation data will be used, and plan for commissioning specific evaluations to help us learn and adapt throughout the strategy lifecycle (see Appendix A).

- *Specific evaluation focus:* This includes articulating evaluation questions, planning for the time and budget to engage grantees, finding and contracting with an appropriate evaluation partner, and being well prepared to use and share the findings.

BEGINNING EVALUATION PLANNING EARLY

Even before commissioning a specific evaluation, **evaluative thinking** can help sharpen a theory of change—an articulation of beliefs and assumptions, explaining why proposed activities are expected to contribute to outcomes and long-term goals. As part of the OFP process, for example, a program team should consider and make explicit its key assumptions—often the assumptions that link our theories of change together. For instance, consider this example of a simplified generic theory:

- If we invest in an innovative model, we hope and plan for it to be successful, and...
- If proven successful, it will be expanded to reach many more people.

In between each link are potential assumptions to be tested:

- This innovative approach can be successful.
- Effective organizations exist that can implement this approach.
- This approach can become a "model" and not just a one-off success.

- Others will be interested in adopting and supporting the model.

- Resources for growth and expansion exist to expand the model.

- When resources for growth are invested, the approach can be widely and effectively implemented with high quality.

As with many strategies, each link builds on the one before it.

Why Start Evaluation Planning Early?

The Hewlett Foundation's Organizational Effectiveness (OE) strategy, which began in 2004, seeks to help nonprofits become high-performing organizations that are healthy, sustainable, and successful in achieving their goals. OE provides relatively small, targeted grants to help build Hewlett Foundation grantees' internal capacity in areas like strategic planning, board development and governance, fundraising, and communications. In 2014, the foundation commissioned its first-ever evaluation of the OE strategy. A new OE officer had many questions—including understanding what was working, what was not, why, and whether OE grants were strengthening grantees' health and resiliency in both the short and longer terms.

The evaluation[16] found that, by and large, OE grants deliver what they promise in the near term: solid strategic plans, fundraising campaigns, leadership transition plans, and so forth. The evaluation also inspired new research into organizational assessment tools[17] that might be useful for future assessment and evaluation. But the evaluators were not able to address other important questions, including whether OE support also provides positive, broader, and longer-term value to grantees after the grant term. Crucially, the evaluators did not have the information they needed because the program had not planned for evaluation early enough and had not been collecting data consistently and systematically. They may or may not have been able to answer vexing questions about broader and longer-term value, but at least they would have been equipped to try.

Starting evaluation planning early, including sharing those plans with grantees and others who will likely be involved (e.g., other funders), protects against four common pitfalls: (a) missing a "baseline"; (b) not having data available or collected in a useful common format; (c) surprised, unhappy, or unnecessarily burdened grantees; and (d) a strategy or evaluation that is not optimally designed to generate the desired information to inform learning and decision making. It also helps identify opportunities for small tests or experimentation in a strategy that could make a big difference for the strategy's long-term trajectory.

Planning for evaluation does not mean casting those plans in concrete. In fact, given that our strategies typically unfold dynamically, it is essential to revisit and modify evaluation plans, as a strategy matures and refreshes.

Purpose-Driven Evaluation Will Shift Focus Over Time

The Madison Initiative[18], which focuses on strengthening U.S. democracy and its institutions—especially Congress—in a time of political polarization, commissioned the Center for Evaluation Innovation to work closely with foundation staff throughout the initiative's initial three-year exploratory period. During these early years, the Madison Initiative team selected a developmental evaluation design, an approach that embeds evaluation in the process of strategy development and implementation. The role of developmental evaluators is to be a "critical friend" to the strategy team, asking tough evaluative questions, uncovering assumptions, applying evaluation logic, and collecting and interpreting evaluative data to support strategy development with timely feedback.

Early in the evaluation, the evaluators developed a systems map. This map helped the Madison team establish a common understanding of what the initiative was trying to do—and the key variables both inside and outside of the Madison Initiative's funding areas. Working with the map helped the team recast its thinking and see holes and gaps in different ways, which then allowed the team to change the grantmaking avenues it pursued. However, the map was less helpful for informing decisions specific to what the foundation could do, relevant to their grant clusters.

Thus, as the strategy matured, smaller evaluations were commissioned of specific grant clusters working toward similar outcomes. These cluster evaluations informed strategic pivots and grantmaking decisions. As an example, by early 2016, the Madison Initiative had been investing in campaign finance grantees for several years, and began wrestling with whether, in light of technological advances and talk of "the big data revolution," foundation support for basic campaign finance data collection and curation was still necessary. Findings provided by the evaluator affirmed the team's convictions that these grantees in fact do play an important role in reform; a decision was made to support large, long-term funding to those grantees.

After a strategy refresh in 2016, the team continued its focus on smaller, targeted evaluations of grantee clusters, working toward specific outcomes. A series of sequenced evaluations will examine questions about what is, and is not, working. These evaluations are timed to produce findings at key junctures in the grantmaking process, in order to position the Madison Initiative and its grantees to act on lessons learned.

CHOOSING WHAT TO EVALUATE THROUGHOUT THE STRATEGY LIFECYCLE

We cannot (and need not) evaluate every aspect of a strategy, nor do we evaluate every grant. One criterion for what program staff choose to evaluate is their openness to change—and readiness to **challenge strongly held beliefs.** Often the most strongly held beliefs are those assumptions embedded in the theory of change—that an innovation will succeed, for instance, or that others will adopt a "successful model."

Several other criteria guide our decisions about where to put evaluation dollars. Highest priority is given to the following considerations:

- Urgency for timely course correction or decisions about future funding.
- Opportunity for learning, especially for unproven approaches.
- Risk to strategy, reputation, or execution.
- Size of grant portfolio (as a proxy for importance).

Program officers, with the supervision of program directors, determine what aspects of the strategy to evaluate and when. Teams submit an evaluation plan on an annual basis and update these plans each year.

Most of the time, program staff will plan for an evaluation to focus on a strategy, substrategy, or cluster of grants that meet these criteria, rather than focusing on a single grant. The exception is when a grant is essentially operating as an initiative or cluster in and of itself.

OUTCOME-FOCUSED PHILANTHROPY STRATEGY HIERARCHY

PROGRAM
STRATEGY OR INITIATIVE
SUBSTRATEGY
GRANT CLUSTER
INDIVIDUAL GRANT
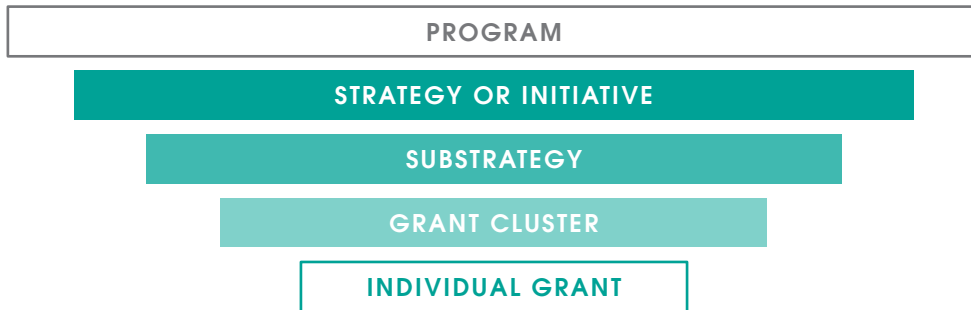
While we encourage choosing strategically what to evaluate, we recommend that all strategies should begin an evaluation (**in whole or in part**) **within three years of origination or refresh.** Planning for an evaluation within that time frame encourages us not to let too much time go by without an external set of eyes on progress; doing so also sets us up with evaluations that can inform strategy refreshes—which most strategies go through roughly every five years.

Most strategies operate with several substrategies, often with multiple clusters nested in each. **Many program staff identify smaller substrategy and grant cluster areas as highest priority for evaluation, to inform strategy and grantmaking decisions.** For example, the Cyber Initiative chose to evaluate its work on network building[19] early in the life of the strategy, since that work was identified as a necessary element to support the overall strategy goal. After a strategy refresh, the Global Development and Population Program's International Reproductive Health team identified for evaluation several substrategies it was introducing. These included testing new tools and approaches,[20] working in Francophone West Africa,[21] and supporting local advocacy in sub-Saharan Africa.[22] These substrategies were identified because they held more risk for successful implementation and because the team believed there were benefits to early learning and adaptation. In fact, that plan turned out perfectly, as the specific evaluations for these substrategies did, in fact, substantially inform decisions during implementation.

When it comes to strategy-level evaluation, typically, no single evaluation can tell us if a strategy has been successful or is on track. Such a comprehensive assessment—most commonly used to inform a strategy refresh—likely requires synthesis of multiple evaluations of smaller, cluster-level evaluations; summary and analysis of relevant implementation markers; and active reflection on, and interpretation of, the results in context. This process can be more of a "quilting art" than an exact science. There is value in having a third party assist with such an evaluation to increase objectivity. The 2018 Western Conservation strategy refresh,[23] for example, relied on a third-party consultant to weave together a comprehensive retrospective evaluation[24]; a set of cluster-specific evaluations; a deep dive into equity, diversity, and inclusion issues among grantees; and research on best practices for effectively communicating and collaborating with rural Western communities.

Frequently, the foundation uses regranting intermediaries to extend its reach and increase the impact of its grant dollars and results. **Because we are delegating to these intermediaries what might be considered our stewardship role, we have an even greater responsibility to evaluate their efforts.** By definition, large intermediaries rank high on the risk and size criteria above, and evaluating them typically offers important learning opportunities. Also, whenever we contribute to creating a new inter-

> When it comes to strategy-level evaluation, typically, no single evaluation can tell us if a strategy has been successful or is on track.

mediary organization or fund the launch of a major new initiative, it is important to evaluate not only the strategic elements, but also issues of organizational health (e.g., leadership and board development) and effective execution (e.g., to what extent is something happening, as planned)—challenges that vex many startups. In some cases, we commission an evaluation ourselves, such as the evaluation of the Open Educational Resources Research Hub;[25] in other cases, particularly for funder collaborations, we may contribute to pooled funding for evaluations. This was the case with the evaluation of the Fund for Shared Insight,[26] where many funders contribute and the intermediaries commission the evaluation. When we are interested but will not be involved in a decision-making role, we might give a supplemental grant to a grantee for them to commission an evaluation and serve, for example, on an evaluation advisory committee. As with all of our evaluations, it is important to be clear on the purpose and to weigh the benefits and challenges of each approach—with regard to quality, buy-in, likelihood of use, and broad sharing.

Multiple Evaluations Across the Strategy Lifecycle Strengthen Learning and Adaptation

When the Hewlett Foundation introduced its Deeper Learning strategy in 2010, the goal was to spread a concrete set of skills that American students should be learning. The strategy anticipated that high schoolers who gain academic knowledge alongside inter- and intrapersonal skills will be better prepared as college students, workers, and citizens. The Deeper Learning team commissioned multiple evaluations over its first six years.

