Table of Contents

Introduction
Why Care About Leadership Development Evaluation?
How is This Guide Organized?

Part 1: Framing Leadership Development as a Social Change Strategy
Leadership Development Theory of Change
Status of Leadership Development Evaluation
A Framework for Building a Customized Theory of Change
A Construct for Evaluating Leadership Development

Part 2: Implementing Evaluation for Leadership Development
Individual and Organizational Capacity for Leadership
Collective Capacity for Leadership
Social Change Levers

Part 3: Putting It All Together
Three Real-life Case Examples
Hypothetical Case Example
### Table of Tables

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>1a Individual Leadership Capacity: Example Measures and Potential Methods</td>
</tr>
<tr>
<td>15</td>
<td>1b Organizational Leadership Capacity: Example Measures and Potential Methods</td>
</tr>
<tr>
<td>17</td>
<td>2 Collective Leadership Capacity: Example Measures and Potential Methods</td>
</tr>
<tr>
<td>19</td>
<td>3 Social Change Levers: Example Measures and Potential Methods</td>
</tr>
<tr>
<td>26</td>
<td>4 Methods and Data Sources to Evaluate Leadership Outcomes at the Individual Level</td>
</tr>
<tr>
<td>30</td>
<td>5 Methods and Data Sources to Evaluate Leadership Outcomes at the Organizational Level</td>
</tr>
<tr>
<td>33</td>
<td>6 Methods and Data Sources to Evaluate Collective Leadership</td>
</tr>
<tr>
<td>36</td>
<td>7 Methods and Data Sources to Evaluate Contribution of Leadership Development to Social Change</td>
</tr>
</tbody>
</table>
Philanthropic funders increasingly recognize that committed and effective leadership can be a powerful vehicle for driving social change.

This has led to an increase in leadership development programs that aim to strengthen the leadership capacity of individuals, organizations, networks, and the communities and/or sectors in which they work.¹

Leadership development is the process by which the skills and capacities of leaders are intentionally developed in furtherance of a goal. Developing leaders for social change transcends a traditional focus on managerial leadership within an organizational hierarchy. Increasingly, funders realize that fostering a mindset that is directed toward solving the complex, systemic issues facing society today requires an expanded set of leadership competencies.

Why Care About Leadership Development Evaluation?

Evaluation of leadership development in the social sector presents challenges. Evaluators should be encouraged to look beyond the immediate outcomes associated with individual leaders (outcomes largely within direct programmatic control) to consider the contribution of leaders to the issues, sectors, and communities they are trying to affect. In essence, evaluation must be guided by the larger question of leadership for what? What changes do we want to see as the result of our leadership development both immediately and in the long term?

Using a broader set of evaluation questions and approaches will ultimately support the field’s ability to learn from its leadership development efforts. By compiling systematic evidence of contribution (in addition to changes directly attributed to the program), evaluation can help the field learn which programmatic elements lead to short-term and interim outcomes, and which outcomes best predict longer-term goals, thereby increasing the effectiveness of leadership for social change.

Making these connections is not an easy task. Those interested in measuring leadership development have choices to make about what to measure and how. Appropriate choices depend on the nature of the program, the broader purpose that leadership is intended to address, the relationship between philanthropy and the leaders they are supporting, the historical and cultural context of leadership, and beliefs about how change occurs. This guide is intended to clarify these choices to support good evaluation practice.

¹ In 2011, estimated foundation spending on leadership totaled $400 million, excluding grants to individuals (McKinsey, 2014).
The content of this guide is informed by our evaluation practice, but relies heavily on an extensive review of the literature on leadership development, social change, and evaluation. As such, any wisdom imparted here owes a great debt to others in the field who have sought to systematically capture their experiences, share lessons learned, and develop frameworks to support this work.

We explicitly acknowledge that the role of philanthropy in developing leaders has been neither static nor without controversy. The choice of who participates in leadership development, what strategies will be employed, and what outcomes are sought, have often been closely aligned with the priorities and values of funders. To undo inequitable power structures, philanthropy should make an intentional effort to let go and enable the leaders themselves to make these choices and to incorporate community or stakeholder engagement strategies. Some funders are doing just that. Our belief is that theory of change and evaluation can play a valuable role in identifying the explicit and implicit assumptions that shape leadership development, uncovering diverse views, and assessing outcomes that matter the most to those affected by the effort, directly and indirectly.

We offer this guide in the spirit of promoting shared learning for the field of leadership development evaluation. The primary audience is evaluators with responsibility for evaluating investments in leadership development for social change. Funders interested in telling the story of their investments might be interested in using a typology (See ORS Impact [2018]. A typology of investments in leadership development. Available at: http://orsimpact.com/directory/a-typology-of-investments-in-leadership-development.htm). The focus of this guide is on evaluation of outcomes and impacts rather than program design or implementation. Accordingly, we do not provide guidance on program design or the selection of leaders, important issues that are addressed elsewhere. See, for example, Rockwood Leadership Institute, CompassPoint, Leadership Learning Community, and Bridgespan Group’s Nonprofit Leadership Development Toolkit.

How is This Guide Organized?

*This guide contains three parts:*

1. Part 1 provides guidance to help evaluators conceptualize leadership development, emphasizing both the nature of the leadership development program and beliefs about how change occurs.

2. Part 2 provides details to help evaluators select measures and methods to support good evaluation practice.

3. Part 3 provides concrete case examples that demonstrate the application of the guidance provided in Parts 1 and 2.
Framing Leadership Development as a Social Change Strategy

Leadership Development Theory of Change
Status of Leadership Development Evaluation
A Framework for Building a Customized Theory of Change
A Construct for Evaluating Leadership Development
There is a wide range of programs and strategies for building the capacity of leaders, reflecting assumptions about who leaders are and how change is expected to happen. Regardless of format, leadership development is rarely seen as an end in and of itself. Rather, it represents an investment in individuals or groups of individuals toward a larger goal or set of goals.

Individuals are selected to participate in leadership development for many reasons. Often it is because they are perceived to be well positioned to advance goals because of their ties to a place (e.g. community, region), a specific organization, a sector (e.g. agriculture, health, education), or an issue (e.g. water conservation, climate change, breast feeding). Individuals may also be selected to intentionally support the development and success of diverse staff and leaders, including people of color and women.

Leadership Development Theory of Change

Traditionally, leadership development has focused on building the capacity of selected individuals so that they can be more effective leaders for social change. The belief is that by increasing their skills and capacities, they will be better able to address myriad social goals. A growing interest in the concept of collective leadership has led more recently to approaches that target cohorts, coalitions, or partnerships with an explicit goal of creating and mobilizing networks for social change. This reflects a belief that addressing complex social problems requires collective action.

Whether investing in individuals or cohorts, leadership development practitioners can choose from many strategies. Some may target individuals who are already leaders in their organizations, communities or professional fields. Others may focus on individuals earlier in their journey who can serve as the next generation of leaders. Some programs, especially those that use a cohort model, intentionally mix seasoned and emerging leaders.
As suggested above, the choice of leadership development strategy reflects an underlying theory or set of assumptions about how change happens. To carry this further, let’s consider common goals associated with leadership development. For leadership to make a difference and achieve a desired population-level change, funders may explicitly or implicitly believe that their leadership development investments will achieve social change through one or more of the following social change levers:

### Advancing an issue or policy

Often, funders invest in leadership to advance a specific policy or policy agenda. An example might be building the specific capacities, behaviors or connections that best support leaders in advancing state-level forest conservation protection solutions.

### Increasing community self-determination

Leadership development is sometimes place-based, with a goal of building the capacity of communities to identify and solve problems. Programs focused on increasing community capacity for self-determination often emphasize inclusion, seeking to listen to and amplify voices that are typically excluded from the negotiating table. An example would be building a network of leaders in a community who work to dismantle systems that perpetuate inequalities and advance new collective visions for their communities.

### Strengthening or reforming a sector

Leadership development is often focused on a specific sector to strengthen capacity generally or to reform or improve specific practices. An example might be providing fellowship support to individuals working in the seafood sector to develop a pipeline of leaders who can promote and sustain sustainable seafood practices.

### Building a large-scale social movement

Leadership development may have a larger goal of fostering or supporting a broad-based social movement to achieve large-scale change. Social movements are a form of long-term, sustained collective action. An example is the National Domestic Workers Alliance Strategy-Organizing-Leadership (SOL) Initiative to develop grassroots leaders representing domestic workers to play a role in the broader social justice movement (Ito et al, 2014).