The team commissioned a strategy-level evaluation[27] in 2013 with a formative purpose: to assess the execution of the strategy during its first four years, assess the viability of the strategy's assumptions, and provide concrete recommendations on how to improve the prospects of attaining the ultimate 2017 goal to ensure that 8 million students (about 15 percent of the K-12 public school population) were taught higher-order skills.

In deciding which aspects of the strategy to evaluate over the years following the 2013 evaluation, the team continued to lead with purpose and looked at which key decisions could benefit most from evaluation. The team also considered how smaller evaluations could serve as critical input to inform a larger strategy-level summative evaluation, which would likely be commissioned in 2016.

Between 2014 to 2016, the team commissioned numerous cluster evaluations. One evaluation was helpful in informing shifts in grantmaking, when program staff needed to reduce (and more strategically focus) its funding of policy work in order to increase funding for other aspects of its strategy—a recommendation that came out of the 2013 formative assessment. Staff used evaluation findings in a number of ways, including to identify the grantees best positioned to successfully advance Deeper Learning-aligned policies—those with both strong capacity for policy impact and high alignment with Deeper Learning goals; the team shifted funding for these organizations toward longer, larger, and more general operating support grants. Another evaluation—of communications efforts—was useful for establishing the (then) current status of how Deeper Learning was communicated and understood as a term and focus in the field, and grantees' roles in shaping it.

As the time approached for their planned summative evaluation[28] of the strategy, the team settled on a set of broader strategy-level evaluation questions, with the intentional purpose of synthesizing progress from 2010 to 2015. As the team described it, "while the substrategy evaluations were precisely designed to test the individual links in our logic model, the broader 2016 summative evaluation would look at what happened in the boxes and interrogate the assumptions about the arrows linking the boxes together." This summative evaluation (along with a host of other inputs) then informed the program's strategy refresh, which began in late 2017.

It is also important to plan for commissioning an evaluation in cases where the foundation exits—to generate lessons that will be useful to multiple stakeholders inside, and potentially outside, the foundation. The Nuclear Security Initiative summative evaluation is a good example of this.[29] The evaluators summed up the lessons learned from this seven-year initiative, with respect to successes and challenges and how well the Initiative handled the exit. The evaluation also provided a broader set of lessons on how the Hewlett Foundation and other funders might track progress and evaluate policy-related efforts.[30] Many of the lessons learned from the Nuclear Security Initiative evaluation are incorporated into OFP guidance on preparing for and handling an exit.

Engaging Grantees is Critical to Building Trust and Buy-In

In 2014, the Global Development and Population program's International Reproductive Health strategy began funding new approaches to increase the effectiveness of service delivery by pairing service delivery grantees with organizations that could bring new insights about service design—drawing from the fields of behavioral economics and human-centered design (HCD). As the strategy began, program staff commissioned an evaluation to understand whether and how this new effort was working.

The program officer took several steps to ensure the success of the evaluation, including developing an RFP, being clear on the evaluation's purpose, identifying a set of evaluative questions—and sharing these items with some grantees for input. But early into the implementation of the evaluation, there were rumblings about how the evaluation was being carried out. Grantees wondered, "Why are these evaluation questions being asked?" "Why are the evaluators taking the approach to data collection they are using?" "Are there important ways in which the evaluators are biased toward their own approach to HCD?" These rumblings disrupted the ability of the evaluators to effectively carry out an evaluation.

It became clear that these evaluators would not be able to build trust and gain enough confidence to gather the data needed for the evaluation. As a result, after the completion of the contract for an initial evaluation design and inception phase, the program officer decided to end that evaluation relationship. Yet evaluating the work remained important. So the program officer tried again—this time, doing even more to get input and buy-in from all the grantees who would be involved in the evaluation. To do so, the program team took several steps: First, they traveled to and met with representatives from each of the grantees involved in the effort and asked what questions they wanted answered and what they would be looking for in an evaluation. Second, based on what the program officer heard, she put together a scoping document that identified overall purpose (key audiences and intended use) along with which questions a Hewlett Foundation-commissioned evaluation would answer, and which might be addressed by other funders, or, if appropriate, by individual grantees to answer as part of their own evaluation. Third, when she received evaluators' proposals from a subsequent RFP process, she ran the finalists by the grantees to hear of any concerns before finalizing the selection. Finally, she developed an advisory committee composed of representatives from each grantee and other funders and requested that the group weigh in at key junctures of the evaluation, including giving input on the design and data collection strategy, getting feedback on initial evaluator findings, and finally discussing findings and recommendations[31] at a "co-creation workshop."

Taking the time to get grantee input early and often and using an advisory group as a mechanism for grantee engagement throughout the evaluation process helped build trust and limit surprises. An advisory committee composed of grantee representatives and other funders can support the use of findings for learning and project improvement.

ENGAGING WITH GRANTEES

Communication with grantees early and often about expectations for evaluation is crucial— though what information is communicated, and how that information is communicated, will depend on the purpose of the evaluation commissioned and the grantee's role in it. Often, this expectation needs to be communicated and reinforced several times—at a grant's inception, again as a specific evaluation is planned, during implementation of evaluation activities and data collection, and during the use and sharing phases. For example, at a grant's inception, program staff need to inform grantees that they may be expected to participate in an evaluation, share data with the foundation and evaluators, and have the results shared publicly in some form—full, executive summary, or presentation. The shared agreement from this discussion is then reflected in the language in the Grant Agreement Letter. As one grantee who reviewed an early draft of this guide advised us, "Don't sugarcoat what the evaluation experience will entail." In the long run, everyone does better when expectations are clear.

Some evaluations involve large numbers of grantees (for example, strategy-level evaluations can include the work of dozens to hundreds of grantees), but even in these cases, to the extent feasible, it is important to get at least some input from grantees who will be most directly involved in an evaluation.

Be Good Clients!

An effective consulting engagement—whether for an evaluation or anything else—requires that we be engaged and thoughtful clients. **For evaluation, this requires planning for and allocating enough time to be a full participant in the process: planning, data collection, analysis and interpretation, and use and sharing.** So what should you do?

1. **Allocate enough time for you to participate in the evaluation as needed.** Depending on the type of evaluation you commission, this tends to be more necessary at the beginning and toward the end of an evaluation process. An evaluation where you want interim results or a more participatory approach will take even more time, for you and the evaluator.

2. **Check in with the evaluator about the time commitment they need from you,** and, in advance, define a process to ensure the evaluators receive the support and information they need to do a great job.

3. **At the start of the evaluation, take the time to orient consultants to the foundation's values and processes.** Team culture, values, and dynamics are important, and it will help the consultant to understand what these are and what issues the team might be grappling with.

4. **Give evaluators the time needed to design, gather data, analyze, reflect, and interpret.** Don't drag your feet in commissioning an evaluation and then squeeze the evaluators with an unrealistic time frame.

5. **Make sure evaluators are aware of, and you and they have the time and budget to address, priorities related to grantee engagement and DEI issues,** including, as relevant, the questions posed, the methods used, and the process for interpreting and using the findings. (Appendix C and the Equitable Evaluation site offer helpful resources for this step.)

ALLOCATING SUFFICIENT TIME

Planning for and allocating sufficient time—for program staff and the evaluator—across the phases of an evaluation's life is a key ingredient for an evaluation that is useful and used. In general, program officers are expected to effectively manage one significant evaluation at any given time; this includes proper oversight and engagement at each stage, from planning through use and sharing of results.

Though evaluations take time, they should be considered an important part of a program team's work at each stage of the strategy lifecycle: interacting with grantees and others involved, learning what is working and what is not, and informing course corrections. Program staff who have engaged in this way with evaluations have indicated the time spent has paid off.

In general, program officers—who take the lead on the evaluations they commission—will spend 5 to 20 percent of their time planning, managing, and determining how to use the results from evaluations. This overall expectation is amortized over the course of each year, though there are periods when the time demands will be more or less intensive. The most intensive time demands tend to occur at the beginning and end of an evaluation—that is, when staff are planning and then using results. During these periods, full days can be devoted to the evaluation. For instance, planning requires considerable time to clarify purpose, refine evaluation questions, pull together the necessary documents for the evaluation, engage grantees, choose consultants, and set up contracts. Program associates also typically spend quite a bit of time during these phases related to contracting, document collection and sharing, and convening activities.

During the use phase (including sharing), the time demand ramps up again as staff spend time meeting with consultants, interpreting results, reviewing report drafts, checking in with grantees and others, communicating good or bad news, identifying implications for practice, and sharing findings internally and externally—including publicly. Typically, the time demand for the program staff is not as intensive while the evaluator is implementing the evaluation data collection activities and doing the analysis—though program staff should not underestimate the time it will take to stay in touch, ask questions of the evaluators and the grantees, discuss draft protocols, and ensure everything is proceeding well.

THE TIME REQUIRED BY PROGRAM STAFF VARIES OVER THE COURSE OF AN EVALUATION. SOME EVALUATIONS, LIKE DEVELOPMENTAL EVALUATIONS, MAY REPEAT THIS CYCLE.

TIME REQUIRED

PLANNING   IMPLEMENTATION   USING AND SHARING
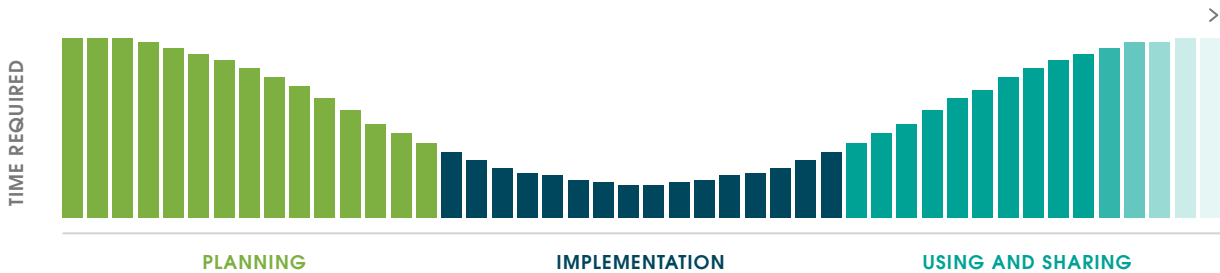
CLARIFYING AN EVALUATION'S PURPOSE

The purpose of an evaluation—which includes its planned **audience, use, and timeline** for when evaluation findings will be most useful—is central. Questions, methods, and timing all flow from a clear understanding of **how the findings will be used, by whom, and when**.

From the very beginning of the evaluation process, it is important to plan how the results will be used; if in the beginning, you take time to imagine how you might respond to different results scenarios, you are halfway toward actually using what you find out!

Below, we note three main uses (not intended to be mutually exclusive) for our evaluations:

- *To inform foundation practices and decisions.* Evaluations with this aim may inform our decision making about funding or adapting an overall strategy or substrategy, setting new priorities, or setting new targets for results. These evaluations are typically designed to test our assumptions about approaches for achieving desired results. While we share these publicly in keeping with our principles around openness and transparency, the primary audience for these evaluations is internal foundation staff.