These social change levers represent common beliefs about social change but are by no means exhaustive. The specific larger-scale population changes that these levers are meant to influence may vary widely depending on the goals of the funder or the leaders themselves, and might include changes in populations (e.g., levels of employment, or decreases in disease incidence) or changes in social or environmental conditions (e.g., increased economic opportunity, improved air quality).
As recently as 2002, few leadership development programs had an explicit theory of change, and most failed to articulate links between individual changes and changes in organizations, communities, systems, or fields (Reinelt, Foster, and Sullivan, 2003). Too often, it is assumed that increasing the skills and capacities of individual leaders will result in desired social change with no attempt to explicitly define those assumptions or pathways. We believe that explicitly articulating assumptions about how leadership contributes to change is essential to informing good evaluation that can help the field learn how best to develop leaders that can contribute to social change.

By articulating the nature of the program, the end goals, and how change is believed to occur; a theory of change can help bring clarity about the appropriate unit of change, timeframe, and where measurement can be directed. Being more systematic about conceptualizing leadership development can help evaluators focus on measurement that supports learning both within and across programs.

Status of Leadership Development Evaluation

Evaluation practice suffers from similar limitations. It is not uncommon for evaluations of leadership development, when they have been conducted at all, to focus on outcomes associated with individual leadership capacity, including knowledge, skills, values, beliefs, identities, attitudes, and behaviors. Most evaluations are short term, lacking both the tools and time horizon to adequately assess long-term outcomes associated with large-scale social change goals.

While short-term individual approaches to evaluating outcomes of leadership development continue to dominate the field, there are signs that this is shifting. For example, the Leadership Learning Community, a national nonprofit, has hosted a series of meetings with leadership funders to talk about both desired changes and those changes that are currently getting measured. They found a noticeable growth between 2008 and 2014 in funders’ interest in facilitating and measuring larger scale results. That is, while most programs still measure individual change, and some measure organizational outcomes, there is increasing focus on measuring changes at the field, network, community, and population level (Meehan, Reinelt, and Leiderman, 2015).
There is growing emphasis on evaluation approaches that go beyond changes in individual leaders to understand the contribution of leadership to more distal outcomes. The growing interest in the concept of collective leadership is also informing evaluation practice. Evaluators are beginning to pay attention to how leaders work with others, measuring changes in connections and relationships. There is also greater attention to being explicit about the links between individual and organizational competencies and the community and systems changes that are the ultimate aim. Technical challenges remain in identifying how individuals, cohorts, and networks contribute to social change, and the relative contribution of leaders in relation to other influences.

Seminal work in this area includes several efforts to categorize the types of outcomes associated with leadership development. A 2002 scan of leadership programs organized outcome measures into five domains that programs may hope to achieve: individual, organizational, community, field, and system (Reinelt, Foster, and Sullivan, 2003). In 2008, the Leadership Learning Community developed a leadership development investment framework that summarizes where results are likely to occur, specifically looking at capacities at the following levels: individual, team, organization, network, and systems. This framework was further developed in a publication by McGonagill and Reinelt (McGonagill and Reinelt, 2010).

Other work has focused specifically on social movements. The Building Movement Project includes a strategy for supporting young leaders who bring new ideas and energy to social change movements (Building Movement Project, 2010). Pastor et al (2011) provides a summary of metrics for movement building that centers around organizing, civic engagement, leadership development, alliance building, and campaigns, with a call to action for philanthropy to focus on what matters most in the long term in their leadership investments.

Other funders are challenging the traditional relationship between philanthropy and community by considering a broader set of roles for funders in developing leaders, including acting as convener, organizer, relationship broker, and knowledge disseminator. This opens the door for co-creation of both programs and their evaluations (Grant, 2014).

These boundary-pushing developments are part of an effort to move both funders and evaluators toward greater social relevance. By increasingly striving to link leadership development programs with outcomes and confronting head on the challenges of demonstrating attribution and contribution, we believe that as evaluators we can increase the value of our practice for social change and more effectively support the funders that are investing in leaders for this purpose.
A Framework for Building a Customized Theory of Change

The evaluation community has begun to define leadership competencies and revise frameworks and approaches that suggest where evaluators should look for changes that result from leadership development. There has been a corresponding expansion in the suite of tools and approaches used by evaluators. While the field has not yet coalesced around an agreed-upon set of measures or tools, there are helpful frameworks, indicator lists, and resources that can help evaluators make the best choices.

In this section, we lay out the building blocks for creating a customized theory of change and identifying relevant measures to support an evaluation. A theory of change is a conceptual model that articulates the linkages among strategies, outcomes, and goals that support a mission or vision for achieving social impact. It is typically presented as a visual diagram accompanied by a narrative detailing the underlying assumptions. We hope you will see your own investments or programs reflected in these pages and will be inspired to think carefully about the types of change you hope to see and how your evaluation strategy can explicitly capture and measure those changes. We include key evaluation questions and sample measures and methods for each level of change:

1. **Individual Leadership and Organizational Capacity**
2. **Networks and Collective Action**
3. **Social Change Levers**

Our goal in presenting these materials is to provide a framework and tools that will lead to good evaluation decisions. We have purposefully provided more options than could or should be utilized in any given evaluation. By providing more options rather than fewer, we aim to support better choices that truly reflect the nature of the program and beliefs about what results can be expected. While we strongly believe that evaluation should consider multiple levels of change, we also caution against over building and instead encourage you to “right-size” and choose those options that best fit the goals and expectations for change at any given point in time. We hope the example we present will help you visualize how to right-size your evaluation strategy.

3 ORS Theory of Change guide available at www.orsimpact.com
While individual leadership growth remains at the heart of leadership development, two trends merit attention.

First, the concept of who is a leader has expanded. No longer are investments made only in individuals who hold senior leadership positions. There has been increasing focus on non-traditional leaders including grassroots leaders, youth, other individuals in informal leadership positions, and individuals traditionally excluded from leadership opportunities. Second, the sophistication of thinking around measuring the attributes of an effective leader has continued to evolve along with ideas about 21st century leadership, collective leadership, and beliefs about how social change occurs. Leadership competencies increasingly include attributes related to values, styles of engagement, vision, adaptability, and ethics, such as systems thinking or the ability to apply a racial equity lens. At the organizational level, competencies associated with leading change include creativity and innovation, external awareness, resiliency, team building, developing others, and leveraging diversity (U.S. Fish and Wildlife, 2010).

In their Leadership for a New Era series launched in 2009, the Leadership Learning Community explores leadership approaches that are “more inclusive, networked and collective” with an explicit focus on leadership and race (2010), leadership and collective impact (2012), leadership and networks (2012), and leadership and large-scale change (2015). In each initiative, they shine a light on important leadership competencies including systems thinking, network weaving, risk taking, organizing, learning and reflection, bridging, meaning making, collaboration, civic engagement, innovation thinking, design thinking, vision and framing, diverse stakeholder engagement, commitment, alignment and accountability, and leveraging leadership to create a more just and equitable society.

Some leadership programs seek not only to improve the capacity of individual leaders, but also to improve the functioning and effectiveness of the teams and organizations in which leaders work. There are several ways that changes in individual leadership capacity might translate into gains at the organizational level.

**Examples include:**

- **Knowledge and skills gained in trainings might influence an organization’s adoption of new policies that improve their ability to carry out key functions such as community engagement or communications.**

- **Partnerships and relationships formed during leadership development might translate into cross-organizational agreements and new joint endeavors or collaborations.**

- **Investment in organizational leaders may strengthen the long-term health and viability of the organization, including building a robust pipeline of future leaders. These kinds of changes may be particularly relevant when the purpose of the investment is to build or strengthen an entire sector.**
Thus, evaluations may also focus on the capacities of the teams or organizations in which leaders work. Evaluation of leadership outcomes at the organizational level, while dependent on the type of capacity change sought, might assess outcomes such as the organization’s staffing mix, resiliency, and ability to develop future leaders. It might also focus on the organization’s ability to respond to community needs or create environments where best practices thrive. As with individual measures, there is increasing attention on an organization’s ability to collaborate, develop partnerships, operate in a diverse and global environment, and adapt to change—attributes seen as important for success in today’s environment.

Evaluation of leadership growth can look at changes in individual or organizational capacity in response to a specific training, as well as changes over time resulting from sustained participation in fellowships, communities of practice, or other long-term programs.

**Evaluation questions**

Relevant evaluation questions for changes in individual leadership capacity can range in focus from individual leadership skills and knowledge enhancement to broader transformative changes in leadership styles, values, and influence. Questions related to changes in organizations can range in focus from changes in internal or external leadership practices, to resiliency of the organization to sustain shocks, to perceptions of the organization as a leader in a sector or on an issue.