- *To inform grantees' practices and decisions.* At times, the foundation may want to fund or commission evaluations that can be used by grantees to improve their practices and boost their performance. When the interests of the foundation and grantees overlap, it can be worthwhile to commission evaluations designed to be of value to both. Collaborating in this way can promote more candor and buy-in for the ways data are collected and results are used. In these cases, it is worthwhile to consider ahead of time the value of having an evaluator prepare an overall public report as well as individual private reports for each or select grantees. This costs more, but there is also typically greater benefit to the individual organizations.

- *To inform a field.* Evaluations that include informing a field or other funders as a distinctive purpose should be intentional about involving others (such as peer funders, or others who we hope will also benefit from the evaluation) in considering what questions would be most valuable to address. These evaluations are typically designed with influence in mind, so that others can learn what we are learning and help shape field-building or build trust in an idea or approach. Although all our evaluations involve sharing publicly, when informing a field is identified as an important purpose, it is worthwhile to work ahead of time with the evaluator to determine whether it would also be helpful to consider additional support for preparing final products for dissemination (e.g., from communications experts, editors, or data visualization specialists).

From the very beginning of the evaluation process, it is important to plan how the results will be used; if in the beginning, you take time to imagine how you might respond to different results scenarios, you are halfway toward actually using what you find out!

Most of the evaluations we are covering in this guidance have the dual purpose of informing foundation decisions and practices and those of our grantees—in both cases to support ongoing learning, adjustment, and improvement.

We Distinguish the Evaluations We Commission from the Research We Fund

There is a lot written on the differences between evaluation and research.[32] Almost every program funds research as part of a strategy itself: to identify new opportunities and best practices in an area, to identify gaps in a landscape, or to generate knowledge for a field—and to have that knowledge shape policy and practice. For example, the Madison Initiative funds research on digital disinformation; the Knowledge for Better Philanthropy strategy funds research on best practice in philanthropy; and the Global Development and Population Program's Evidence Informed Policymaking substrategy funds organizations committed to commissioning impact studies.

The research studies funded are typically distinct—in terms of purpose, process, and audience—from the evaluations we commission to understand to what extent, for whom, how, and why a strategy is working or not. Indeed, because these research studies are part of our strategies, they are often subjects of the evaluations that we commission. For example, in the evaluations of the Knowledge Creation and Dissemination[33] and the Population and Poverty Research[34] strategies, program staff addressed evaluation questions about the quality and reach of the research and adoption by intended audiences.

STRATEGY	RESEARCH (PART OF STRATEGY)	EVALUATION QUESTIONS
Knowledge for Better Philanthropy	In addition to supporting basic and applied research on philanthropy, grants supported leading journals in the field that disseminate knowledge and efforts to create new systems and platforms that would provide better solutions to learning and professional development. • Stanford Social Innovation Review • Center for Effective Philanthropy • Bridgespan Group • Issue Lab • National Committee for Responsive Philanthropy	• How do grantees measure and understand their impact to-date related to knowledge production and dissemination aimed at informing, influencing, and improving donors'/grantmakers'/funders' thinking and decisionmaking? • What knowledge on philanthropy and other aspects of the social sector is being produced by grantees? • How have grantees disseminated knowledge on philanthropy and other aspetcs of the social sector? • Who is using the disseminated knowledge and how is it being used?
Population and Poverty Research Initiative (PopPov)	Between 2005-2014, the Hewlett Foundation invested more than $25 million in a body of research on the relationship beetween population dynamics and micro- and macroeconomic outcomes • Center for Global Development • Population Reference Bureau • World Bank	• To what extent did PopPov strengthen the field of economic demography? • What contribution has PopPov made to the evidence base? • To what extent did PopPov yield policy-relevant research? • How did the design and implementation of PopPov affect outcomes?

When considering use, it is important that we examine our level of openness to a range of results and pin down what degree of rigor and strength of evidence in the evaluation would be required to change the minds of the various users of the evaluation. Evaluation is worthwhile only if one can imagine being influenced by the findings. What strength of evidence will change our minds and the minds of the others we hope to engage? If we are not willing to change our minds, or the degree of evidence to change our minds is too costly or time-consuming to obtain, we should reconsider the value of spending money on evaluation.

What is the timeline for when the findings will be most useful? One criticism of evaluation is that results often come **too late to be useful**. But that is in our control. There are trade-offs to keep in mind, but it is important not to sacrifice relevance by having evaluation findings be delivered too late to matter. Of course, considering evaluation early as part of strategy development will help define when specific information will be needed.

Consider the following set of questions in preparing a timeline for evaluation—one that includes important dates, decisions, and a cushion for inevitable lags. If we want to inform foundation decisions, what is our timetable for receiving at least preliminary results? How rigid is that timetable? Backing up from there, when would we need to have results in order to make sense of them and to bring them forward for funding considerations? Backing up again, how much time do we need to find the right evaluator and give the evaluator enough time to design an effective evaluation, then gather, analyze, synthesize, and bring those results to us? If we want actionable information, it is essential to grapple with what is knowable in what time frame. If we are aiming to inform grantees, how might their budgets or program planning cycles affect the evaluation timetable? Grantees also need time to make sense of findings and act upon them. If our purpose is to inform the field, or to engage other potential funders in an area, are there seminal meetings or conversations that we want an evaluation to influence? Are there planning processes or budget cycles that might be important to consider in our evaluation planning?

Be sure to consider how others in the foundation—program peers, the evaluation officer, communications staff, and at certain points grants management and legal—would also be helpful or necessary. For example, if evaluation questions refer to legislation, ballot initiatives, or campaigns and elections, or if evaluators will interview U.S. or foreign legislators, you must talk with legal before getting started. Leave a couple of weeks for contracting and contract review!

Write RFP, Find Evaluator, and Contract	Design and Data Collection	Discuss and Interpret Findings	**KEY JUNCTURE!** Apply the Findings in Practice	Sharing and Follow-Up
3-4 MONTHS	**3-4 MONTHS**	**3-4 MONTHS**		**3-4 MONTHS**

DEFINING EVALUATION QUESTIONS

Our evaluations begin with and are guided by clear, crisp questions. **Crafting a short list of precise evaluative questions increases the odds of receiving helpful answers—and a useful evaluation.** It also increases the odds that evaluation will support learning, because the program officer (and grantees) can ask questions that will be most informative for their learning. Well-designed questions can not only clarify the expected results, but also surface assumptions about design, causality, time frame for results, and data collection possibilities. These surfaced assumptions and questions can then help sharpen a theory of change and ensure effective planning for evaluation and learning.

Unfortunately, many evaluations begin to go awry when questions are drafted. It is useful to start by distinguishing between the following areas of inquiry. **Although not every evaluation should seek to answer this full range of questions, the categories below offer suggestions for effective investigation, depending on the nature and maturity of the strategy or area of interest selected for evaluation.** These are general examples of questions and should be tailored to be more precise.

- *Implementation:* How well did we and our grantees execute on our respective responsibilities? What factors contributed to the quality of implementation? *In much of the social sector, evidence shows that most program strategies and program interventions fail in execution. This makes evaluating implementation very important for driving improvement, understanding the ingredients of a successful or failed approach, and replicating or adapting approaches over time.*

- *Outcomes:* What changes have occurred and why? If not, why not? How do the changes compare with what we expected and the assumptions we made? To what extent and why are some people and places exhibiting more or less change? To what extent is the relationship between implementation and outcomes what was expected? *To be able to answer these questions, it is enormously helpful to have planned an evaluation from the outset so that measurements are in place and changes are tracked over time. While we tend to use implementation markers to track progress toward outcomes, we use evaluation to analyze more systematically underlying issues related to what caused or contributed to changes in these outcomes, why, why not, and for whom.*

- *Impact:* What are the longer-term sustainable changes? To what extent can these changes be attributed to the funded work or could the work be determined to have contributed to these impacts? *Impact questions are typically the most complex and costly to answer, particularly for much of the field-building, systems-level and research- and advocacy-related work that we fund.*

- *Context:* How is the (political, policy, funding, country) landscape changing? Have changes in the world around us played an enabling or inhibiting role in the ability to effect change? *Often our theories of change involve assumptions about how the world around us will behave, and unanticipated events—conflicts, new governments, social protests, disease, technological or scientific breakthroughs— can accelerate or slow progress toward long-term goals. Understanding these interplays can help us avoid false conclusions.*

- *Overall Strategy and Theory of Change:* Did our basic assumptions turn out to be true, and is change happening in the way we expected? *In order to answer questions about the overall strategy, it is likely that you will draw from more than one evaluation—with findings synthesized in order to gain deeper insight. It is helpful to develop overarching evaluation questions early on in the strategy process, however, to ensure that you are gathering the pertinent information you may need from each.*

What can you do to make these general evaluation questions more precise and evaluative?

- Make more specific (e.g., be more specific about what is meant by a word or phrase)

- Tie more closely to intended use (e.g., add a note about why answering this question will be helpful and for whom)

- Make more realistic (e.g., add time frame, location; narrow scope)

- Add "compared to" as part of the question (e.g., compared to other approaches, compared to where we expected)

- Ask to what extent? why/why not? How? For whom? (rather than just "did x happen?")

- *Special Case: Evaluating a Regranting Intermediary. This can require an additional set of questions:* How and to what extent is the intermediary adding value to its grantees? Is it just a go-between, supporting the transaction of regranting funds without adding significant additional value? Or is the intermediary able to offer important technical assistance to organizations by virtue of being closer to the ground? Where and for whom is the intermediary adding the most value, and where is it adding the least? What are the enablers and inhibitors to an intermediary's high performance? How does this intermediary's performance compare to other intermediaries? How transparent, efficient, and well managed is the subgranting process and related communications, and what is the subgrantees' experience? Did they get clear communications from the intermediary, or support to figure out how to prepare budgets that reflected their full costs?

- *DEI Issues: When we consider issues of diversity, equity, and inclusion in our strategies, we should also consider how DEI shows up in our evaluation questions.* Include questions to understand the perspectives and insights of those whose voices are least heard. As a hypothetical example, a gender-blind evaluation may ask: To what extent were the goals of the program met? Why or why not? A gender-inclusive evaluation may instead ask: To what extent were the goals of the program met—and how did the effects differ among males, females, and others? When changing systems that drive inequity is part of the strategy, then the evaluation might ask questions about whether and how those systems have changed, why, why not, and for whom? **At the very least, include questions about whether there were any unintended consequences—and for whom.**

**Integrating Diversity, Equity, and Inclusion into the
2018 Western Conservation Strategy Evaluation and Refresh**

The Western Conservation strategy's long-term goal is the preservation of biodiversity and the conservation of the ecological integrity of the North American West for wildlife and people. In preparing for its most recent strategy refresh, program staff commissioned evaluations to assess the strategy's short-term, time-bound initiatives such as California Drought and Canadian Boreal Forest, as well as an evaluation of the overall strategy.[35]

Among other evaluation questions about progress, successes, and challenges, the team included questions related to issues of diversity, equity, and inclusion. The team sought an evaluation that would: (a) unpack the foundation's blind spots related to the diversity of the current Western Conservation grantee portfolio; (b) identify the relationship of these DEI values to the strategy's policy objectives; and (c) recommend how the Hewlett Foundation could be more deliberate about supporting grantees led by and serving communities of color, and those working to make greater progress by embracing the values of equity and inclusion within their organizations and in how they approach their work in the world.