*Examples include:*

- How have leaders’ awareness, knowledge, or competencies changed?
- Which training elements are most associated with those changes?
- To what extent have organizations improved their capacity to thrive and provide leadership in their sector?
- Have leaders (or organizations) undergone sustained shifts in values, behaviors, cultures, or ability to influence outcomes?
- Which individual (or organizational) changes are making a difference for the goals of the leadership program?
- How well do internal systems, policies, and practices support the development and success of diverse staff and leaders, including people of color and women?
- To what extent is the individual or organization recognized as a leader in the sector or issue area?
Example measures and methods
Table 1 provides example measures and data collection methods for evaluating leadership capacity of individuals and organizations. The evaluation questions and measures and methods presented above provide a sample of approaches for measuring leadership capacity. More guidance on how to make the right choices for a given evaluation is provided in Part 2.

### Table 1(a) Individual Leadership Capacity: Example Measures and Potential Methods

<table>
<thead>
<tr>
<th>Example Measures</th>
<th>Potential Methods</th>
</tr>
</thead>
</table>
| **Increased awareness of leadership style** |  » Assessments of leadership style  
» Self-reports of individual leaders |
| » Consciousness of self  
» Awareness of strengths and weaknesses |
| **Increased understanding of leadership skills and mindsets** |  » Multi-rater capacity assessments (360° assessments)  
» Self-reports of individual leaders, peers, or program staff  
» Participatory methods  
(e.g. appreciative inquiry, most significant change)  
» Case studies |
| » Advocacy skills  
» Systems thinking  
» Adaptive leadership  
» Collective impact  
» Problem solving  
» Racial equity  
» Participatory methods |
| **Changed behavior or sustained growth** |  » Multi-rater capacity assessments (360° assessments)  
» Self-reports of individual leaders, peers, or program staff  
» Participatory methods  
(e.g. appreciative inquiry, most significant change)  
» Case studies |
| » Self-efficacy  
» Cross-cultural competence  
» Social capital  
» Ability to collaborate with others  
» Management ability  
» Participatory methods |
| **Career advancement** |  » Career position tracking  
» Self-reports of individual leaders  
» Case studies |
| » Increased responsibility and influence  
» New positions |
### Table 1(b) Organizational Leadership Capacity: Example Measures and Potential Methods

<table>
<thead>
<tr>
<th>Example Measures</th>
<th>Potential Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Increased organizational capacity</strong></td>
<td>» Self-reports of organizational leaders and staff</td>
</tr>
<tr>
<td>» Ability to motivate and mobilize constituents</td>
<td>» Organizational capacity assessments</td>
</tr>
<tr>
<td>» Financial stability</td>
<td>» Review of organizational internal documents, budgets, reports or media coverage</td>
</tr>
<tr>
<td>» Fundraising ability</td>
<td>» Participatory methods (e.g. appreciative inquiry, most significant change)</td>
</tr>
<tr>
<td>» Marketing and communication expertise</td>
<td></td>
</tr>
<tr>
<td>» Ability to utilize data for internal decision making</td>
<td></td>
</tr>
<tr>
<td>» Ability to promote equity and inclusion</td>
<td></td>
</tr>
</tbody>
</table>

| **Changes in internal culture**                        |                                                                                 |
| » Learning and innovation                              |                                                                                 |
| » Equitable policies and practices                     |                                                                                 |

| **Changes in external perception**                     |                                                                                 |
| » Organization recognized as leader                    |                                                                                 |

| **Changes in connections**                             |                                                                                 |
| » Number of partners                                   |                                                                                 |
| » Quality of partnerships                              |                                                                                 |

| **Improved organizational infrastructure and technology**|                                                                                 |
| » Systems and structures                               |                                                                                 |
| » Technology upgrades                                  |                                                                                 |
Increasingly, funders appreciate that powerful social change often occurs through collective action.

To support and encourage collaborative leadership, leadership development may focus explicitly on a cohort or coalition: creating network mindsets, building connections between individuals, creating robust local or national networks, and mobilizing networks for collective action. This approach de-emphasizes the idea of the individual leader in favor of a collective model of leadership that is more inclusive of a wide range of actors and is better suited to tackling complex, systemic problems. These network leadership strategies are thought to build social capital, catalyze community engagement, mobilize people to a cause, bring projects to scale, and foster greater equity (Meehan and Reinelt, October 2012).

Programs that seek to build networks and help leaders adopt a network mindset may do so through developing specific skills, supporting processes that build relationships, and/or providing space or resources to apply skills and relationships in the planning or execution of joint projects. The specific aim of the program may be to build a network to accomplish a short-term goal, or the focus may be on building longer-term connections to sustain a movement.

Depending on the program and goal, evaluation of a coalition or network might focus on its membership, structure, or operations, as well as on the ability of collaborators to align goals, coordinate resources, or organize around shared goals. Increasingly, leadership development evaluations incorporate social network analysis to evaluate the extent to which relationships between leaders have formed, and the strength of those relationships. Network analysis supplements more traditional data collection methods such as surveys, interviews, and collaboration rubrics.

Evaluation questions

Relevant evaluation questions for changes in collective leadership capacity range from changes in the structure and operations of formal or informal networks to questions about network leverage and influence.

Examples include:

» Have the number of connections among individuals or organizations expanded to include those most relevant to their work?

» To what extent are people in the network building trust and sharing information?

» To what extent have members aligned their goals and resources?

» Has the network become a sustainable force that can catalyze systemic changes?

» Does the network include members who are able to identify and challenge sources of inequity?
Example measures and methods
Table 2 shows example measures and data collection methods for evaluating collective leadership capacity.

Table 2  Collective Leadership Capacity: Example Measures and Potential Methods

<table>
<thead>
<tr>
<th>Example Measures</th>
<th>Potential Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Increased growth in members and connections</em></td>
<td>» Self-report of participants in collective efforts</td>
</tr>
<tr>
<td>» Increased number of members</td>
<td>» Measurement of network composition</td>
</tr>
<tr>
<td>» Increased representation or engagement of priority populations</td>
<td>» Network analysis</td>
</tr>
<tr>
<td>» Increased connectivity and density</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Increased quality of network interactions</em></td>
<td>» Self-report of participants in collective efforts</td>
</tr>
<tr>
<td>» Increased trust</td>
<td>» Levels of collaboration rubrics</td>
</tr>
<tr>
<td>» Increased communication and information sharing</td>
<td></td>
</tr>
<tr>
<td>» Increased collaboration</td>
<td></td>
</tr>
<tr>
<td>» Distribution of power within the network</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Increased collective action</em></td>
<td>» Self-report of participants in collective efforts</td>
</tr>
<tr>
<td>» Increased alignment of goals and resources</td>
<td>» Key informant interviews</td>
</tr>
<tr>
<td>» Increased innovation</td>
<td>» Case studies</td>
</tr>
<tr>
<td>» Increased visibility and diffusion of new ideas</td>
<td></td>
</tr>
<tr>
<td>» Increased amplification of best practices</td>
<td></td>
</tr>
</tbody>
</table>
Leadership for what purpose?

While individual, organizational, or collective changes targeted through leadership development programs are important, we can’t lose sight of the larger social change goals. In other words, leadership for what? What changes do we want to see in systems, fields, or the community at large? Evaluation of the contribution of leadership to these social goals requires a different set of questions, measures, and methods. Using social change levers provides a framework for defining relevant evaluation questions and considerations suited to explicitly evaluating the contribution of leadership to social change. Answering many of these questions pushes evaluation beyond immediate outcomes that are within direct programmatic control.

Evaluation questions
The list of relevant questions can be endless. Here are a few to help you get started.

Examples include:

<table>
<thead>
<tr>
<th>Advancing an issue or policy</th>
<th>Strengthening or reforming a sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>How has political will shifted among decision makers? Among the public?</td>
<td>How are policies and practices changing within the sector?</td>
</tr>
<tr>
<td>What new policies have resulted?</td>
<td>How is new leadership transforming the sector?</td>
</tr>
<tr>
<td>To what extent are policies being implemented?</td>
<td>What changes are occurring because of changes in the sector?</td>
</tr>
<tr>
<td>What impact have policies had, intended or unintended?</td>
<td></td>
</tr>
</tbody>
</table>

Increasing community self-determination

To what extent have structural barriers to community engagement or decision making been dismantled?

To what extent has participation in community decision making increased?

How are new voices, including underrepresented voices, shaping community decisions?

What other changes have resulted from the efforts of community members?

Building a large-scale social movement

To what extent have individuals mobilized around the issue?

How are social norms shifting?

To what extent are movement leaders recognized by public and political institutions?

To what extent has the movement built political will?