The strategy-level evaluation paid careful attention to soliciting diverse perspectives. For example, the evaluators talked to Hewlett Foundation grantees, farmers and ranchers, tribal governments, sportsmen and women, Latino and African-American organizations, faith communities, and youth and included their direct feedback, along with other qualitative and quantitative data. Paying specific attention to whom they were hearing from—with foresight, time, and intentionality—proved valuable for providing new insights.

Perhaps most important among the many lessons from this intensive evaluation process was new understanding of the critical role of inclusivity in securing lasting conservation outcomes. One ah-ha moment for the team from the evaluation: Equity and inclusion efforts must be paramount in the grantmaking strategy and precede a diversity push in order to intentionally signal to the field their importance and to avoid tokenism in hiring and outreach strategies (which might come from a focus on diversity alone). The evaluation findings also reaffirmed the value of diversity, equity, and inclusion as "not simply a moral issue, but a policy-making imperative." Building on these findings, the new strategy argues that to endure the winds of political change, conservation solutions must be place-based and account for the diverse cultural, economic, social, and ecological needs of a community, which requires engaging a broader range of stakeholders and constituencies in the development and defense of conservation solutions. Today, the Western Conservation grantmaking portfolio and strategy[36] reflect a vision of a more inclusive conservation movement for the long-term benefit of western communities, ecosystems, and wildlife.

'Teacher as Learner' Initiative Lessons on Crafting Precise Evaluation Questions

Imagine that we are supporting a new initiative called "Teacher as Learner" that aims to improve the quality of teaching and learning for students of all backgrounds in different places via a network of 100 self-organized groups called "communities of practice." Each group of local teachers is professionally facilitated and focused on their specific capacity needs. Having organized themselves around issues of local importance, the communities of practice draw on regional resources as needed. The initiative's key assumption, based on some evidence in other fields, is that a blend of professional support and local ownership will lead to improved outcomes. If this approach seems successful after an initial period of innovation, we might develop an experiment to rigorously assess impact.

What are our evaluation questions?

POOR SAMPLE QUESTION

Was the Teacher as Learner theory of change successful?

This question has limited value for several reasons. **First, it is vague.** Usually a theory of change has multiple dimensions and contains many assumptions about how change will happen. A useful evaluation question is explicit about which interventions and assumptions it is exploring or interrogating. A vague question gives the evaluator too much discretion. This often sets us up for potential disappointment with the findings when we receive an evaluation report that is not useful and does not answer questions of importance to us.

A second, related, point: it is unclear whether the question is aimed at issues of execution (e.g., Did x happen?) or issues related to the "causal chain" of events (e.g., If x happened, did it catalyze y?). It is often useful in an evaluation to look at execution and outcomes with a distinct focus, as well as the relationship between them.

Third, the definition of success is unclear, allowing the evaluator too much discretion. Does success mean that 80 percent of what we hoped for happened? What if 60 percent happened? What if two out of three components progressed exactly as planned, but a third, delayed by an unforeseen personnel challenge, has not yet been implemented? Asking a dichotomous Yes/No question about an unspecified notion of "success" will be less helpful than a few focused questions that precisely probe what we want to learn and anticipate how we might use the answers.

GOOD SAMPLE QUESTIONS

About implementation:

1. How and to what extent did the Teacher as Learner initiative create a network of local, self-organized communities of practice?

2. What was the nature of the variation in areas of focus for the communities of practice?

About intermediate outcomes:

3. To what extent did teachers adopt or adapt improved teaching methods after participating in the communities of practice?

4. What were the key factors that enabled or inhibited teachers from adopting new teaching methods?

About outcomes:

5. In what ways and by how much did these teachers' students improve their learning?

6. Is there any variation in students' learning gains— for example, by gender, or by students with disability? If so, what are possible explanations for that variation (including student, teacher, or community characteristics, and features and approaches used in the communities of practice)?

Why are these better questions? As a valuable beginning, they break one vague question about success into clear, specific ones that generate insight about different steps in the initiative's causal chain: which parts may be working well and as expected, which less well, and possible explanations for this. They give more direction to the evaluator about our specific areas of interest. And, although they still need to be elaborated on with specific measurement indicators and methods of data collection, they are designed to generate data that can be used to correct course.

IDENTIFYING METHODS

Most strong evaluations use multiple methods to collect and analyze data. The process of triangulation allows one method's strengths to complement the weaknesses of another. For example, randomized experiments can determine whether a certain outcome can be attributed to an intervention, but complementary qualitative methods are also needed to answer questions about how and why an intervention did or didn't work—questions that are central to replication or expansion.

Multiple methods help reduce bias, as does active consideration of how the methods are applied. For instance, if an evaluation is primarily based on qualitative key informant interviews, it will be important to include respondents who are not cheerleaders but may offer constructive critiques.

The evaluator should be primarily responsible for identifying the methods, but it is helpful to set expectations that the evaluation will include **multiple methods and capture diverse perspectives,** and that it will use both qualitative and quantitative data collection.

Grantees are good sources regarding methods. Once an evaluator is selected, the evaluator should review data collection procedures and protocols with (at least some) grantees in order to assure consistency and applicability. Sometimes, grantees might give input on methods: for example, if they are aware that a certain methodology would prove ineffective or the language in an interview or survey might need adjustment.

Our goal is to maximize rigor without compromising relevance. While most evaluations of our strategies cannot definitively attribute impact to the funded work, the essence of good evaluation involves some comparison—against expectations, over time, and across types of interventions, organizations, populations, or regions. Even when there is no formal counterfactual (i.e., example of what happened or would have happened without the strategy), it can be helpful to engage in discussion of what else might be happening to explain findings, in order to challenge easy interpretations of data and consider alternative explanations.

Evaluators should be expected to take a systematic approach to causal inference. At the very least, this includes checking that the evidence is consistent with the theory of change and identifying alternative explanations to see if they can be ruled out as the cause of any observable results. Given the nature and complexity of many Hewlett Foundation strategies, it is likely that evaluators will need to identify noncounterfactual evaluation approaches that go beyond simple comparison. For example, contribution analysis[37], process tracing[38], qualitative comparative analysis[39], and QuiP[40] are techniques that have been developed to do this. It is not necessary for program staff to know the details of these approaches, but it can be helpful to surface in discussions with potential evaluators why they are suggesting a specific approach. A good resource for learning about different types of evaluation is BetterEvaluation.org.

The essence of good evaluation involves some comparison—against expectations, over time, and across types of interventions, organizations, populations, or regions.

Increasing Rigor and Relevance

In 2015, the Performing Arts program—which aims to sustain artistic expression and encourage public engagement in the arts in the Bay Area--commissioned a midpoint evaluation of their strategy.[41]

Two key questions in commissioning the evaluation were: which geographic and demographic communities have benefitted from Hewlett Foundation support, and where are the gaps? Data from an audience research project the program had previously supported for a group of grantees—the Audience Research Collaborative (ARC)—proved very informative for addressing this question. The evaluators were able to compare the demographics of the audiences of the grantees to the broader demographics, using census data for the area. As a result, the Performing Arts program could see where they had strengths and weaknesses, and where they could diversify their portfolio to better reach the participants and artists they hoped to reach. Since they had anticipated early on that demographic data would be valuable, they were able to build in a comparison point as part of their evaluation.

It's important to note that the data collection strategy was not without its limitations—which is the case with all evaluations. For example the survey methods used may have introduced bias towards more white, female and highly-educated respondents. What's more, U.S. Census categories themselves have not kept pace with rapid demographic change, and don't reflect the true diversity of our society, something many participants in ARC addressed by collecting data about both the census categories and expanded categories that better reflect the communities they serve (for gender identity, for example). Nevertheless, the findings were valuable for the program and the planning and foresight paid off—and they went in with eyes open to the limitations.

As noted in the section on purpose, **it is worth checking with intended users** (**including yourself as commissioner**) **to understand what degree of rigor is necessary for the findings to be useful for those who are in the position to use and make changes based on the findings.** An evaluation is worth doing only to the extent that it is going to affect decisions or challenge beliefs. In some cases, this means the strength of evidence needed will be time-consuming and costly to obtain. In other cases, the users might agree that less evidence is necessary to inform course correction or decision making.

Be prepared to discuss with the evaluator how the data will be gathered, analyzed, and shared with regard to confidentiality. Will the data be kept confidential, with no names or unique identifiers attached? Will grantee organizations be identified? There are pros and cons to different types of data gathering. If an evaluator tells the respondent the information gathered will be kept confidential, they cannot then turn around and share the name of the grantee or others who responded a specific way. If the information is only going to be useful with identifiers attached, then the pros of gathering it that way may outweigh the potential cons of too much courtesy bias.

CRAFTING AN RFP FOR AN EVALUATOR

Once you have identified the purpose and an initial set of evaluation questions, it is time to identify a third-party evaluator. Start by crafting a thorough request for proposal (RFP). Evaluations chosen through competitive selection processes—even if only involving conversations with two or three different evaluators, rather than a formal paper proposal process—tend to offer greater opportunity to find an evaluator that will make the best partner for the evaluation project. At a minimum, if an evaluator is chosen without an RFP (perhaps in cases where the evaluator's work is already well known to the person commissioning the evaluation, or the evaluation is a continuation of earlier evaluation work), create an evaluation purpose document for discussion with the evaluator. **Establishing clarity about**

the expectations for the project (**on all sides**) **helps to make sure that all parties are in agreement regarding the purpose, approach, roles, and time commitments.**

The basic elements of an RFP to engage an evaluator include background information about the strategy, substrategy, cluster, or grantee; background information about the evaluation, its purpose, intended audiences, and anticipated use; key evaluation questions; known available data sources; time frame for receiving results; preferred deadline for the deliverables and types of deliverables required (including internal foundation, external grantee, field, and public-facing deliverables); evaluator qualifications and roles/expectations; and evaluation budget (amount of available funding).

Include language in the RFPs, and ultimately the contract scope of work, so that the evaluator is aware of the foundation's expectation that there will be a product that will be shared publicly.

During this planning phase, grantees can suggest evaluation questions, weigh in on terms of reference, and help identify potential evaluation consultants. (See Appendix C.) Some program staff have engaged grantees in this part of the process by circulating draft scoping documents that summarize purpose, key evaluation questions, and proposed timelines for data collection, synthesis, and reporting. Grantees will likely offer useful feedback on opportunities or challenges for timing the collection of data.

CONSIDERING WHETHER OR NOT—AND HOW—TO HAVE THE EVALUATOR OFFER RECOMMENDATIONS

One important area to be clear on before beginning an evaluation is whether you want the evaluator to include recommendations along with the findings. **Sometimes asking an evaluator not to provide recommendations works best, since the evaluator may not be in a position to truly understand the context within which program staff will be making strategic decisions.** In fact, we have found that recommendations can sometimes be counterproductive, because if they are off-base or naive, they can undermine the credibility of the overall evaluation.

CHOOSING AN EVALUATOR AND DEVELOPING AN AGREEMENT

The ideal evaluator is strong technically (typically in social science research techniques), has subject matter expertise, is pragmatic, and communicates well, both verbally and in writing—whether about good or bad news. Cultural awareness and sensitivity to the context in which nonprofits are operating are also very important, as is the ability to work well with grantees and to find ways to communicate results in ways that are useful. If we cannot find that full package, it is sometimes appropriate to broker a relationship and bring together people or teams with comple-

> During this planning phase, grantees can suggest evaluation questions, weigh in on terms of reference, and help identify potential evaluation consultants.

mentary skills. For example, when commissioning the evaluations of our Cyber and Nuclear Security initiatives, pairing firms that had technical expertise in evaluation with specialists in these respective fields proved valuable. Choices always involve trade-offs; it is important to manage their risks. If we arrange the partnership, are we prepared for the extra time it will take for the partners to collaborate? Are we and the partners clear on what roles each will play, and are the roles well defined and clear?

Part of our commitment to diversity, equity, and inclusion includes consideration for the evaluation team with whom we work. When selecting an evaluator, build in enough time to look for candidates from a broad pool of qualified applicants with diverse backgrounds and experiences, and reflect on the evaluation team's ability to draw on knowledge of local context.

Grantees can be helpful in the evaluator selection process. Including grantees not only may help get them invested in the effort, but they also often contribute a useful pragmatic perspective. At the very least, it is important to introduce a selected evaluator to grantees and let them know of the time commitment and expectations for the evaluation—and what they can expect in terms of their involvement.

Getting the Right Evaluator/Evaluation Team: Technical, Context, Strategy, Content, and Other Considerations

The Cyber Initiative seeks to cultivate a field that develops thoughtful, multidisciplinary solutions to complex cyber challenges and catalyzes better policy outcomes for the benefit of society. A couple of years into the strategy, the Cyber team commissioned an evaluation[42] of its nascent effort to foster a network of cyber policy experts. They selected it as an evaluation focus area because of its centrality to the success of the overall strategy and because it offered an early opportunity for learning and course correction.

As the team put together a request for proposals, they crafted a set of evaluation team criteria. Subject matter knowledge of cyber policy was critical for this project, as was experience with evaluation and strategy development, so the team invited proposals that would incorporate partnerships between consultants.

In the end, they selected a proposal in which two firms partnered—one with deep evaluation expertise and the other with deep cybersecurity policy knowledge—enabling the evaluation to have the degree of rigor an evaluator brings, along with cyber content knowledge.

If the evaluator doesn't understand the work, there can be serious ramifications for building trust, accessing relevant subject matter experts, recognizing concepts, and determining appropriate evaluation designs. If the evaluator is exclusively a content expert, there can be serious consequences—predetermined ideas about how things should work, a lack of independence, and frequently little to no evaluation technical expertise.

Selecting an Evaluator

Selecting the right evaluator is hugely important to the success of your evaluation. It's no easy task—an evaluator is charged with possessing the right mix of evaluation technical expertise, content and context knowledge, and cultural competence. The evaluator should understand how to convey evaluation findings to you, the client, in the ways that work best for you. Of course, you have a role to play in making that all work—but it's important to select the right partner. There are three tips that might help you:

- First, think through what qualities are likely to make an evaluation team be successful given your evaluation questions, time frame, and the context within which the strategy is situated.

- Second, talk to more than one evaluation team to understand the variety of approaches they bring to addressing the evaluation questions and collecting and analyzing data. Regardless of whether your evaluator is chosen competitively, make sure to conduct an interview with them. Interviews can help you feel out whether the evaluation team is a good match for your needs.

- And, finally, one idea is to do a trial run: Hire an evaluator to do just part of the evaluation process, such as the design phase. If both parties feel the partnership is working, you can extend the engagement. If you don't benefit from working with together, you can cut your losses early.

Implementation

Sometimes an evaluation's implementation does not go precisely as planned. Staying connected with the evaluator and the evaluation during implementation can go a long way toward ensuring responsiveness and a generally higher-quality evaluation.

MANAGING AND STAYING INVOLVED WITH AN EVALUATION

As mentioned above, less staff time is usually required during implementation of an evaluation, while evaluators are collecting data in the field. Ongoing management of their work takes some time but, in general, less than during planning and use. That said, active management is essential. Talk with the evaluator regularly and ask what is, or is not, going well. What are they finding? Are the findings making sense? Are there data missing, or might the data interpretation be off or incomplete due to the complexities of the strategy or other issues?

Request periodic updates to document progress and any obstacles the evaluator is facing in data collection, data quality, or other areas. These exchanges can be useful forcing functions to keep an evaluation on track and to start troubleshooting early. Often the data collection in an evaluation mirrors some of the challenges faced by a program in other facets of its work, so evaluation progress updates can be helpful in many ways.

Some program staff review and provide feedback on evaluation tools. This can be a useful way to ensure that everyone is on the same page about what data will be gathered and why.

Support the Evaluator's Success

As the evaluation commissioner, program staff have a significant role in the success of an evaluation: whether and how the evaluation ultimately provides information that will be used and shared. We hire independent third-party evaluators. Therefore, it is essential for program staff to:

- Provide the important background information, contextual information, and introductions to grantees or other key stakeholders, to give the evaluators a good chance to gain the requisite knowledge to proceed effectively. Help them understand the culture of the foundation and the approach to strategy, grantmaking, and evaluation. For example, share these Evaluation Principles and Practices, and the Outcome-Focused Philanthropy Guidance.

- Give the evaluator enough time to dig into the background information, finalize the evaluation design, collect data, analyze, and interpret—and the time and attention to get your feedback. It might have taken you a while to plan for the evaluation and to hire the evaluators. Allow them the needed time to carry out the evaluation.

- Keep the evaluators apprised of changes to the strategy, new information about grantees, or issues that develop that may affect either the evaluation design or the interpretation of the findings.

Keep in mind: Your evaluators should be asking you for information and feedback as they proceed. These kinds of conversations ensure that the evaluation is on the right track. Don't let too much time go by before you have a conversation about what types of information sharing will be most helpful.

It can be especially useful to set an expectation of baseline data summaries, or interim evaluation reports on preliminary findings. This will keep an evaluation on course, engage foundation staff in the discussion of early findings, and make space for any needed course corrections. For example, as part of the evaluation of the Fund for Shared Insight,[43] the evaluator has produced numerous products "along the way" that have been helpful for discussions with funders, grantees, and prospective funders. The evaluations of the Transparency, Participation and Accountability and Supporting Local Advocacy in Sub-Saharan Africa strategies similarly have provided materials such as "baseline" summaries and annual reports of progress, which prompt topics for discussion at convenings.

Make sure to take the time to provide updates to the evaluator, so they are informed and can adjust. In cases where the strategy is evolving and the evaluation is being conducted alongside that evolution, it is important for program staff to provide evaluators with timely updates on relevant developments that may affect either the evaluation design or the interpretation of the findings.

As important as managing the process is talking through the findings, early and often. What do they mean? Do they resonate? Is the evaluator fully aware of the context? Is there any reason to make changes to the data collection plan or methodology to capture lessons on emerging areas of interest? Here you can consider whether an advisory committee would be worthwhile—to bring in others to the discussion of findings and to consider alternative perspectives on how the findings might be used.

Using an Advisory Committee to Build Buy-In and Enhance Practicality, Quality, and Use

Advisory committees are a great way to engage others—funders, grantees, other external stakeholders, and others internally—in an evaluation. Developing and using an advisory committee has clear benefits: In particular, it can help increase the likelihood that the findings are used by and shared more quickly with intended audiences.

1. **Who?** You should think critically about who is on the committee and what role they play. Which funders, grantees, and others do you want to have early access to your findings? Who can give input on making the evaluation more practical and useful? Whose perspectives or voices might be left out?

2. **Why?** You can choose your members to serve various purposes. You may want to get buy-in from some constituents or feedback on the level of rigor needed to be convincing—this is a way to do it. Or sharing internally may be a priority, so engaging others internally may help.

3. **How?** Remember: You determine how and when you want them to engage. Typical times are when discussing the design of the evaluation, early findings (so they can be your test audience/sounding-board), and recommendations and dissemination. Also, be mindful of how much you are asking of them; consider an honorarium.

OTHER CONSIDERATIONS

The advisory committee will need management, so it is important to figure out whether this will be part of the evaluator's responsibility and paid for in the evaluation contract, or whether the program officer will be responsible. If the program officer is responsible, be aware that the management takes additional time.

Continue to check in with and engage grantees in the evaluation. A reviewer of this guide said, "The relationship between the evaluators and the implementers is KEY" to successfully conducting an evaluation and applying the findings in practice. If grantees are engaged in the evaluations that touch them, they will be: (a) more supportive with respect to data collection; (b) more likely to learn something that will improve their work; (c) less likely to dismiss or defend against the evaluation; and (d) better able to help strengthen the evaluation design, especially if engaged early. From a design perspective, this last point is quite important: Grantees can serve as a reality check and deepen understanding of the available data and data collection systems. They might also suggest potential respondents for interviews, or data that may (or may not) be available from their own or others' monitoring systems. As part of the program staff's evaluation management responsibilities, it is helpful to check in with grantees about how the evaluation is going for them, to get feedback on the process and its strengths and challenges. Another idea is to create an evaluation advisory committee that includes representatives from grantee organizations to advise on aspects of the evaluation.

RESPONDING TO CHALLENGES

Any number of challenges can emerge during an evaluation: A data source may be less reliable than predicted; survey response rates may be too low to draw conclusions; or consultant staff turnover in the selected firm may reduce confidence in the evaluation team.

If you hit these bumps or others in the evaluation road, it is important to pause, take stock of the challenges, revisit prior plans, consult appropriate stakeholders, consider alternative solutions, and make necessary course corrections. Don't forget to communicate any changes to everyone invested in the work, including grantees.

Use *(Including Interpreting and Sharing Findings)*

Our first evaluation principle is "We lead with purpose" for a reason: Establishing a clear purpose—including audience, use, and a timeline—sets us up to maximize the usefulness of the evaluations and to live by another of our established principles: "We use the data!" Not using our findings is a missed opportunity for learning, improvement, and course correction—and a waste of time, energy, and resources. And yet, using the data requires additional effort, including active involvement on the part of the evaluation's intended users, and time to reflect and make meaning of the findings. We optimize use by sharing the findings using a variety of formats and communication approaches to reach diverse audiences. Engaging with groups to discuss shared findings, taking time to discuss and interpret findings from the evaluation as it progresses, and asking yourself "Now what?" go a long way to supporting use.

From the very beginning of an evaluation process, it is important to plan how the results will be used; **along the way, it is wise to remind yourself of those intended uses.**

As noted earlier, our evaluations tend to have a primary dual purpose of informing Hewlett Foundation staff and grantees, and sometimes informing the field. Common uses within those categories include informing:

- strategy-level decisions
  (*making course corrections, ramping up or strengthening aspects of a strategy, testing assumptions, or exiting aspects of a strategy*)

- future evaluations
  (*commissioning smaller substrategy or cluster evaluations as inputs to inform a larger, strategy-level summative evaluation, establishing a baseline, or helping to refine targets for change*)

- process improvements
  (*improving data collection, grantee proposal or reporting practices, and "beyond the grant dollar" activities, such as convenings and information sharing*)

- grant and grantee-level decisions
  (*helping to shape renewals of grants, closing a grant, developing OE grant opportunities, switching from project grants to general operating support*)

- other uses
  (*summing up lessons learned as we exit a strategy or substrategy, developing frameworks for evaluation and assessment that inform other strategies at the foundation, or engaging other funders in discussions of findings*)

- grantees' decisions
  (*for their own program improvement*)

- field use
  (*others learning from what we are doing and how*)

Using results is often messier than anticipated. Sometimes staff expect more confirmation of success—or for an evaluation to uncover more surprises or "aha" moments for them—than an evaluation typically delivers. Sometimes an evaluation is not especially well done, and the results inspire limited confidence. Other times, staff simply are not sure how to apply the lessons. They are uncertain how best to shift or overhaul a strategy or substrategy.

Use requires that teams pause to reflect and grapple with all of the "What? So what? Now what?" questions. Evaluators often provide the "what" in terms of the findings, but the program teams need to grapple with the "so what" and "now what," in order to make sure those findings are used. This takes dedicated time and effort.

Grantees, because of their knowledge of content and context, play an important role in verifying accuracy and helping interpret and make meaning of the findings. Program staff can support use and sharing by convening grantees to discuss findings and sometimes engaging them in a process of co-creating recommendations. Or staff can meet with grantees at key junctures, when reflection on findings can serve to stimulate questions, ideas, and opportunities for improvement.

Sharing what we learn is essential not only at the conclusion of an evaluation but throughout the process. Sharing is not only a way to ensure that our evaluations maximize their utility, it is also consistent with our foundation values and principles of openness and transparency. Every program team should plan to publicly share findings from every evaluation commissioned, in some form.

Issues of diversity, equity, and inclusion are relevant when discussing, interpreting, and sharing findings. Through what lens are the evaluation results interpreted, used, and shared? Who is involved in the discussion? If recommendations are made, how do these factors come into play? Is sharing done in a variety of formats and made accessible to multiple groups?

PAUSE TO REFLECT

- **WHAT**
  What did we try with the strategy? What results are we seeing?

- **SO WHAT**
  What seemed to drive those results (positive and negative)?

- **NOW WHAT**
  What do we take away from that? How do we apply what we've learned going forward?

Adapted from *Adaptive action: Leveraging uncertainty in our organization.*[44]

Some Ways to Share (Internally and Externally)

- Convene grantees to discuss the results and recommendations

- Organize an internal briefing (e.g., at a Shoptalk or All Staff) to share with your colleagues what you've learned, both about your strategy projects and the evaluation process itself

- Discuss with your program's board advisory committee how the evaluation results will affirm or inform changes in your strategy or grantmaking approach

- Share a version of the evaluation with targeted external audiences, accompanied by a memo detailing how you are applying the findings in practice

SHARING RESULTS INTERNALLY

Sharing the results of an evaluation with foundation colleagues brings many benefits, and it is worthwhile to build this step into your process. For staff managing an evaluation, these discussions can crystallize the results, lead to a deeper understanding of those results, and force some grappling with what is not yet understood. It can also help staff think through what the results mean programmatically and how to apply them in practice. For members of other teams, review of the evaluation results can generate insights about their own programs, grantmaking approaches, or evaluation designs.

An internal debrief at the end of each evaluation to discuss key lessons learned, what went well and what did not, and actions taken will help advance the foundation's evaluation practice and keep us focused on designing evaluations with action in mind. If another funder has collaboratively supported an evaluation, it is often appropriate to consider that partner an internal colleague with respect to sharing results and surfacing implications.

Sharing When Findings are Not All Positive

Most times we commission an evaluation, we ask evaluative questions to help us and grantees understand both successes and challenges. Yet, there are times when the findings are more negative than we expected. What do we do in those circumstances? Consider the following:

- The evaluation officer and communications teams are here to help. They can help you plan, from the beginning, what and how to share, to address sensitivities and protect reputations—so that we share lessons learned in the most appropriate and effective ways.

- There are external resources that can help you. This article,[45] by BetterEvaluation, is a good place to start. It includes tips for when you might be in this situation—such as using a participatory approach from the start, limiting surprises, discussing the possibility of negative results from the start, framing as lessons learned, and considering ways to overcome the obstacles that are surfaced—but also emphasizes the need to fully report all findings truthfully, even negative ones.

SHARING RESULTS EXTERNALLY

Our intention is to share evaluation results—positive, negative, and in-between—so that others may learn from them. Out of respect, we communicate with our grantees early on about our intention to share our evaluations, and we listen to any concerns they may have about confidentiality. Grant agreement letters specify an organization's likelihood of participating in an evaluation (which, as noted above, are not typically of just one particular grantee but are usually of numerous grantees who are part of a strategy, substrategy, or cluster of grants). As an evaluator comes on board, it is important to clarify and share with grantees the evaluation's purpose (including any anticipated effect on the grantee), the process for making decisions about it, and each party's required role and participation. It's also important when possible to come to an agreement regarding the level of findings (full evaluation results, an executive summary, or a presentation) that will be shared, with which audience. You might also negotiate that individualized results will be given to grantee organizations and general results will be shared with a cohort and with selected audiences or publicly.

We consider the question of in what form we will be share evaluation findings on a case-by-case basis, with care given to issues of organizational confidentiality. For instance, if an evaluation is in part focused on questions of organizational development, it may be more useful for the findings to be shared in full with that grantee, so the grantee can use the results to drive improvement without having to take a defensive public stance, and also to work with the evaluator to surface broader lessons to be shared publicly.

The foundation expects that some product—whether it's a full evaluation report, an executive summary, or a presentation—from the process will be made public to support openness, transparency, and learning. When planning how to share results publicly, program staff should consult with the foundation's communications staff—ideally, early in the process and over the course of the evaluation—to determine the best approach. They should review documents for sensitive issues and flag those for legal review before sharing results publicly.

Sharing Evaluation Findings in Ways that Work Well for Different Audiences

The Knowledge for Better Philanthropy strategy supports the creation and dissemination of knowledge about how to do philanthropy well. From an earlier evaluation, the program team learned that the grantees were producing high-quality knowledge and distributing it widely. But they were missing key pieces of information: how foundations find knowledge resources, and whether these knowledge products inform or influence their philanthropic practice. These pieces are central to the strategy but had never been systematically assessed. As the philanthropy grantmaking team embarked on this new evaluation, they took time to ask the grantees what they would like to learn as part of the evaluation, included their questions where possible, involved them in an evaluation advisory committee, and established a process for sharing the findings.

The evaluators produced a summary report[46] highlighting key findings. But the team did not stop there. First, the program officer hired a graphic design firm to help translate the report findings into easily digestible bites.[47] This was, in fact, a finding from the study itself: Funders prefer easily digestible formats that are practically applicable and relevant to their work. These resulted in easily shareable, visually friendly snapshots of the data. Second, the program officer ensured that the grantees who were included in the study were also able to make best use of the work. To do so, each grantee received a confidential, individualized report, which included that organization's data as compared to others' (presented anonymously). These two extra steps—making the work more visually accessible, and relevant reporting to grantees who participated—led a grantee to publish this commendation[48]: Finally, a Foundation Commissioned Study Which Actually Helps its Grantees.

Doing this was not free of cost. To prepare both the public and the individualized reports cost more time and more money. It also required both upfront planning and some level of flexibility. The team didn't know exactly how they hoped to share the findings right at the start, but they built in the financial and timeline cushion to explore. They ultimately had to amend their contract with the evaluators to make this possible—but to the grantees involved in the Knowledge work and the broader field, these steps made the report more useful and relevant.

The program officer also looped in the Hewlett Foundation communications officer, who was able to provide input in a timely way about effective email, web, and social sharing.

Key Terms

ACTIVITIES The actions taken by the foundation or a grantee to achieve intermediate outcomes and make progress toward the achievement of goals. *[Gates Foundation glossary]*

BASELINE An analysis or description of the situation prior to an intervention, against which progress can be assessed or comparisons made. *[Gates Foundation glossary]*

EVALUATION An independent, systematic investigation into how, why, to what extent, and for whom outcomes or goals are achieved. It can help the foundation answer key questions about strategy, substrategy, clusters of grants, or sometimes a single grant. *[Variant of Gates Foundation glossary]*

DEVELOPMENTAL EVALUATION A "learn-by-doing" evaluative process that has the purpose of helping develop an innovation, intervention, or program. The evaluator typically becomes part of the design team, fully participating in decisions and facilitating discussion through the use of evaluative questions and data. *[Variant of The Encyclopedia of Evaluation (Mathison, 2005) and Developmental Evaluation (Quinn Patton, 2011)]*

FORMATIVE EVALUATION An evaluation that occurs during a grant, initiative, or strategy to assess how things are working while plans are still being developed and implementation is ongoing. *[Gates Foundation glossary]*

IMPACT EVALUATION A type of evaluation design that assesses the changes that can be attributed to a particular intervention. It is based on models of cause and effect and requires a credible counterfactual (sometimes referred to as a control group or comparison group) to control for factors other than the intervention that might account for the observed change. *[Gates Foundation glossary; USAID Evaluation Policy]*

PERFORMANCE EVALUATION A type of evaluation design that focuses on descriptive or normative questions. It often incorporates before/after comparisons and generally lacks a rigorously defined counterfactual. *[USAID Evaluation Policy]*

SUMMATIVE EVALUATION An evaluation that occurs after a grant or intervention is complete, or in service of summing up lessons to inform a strategy refresh or when exiting a strategy, in order to fully assess overall achievements and shortcomings. *[Variant of Gates Foundation glossary]*

EVIDENCE   A general term that refers to qualitative and quantitative data that can inform a decision.

FEEDBACK   Data about the experience of the people who we hope will ultimately be positively touched by our work. Feedback can provide new information and insight that can be a valuable source for learning; while it may inform evaluation, it is a distinct channel.

GOAL   A general statement of what we want to achieve; our aspiration for the work. *[OFP Guidebook]*

OUTCOME   Specific change we hope to see in furtherance of the goal.

**IMPLEMENTATION   A catch-all term referring to particular activities,
MARKER**   developments, or events (internal or external) that are useful measures of progress toward our outcomes and goal.

GRANT   A sum of money used to fund a specific project, program, or organization, as specified by the terms of the grant award.

INDICATORS   Quantitative or qualitative variables that specify results for a particular strategy, component, initiative, subcomponent, cluster, or grantee. *[Gates Foundation glossary]*

INITIATIVE   A time-bound area of work at the foundation with a discrete strategy and goals. Initiatives reside within a program, despite occasionally having goals distinct from it (e.g., the Drought Initiative within the Environment Program).

INPUTS   The resources used to implement activities. *[Gates Foundation glossary]*

LOGIC MODEL   A visual graphic that shows the sequence of activities and outcomes that lead to goal achievement. *[OFP Overview]*

METRICS   Measurements that help track progress.

MONITORING   A process that helps keep track of and describe progress toward some change we want to see—our goals or outcomes. Evaluation will often draw on monitoring data but will typically include other methods and data sources to answer more strategic questions.

M&E   An acronym used as shorthand to broadly denote monitoring and evaluation activities. It includes both the ongoing use of data for accountability and learning throughout the life of a grant, component, initiative, or strategy, as well as an examination of whether outcomes and impacts have been achieved. *[Gates Foundation glossary]*

OUTCOME-FOCUSED PHILANTHROPY (OFP)

A framework that guides how we do our philanthropic work, from start to finish. It reflects the foundation's commitments to being rigorous, flexible, adaptive, transparent, and open while staying focused on results and actively learning at every juncture. *[OFP Overview]*

STRATEGY

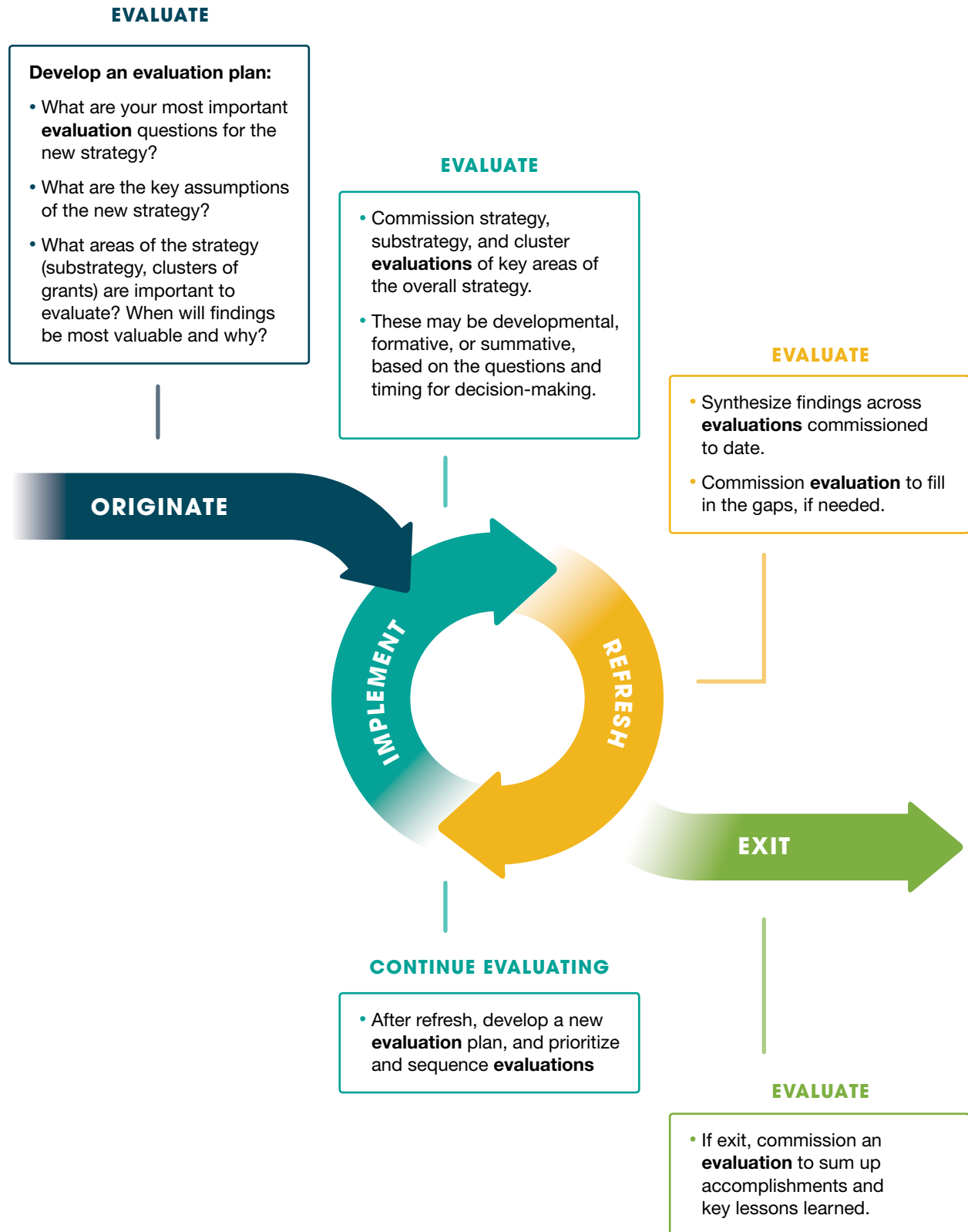A plan of action designed to achieve a particular goal.

SUBSTRATEGY

The different areas of work in which a program decides to invest its resources in order to achieve its goals. Each substrategy typically has its own theory of change, implementation markers, and outcomes—all of which are designed to advance the program's overall goals.

CLUSTER

A small group of grants with complementary activities and objectives that collectively advance a strategy toward its goal.

TARGETS

The desired level for goals the program plans to achieve with its funding. They are based on metrics and should be ambitious but achievable within the specified time frame.

THEORY OF CHANGE

A set of assumptions that describe the known and hypothesized social and natural science underlying the graphic depiction in a logic model. *[OFP Overview]*

APPENDIX A. Evaluate Throughout a Strategy

EVALUATE

Develop an evaluation plan:

• What are your most important **evaluation** questions for the new strategy?

• What are the key assumptions of the new strategy?

• What areas of the strategy (substrategy, clusters of grants) are important to evaluate? When will findings be most valuable and why?

EVALUATE

• Commission strategy, substrategy, and cluster **evaluations** of key areas of the overall strategy.

• These may be developmental, formative, or summative, based on the questions and timing for decision-making.

EVALUATE

• Synthesize findings across **evaluations** commissioned to date.

• Commission **evaluation** to fill in the gaps, if needed.

ORIGINATE

IMPLEMENT

REFRESH

EXIT

CONTINUE EVALUATING

• After refresh, develop a new **evaluation** plan, and prioritize and sequence **evaluations**

EVALUATE

• If exit, commission an **evaluation** to sum up accomplishments and key lessons learned.

APPENDIX B. **Editorial Guidance: What Evaluators Need to Know**

Consistent with the value the Hewlett Foundation places on openness and transparency, we are committed to sharing the results of evaluations of our grantmaking so that others may learn from our experience and hold us accountable for the results of our work. In fact, we presume that results from all evaluations will be shared publicly via our website, with only limited, principled exceptions where sharing could cause material harm to the foundation's grantees or our strategies.

In order to facilitate the foundation review and publication of the evaluation you are preparing for us, we ask you to keep the following points in mind as you draft your report and any related products. They reflect the most common requests we make for edits to draft evaluation reports, and we hope that making you aware of them in advance will help streamline the editing and publication process.

Provide accurate descriptions of grantee advocacy efforts. As you may know, as a private foundation, the Hewlett Foundation must comply with legal rules that preclude using our grant funds for lobbying and partisan election activities. That's the short description. The actual rules regarding grantee lobbying and election activities are quite complicated, and it is important that evaluations of our work reflect what is and is not permissible in our grant agreements. We ask that you flag for us in your draft report any sections where you discuss lobbying or election activities, and also note (either in the body of the report or a footnote) that while the report may describe grantees efforts to affect legislation or government policy, the Hewlett Foundation does not earmark its funds for prohibited lobbying activities, as defined in federal tax laws, and that the foundation does not fund partisan electoral activities, though it may fund nonpartisan election-related activities by grantees.

Provide accurate descriptions of funder/grantee relationship. The Hewlett Foundation believes strongly in treating our grantees as partners rather than contractors carrying out our directives. They are the ones with the expertise and front-line perspective, and we strive to work with them in ways "that are facilitative rather than controlling," as our guiding principles[49] put it. Because evaluations we commission often look through the lens of our grantmaking strategy, the nature of our relationship with grantees is sometimes lost, and grantees are portrayed in a manner that is more instrumental than collaborative. It's accurate to describe shared goals between the foundation and our grantees, but we ask that you avoid descriptions that paint us as "puppet masters" or strategists moving pieces on a chess board. To be sure, we may not always achieve our goals regarding collaboration and partnership, and if it's accurate and appropriate to report that, please do so. However, we do not describe our grantees' work or achievements as our own, and we prefer that evaluations not do so either. It is the work and achievements of grantees that we have strategically chosen to support.

Avoid undue flattery. There's a natural tendency in preparing a report for a client to sing the praises of that client. Sometimes, that results in puffery: evaluators giving undue praise or trying to flatter the foundation. This is wholly unnecessary and a bit off-putting. It also conflicts with our goal in commissioning an evaluation, which is to get constructive, independent feedback on work we support so that we and others can learn and improve. We ask that you strive for a dispassionate and measured tone, acknowledging with candor what we got right and what we got wrong.

Criticism of or confidential information about individual grantees. Two exceptions to our general rule about openness and transparency are information that we have an ethical or legal duty to keep confidential (e.g. staffing changes at a grantee that have not yet been made public) or situations where sharing information publicly could cause material harm to a grantee, such as criticism of an individual

organization's work. For that reason, we ask that you include whatever such information is relevant to your report, but flag it for us in a draft version so we can consider how best to fulfill our ethical and legal obligations to our grantee partners as we share the results in targeted fashion with other partners, or, more broadly, with the field and the public.

Thank you in advance for taking these points under advisement as you draft your evaluation reports and related products.

APPENDIX C. **Engage Grantees in Evaluation**

PLANNING

Get input from grantees about what they hope to learn from an evaluation.

Build grantee questions into the evaluation when possible.

Share draft RFP or Evaluation Purpose and solicit feedback.

Be clear about how the results of the evaluation will be used and shared publicly.

If appropriate, provide opportunity to weigh in on evaluator selection process.

Consider whether there will be an advisory group for the evaluation and how grantee representatives might be involved.

IMPLEMENTING

Introduce the external evaluator before the evaluation begins

Be upfront from the start about the demands of the evaluation (ie. share a FAQ).

Ask for input on methodology and timing of data collection activities.

Keep in touch with grantees about upcoming evaluation activities.

Check in about the evaluation process along the way: how is it going for them?

Share and discuss relevant findings.

USING + SHARING

Share findings with grantees and get feedback on findings and evaluator's interpretation.

Consider opportunities to co-create relevant recommendations.

Provide grantees with the opportunity to verify accuracy of data before finalizing.

Discuss how findings might be used.

Brainstorm settings and opportunities to share findings together.

Convene grantees to discuss the results and recommendations.

APPENDIX D. 7 Principles of Evaluation

1 We lead with purpose.

By anticipating our information needs, we are more likely to design and commission evaluations that will be useful and used.

- Design evaluation with actions and decisions in mind.
- Ask how and when will we and others use the information that comes from this evaluation.

2 Evaluation is fundamentally a learning process.

Establishing evaluation questions early in the strategy lifecycle helps us clarify and refine how, why, for whom, when, and to what extent objectives or goals are expected to be achieved.

- Actively learn and adapt, as we plan, implement, and use evaluations.
- Use evaluation as a key vehicle for learning as we implement our strategies, to bring new insights to our work.

3 We treat evaluations as a key part of the strategy lifecycle.

Building evaluative thinking in throughout the strategy lifecycle helps articulate key assumptions in a theory of change or strategy and establishes a starting point for evaluation questions and a proposal for answering them in a practical, meaningful sequence.

4 We cannot evaluate everything, so we choose strategically.

Undergoing an evaluation either of the whole strategy or of a key part within three years ensures we don't go too long without external eyes.

- Criteria guide decisions about where to put our evaluation dollars, including: opportunity for learning, urgency to make course corrections or future funding decisions, the potential for strategic or reputational risk, and size of investment as a proxy for importance.

5 We maximize rigor without compromising relevance.

Selecting evaluation designs that use multiple methods and data sources when possible strengthens our evaluation designs and reduces bias.

- Match methods to questions and do not routinely choose one approach or privilege one method over others.
- Evaluations clearly articulate methods used and their limitations.
- Evaluations include comparative reference points.

6 We share what we are learning with appropriate audiences and share publicly.

We presumptively share the results of our evaluations so that others may learn from our successes and failures.

- Identify audiences for findings as we plan.
- Communicate early with our grantees and co-funders about intention to evaluate and plans to share.
- Share the results of our evaluations, in some form (full, executive summary, or presentation) with care given to issues of confidentiality.

7 We use the data!

Not using findings is a missed opportunity for learning and course correction.

- Take time to reflect on the results, generate implications for our strategies, grantees, policy or practice, and adapt.
- Combine the insights from evaluation results with wisdom from our own experiences.

[1] Evaluation Principles and Practice, First Edition, https://www.hewlett.org/library/evaluation-principles-and-practices/

[2] The Value of Evaluations: Assessing spending and quality, https://hewlett.org/value-evaluations-assessing-spending-quality/

[3] Hewlett Foundation Guiding Principles reference to Outcome-Focused Approach, https://hewlett.org/outcome-focused-approach/

[4] Outcome-Focused Philanthropy, http://www.hewlett.org/wp-content/uploads/2016/12/OFP-Guidebook.pdf

[5] Tracking Progress: Setting, Collecting, and Reflecting on Implementation Markers, July 2018, https://hewlett.org/wp-content/uploads/2018/07/OFP-Guide-Tracking-Progress.pdf

[6] "Time for a Three-Legged Measurement Stool: Going beyond traditional monitoring and evaluation to focus on feedback can lead to new innovations in the social sector," Fay Twersky, SSIR, Winter 2019, https://ssir.org/articles/entry/time_for_a_three_legged_measurement_stool

[7] Outcome-Focused Philanthropy, http://www.hewlett.org/wp-content/uploads/2016/12/OFP-Guidebook.pdf

[8] Hewlett Foundation Guiding Principles reference to Diversity, Equity and Inclusion Principle, hewlett.org/diversity-equity-inclusion/

[9] Hewlett Foundation Guiding Principles reference to Diversity, Equity and Inclusion Principle, hewlett.org/diversity-equity-inclusion/

[10] The Value of Evaluations: Assessing spending and quality, https://hewlett.org/value-evaluations-assessing-spending-quality/

[11] Hewlett Foundation Guiding Principles reference to Openness, Transparency and Learning, https://hewlett.org/openness-transparency-learning/

[12] Hewlett Foundation Guiding Principles, https://www.hewlett.org/about-us/values-and-policies

[13] "5-A-Day: Learning by Force of Habit," https://medium.com/@jcoffman/5-a-day-learning-by-force-of-habit-6c890260acbf

[14] The Value of Evaluations: Assessing spending and quality, https://hewlett.org/value-evaluations-assessing-spending-quality/

[15] American Evaluation Association Guiding Principles For Evaluators, http://www.eval.org/p/cm/ld/fid=51

[16] Evaluation of The William and Flora Hewlett Foundation's Organizational Effectiveness Program Final Report, November 21, 2015, https://hewlett.org/wp-content/uploads/2016/10/Evaluation-of-OE-Program-November-2015.pdf

[17] "Assessing nonprofit capacity: A guide to tools," Prithi Trivedi and Jennifer Wei, October 30, 2017, https://hewlett.org/assessing-nonprofit-capacity-guide-tools/

[18] "Evaluating the Madison Initiative," Daniel Stid, January 24, 2019, https://hewlett.org/library/evaluating-the-madison-initiative/

[19] Evaluation of Network Building, Camber Collective, November, 2016, https://hewlett.org/wp-content/uploads/2018/02/Evaluation-of-network-building-Cyber-2016.pdf

[20] Evaluation Final Report: Hewlett Foundation's strategy to apply human-centered design to family planning and reproductive health, Itad, July 24, 2018, https://hewlett.org/library/evaluation-of-the-hewlett-foundations-strategy-to-apply-human-centered-design-to-improve-family-planning-and-reproductive-health-services-in-sub-saharan-africa/

[21] "Promise and progress on family planning in Francophone West Africa", Margot Fahnestock, July 25, 2018, https://hewlett.org/promise-and-progress-on-family-planning-in-francophone-west-africa/

[22] International Women's Reproductive Health: Supporting Local Advocacy in Sub-Saharan Africa, April 2016, https://hewlett.org/wp-content/uploads/2016/11/Supporting-Local-Advocacy-in-Sub-Saharan-Africa.pdf

[23] Western Conservation Strategy 2018-2023, July 16, 2018, https://hewlett.org/library/western-conservation-strategy-2018-2023/

[24] Best Practices for Enduring Conservation, Hovland Consulting, July 2018, https://hewlett.org/library/best-practices-for-enduring-conservation-five-year-retrospective-of-the-hewlett-foundations-western-conservation-grantmaking-strategy/

[25] Research on Open: OER Research Hub Review and Futures for Research on OER, Linda Shear, Barbara Means, Patrik Lundh, October 14, 2015, https://hewlett.org/library/research-on-open-oer-research-hub-review-and-futures-for-research-on-oer/

[26] Evaluation findings, https://www.fundforsharedinsight.org/knowledge/?t=evaluating-our-work#knowledge-tabs|3

[27] The Hewlett Foundation Education Program Deeper Learning Review: Executive Summary, January 30, 2014, https://hewlett.org/library/the-hewlett-foundation-education-program-deeper-learning-review-executive-summary/

[28] Deeper learning, six years later, Barbara Chow, April 26, 2017, https://hewlett.org/deeper-learning-six-years-later/

[29] The William and Flora Hewlett Foundation's Nuclear Security Initiative—Findings from a Summative Evaluation, ORS Impact, May 4, 2015, https://hewlett.org/library/the-william-and-flora-hewlett-foundations-nuclear-security-initiative-findings-from-a-summative-evaluation/

[30] "The Legacy of a Philanthropic Exit: Lessons From the Evaluation of the Hewlett Foundation's Nuclear Security Initiative," Anne Gienapp, Jane Reisman, David Shorr, and Amy Arbreton, The Foundation Review: Vol. 9: Iss. 1, Article 4, 2017, https://scholarworks.gvsu.edu/tfr/vol9/iss1/4

[31] "Q&A with Margot Fahnestock: A teen-centered approach to contraception in Zambia and Kenya," Sarah Jane Staats, July 25, 2018, https://hewlett.org/qa-with-margot-fahnestock-a-teen-centered-approach-to-contraception-in-zambia-and-kenya/

[32] "BetterEvaluation community's views on the difference between evaluation and research," December 2014, Patricia Rogers, https://www.betterevaluation.org/en/blog/difference_between_evaluation_and_research

[33] Improving the Practice of Philanthropy: An Evaluation of the Hewlett Foundation's Knowledge Creation and Dissemination Strategy, Harder + Co, November 2013, https://hewlett.org/library/an-evaluation-of-the-knowledge-creation-and-dissemination-strategy/

[34] Evaluation of the Population and Poverty Research Initiative (PopPov),Julie DaVanzo, Sebastian Linnemayr, Peter Glick, Eric Apaydin, 2014, https://hewlett.org/wp-content/uploads/2016/08/RR527_REVFINAL-COMPILED.pdf

[35] Best Practices for Enduring Conservation, Hovland Consulting, July 2018, https://hewlett.org/library/best-practices-for-enduring-conservation-five-year-retrospective-of-the-hewlett-foundations-western-conservation-grantmaking-strategy/

[36] A new conservation grantmaking strategy for today's challenges, Andrea Keller Helsel, July 16, 2018, https://hewlett.org/a-new-conservation-grantmaking-strategy-for-todays-challenges/

[37] Contribution Analysis, https://www.betterevaluation.org/en/plan/approach/contribution_analysis

[38] Process Tracing, https://www.betterevaluation.org/evaluation-options/processtracing

[39] Qualitative Comparative Analysis, https://www.betterevaluation.org/evaluation-options/qualitative_comparative_analysis

[40] Quip, https://www.betterevaluation.org/en/plan/approach/QUIP

[41] Taking stock of our Performing Arts grantmaking, John McGuirk, April 2016, https://hewlett.org/taking-stock-of-our-performing-arts-grantmaking/

[42] Evaluation of Network Building, Camber Collective, November, 2016, https://hewlett.org/wp-content/uploads/2018/02/Evaluation-of-network-building-Cyber-2016.pdf

[43] Evaluation findings, https://www.fundforsharedinsight.org/knowledge/?t=evaluating-our-work#knowledge-tabs|3

[44] Adapted from *Adaptive action: Leveraging uncertainty in our organization*. G. Eoyang, R. Holladay, 2013, Stanford University Press

[45] 52 weeks of BetterEvaluation: Week 23: Tips for delivering negative results, Jessica Sinclair Taylor, June 2013, http://www.betterevaluation.org/blog/delivering-bad-news

[46] Peer to Peer: At the Heart of Influencing More Effective Philanthropy: A Field Scan of How Foundations Access and Use Knowledge, Harder+Co and Edge Research, February 2017, http://www.hewlett.org/wp-content/uploads/2017/03/Hewlett-Field-Scan-Report-2017-CCBYNC.pdf

[47] Peer to peer: At the heart of influencing more effective philanthropy, February 2017, https://hewlett.org/peer-to-peer-at-the-heart-of-influencing-more-effective-philanthropy/

[48] Finally, a foundation-commissioned study that actually helps its grantees, by Aaron Dorfman, March 2017, https://www.ncrp.org/2017/03/finally-foundation-commissioned-study-actually-helps-grantees.html

[49] Hewlett Foundation Guiding Principles, https://www.hewlett.org/about-us/values-and-policies