Example measures and methods
The evaluation questions, measures, and methods presented in Table 3 provide a sample of approaches for measuring social change levers and answering questions not only of attribution, but of contribution. The selection of measures should be guided by your theory of change. More guidance on how to make the right choices for a given evaluation is provided in Part 2.
<table>
<thead>
<tr>
<th>Levers</th>
<th>Example Measures</th>
<th>Potential Methods</th>
</tr>
</thead>
</table>
| Strengthening or reforming a sector | » Shared vision for sector reform  
» Improved alignment of programs and services  
» Improved sector collaboration  
» Increased information sharing  
» Clear role definition  
» Increased health of the sector  
» Reduced duplication of effort  
» Strengthened capacity across the sector  
» Increased ability to attract talent and leadership  
» Increased resilience and sustainability | » Population-level surveys  
» Bellwether interviews  
» Key informant interviews  
» Media analysis  
» Case studies  
» Contribution analysis  
» Developmental evaluation  
» Monitoring of secondary data (e.g. industry reports, sector reports) |
| Advancing an issue or policy | » Increased attention to issue and/or potential policy solutions  
» Among policymakers, the public, or within specific groups  
» Increased ability to influence the policy process  
» Issue framing, agenda setting, access to decision makers  
» Increased political and public will  
» Political leaders demonstrate greater willingness to act  
» Legislators sponsor legislation around policy priorities  
» Good policies adopted and/or implemented  
» Changes in social and/or physical conditions result from policies enacted | » Bellwether interviews  
» Key informant interviews  
» Media analysis  
» Case studies  
» Policymaker ratings  
» Advocacy capacity assessments  
» Monitoring of secondary data (e.g. budgets, resources, policies, regulations) |
| Increasing community self-determination | » Increased community engagement  
» Increased civic participation (e.g. voting and volunteerism)  
» Increased community voice in decision making  
» Increased community representation on local decision-making bodies (e.g. councils, legislatures, and school boards)  
» Increased diversity on boards of local nonprofit organizations  
» Increased resources available in the community  
» Increased private investment  
» Increased public programming or public funding  
» New infrastructure is put in place (e.g. roads, etc.)  
» Increased willingness of community members to act | » Population-level surveys  
» Key informant interviews  
» Policymaker ratings  
» Media analysis  
» Case studies  
» Monitoring of secondary data (e.g. voting records, volunteer records, composition of decision-making bodies, quality of life indicators, community planning documents)  
» Participatory methods (e.g. appreciative inquiry, most significant change)  
» Root cause analysis |
| Building a large-scale social movement | » Movement leaders develop shared motivation, values, and interest  
» Increased mobilization of members and allies  
» Members, volunteers, and constituents engage in movement activities  
» Strengthened alliances and partnerships  
» Number, breadth, and quality of relationships is strengthened  
» Increased trust among movement members  
» Movement organizations increase resource sharing  
» Increased influence on political decision making  
» Increased championing of issues by decision makers  
» Movement members increasingly in positions of power  
» Major policy initiatives advance and are enacted | » Key informant interviews  
» Policymaker ratings  
» Media analysis  
» Case studies  
» Rapid outcome assessment  
» Process tracing  
» Developmental evaluation  
» Monitoring of secondary data (e.g. budgets, resources, policies, regulations)  
» Participatory methods (e.g. appreciative inquiry, most significant change) |
A Construct for Evaluating Leadership Development

**Individual and Organizational Outcomes**

**Individual Capacity**
- Increased awareness of leadership style
- Increased understanding of leadership skills and mindsets
- Changed behavior or sustained growth
- Career advancement

**Organizational Capacity**
- Increased organizational capacity
- Increased learning and innovation
- Improved culture
- Improved organizational infrastructure and technology

**Collective Outcomes**

**Partnership / Networks**
- Increased growth in quality partnerships and local/national networks
- Increased engagement of diverse populations

**Collective Action**
- Increased alignment of goals & resources
- Increased innovation
- Increased visibility and diffusion of ideas
- Amplification of best practices
- Increased mobilization and influence on decision-making

**Building Blocks for Evaluating Leadership Development**

**BUILD CAPACITY**

**CREATE NETWORKS**

**MOBILIZE NETWORKS**

**More Effective Individual Leaders**

**More Effective Organizations**

**Stronger Networks**

**More Effective Networks**
This generic leadership theory of change is comprised of the building blocks introduced in Part 1 of this guide. Each block is firmly grounded in theory and practice. Not all leadership development efforts are built using the same blocks or the same configurations, but we hope that you can see aspects of your own leadership development programs in this visual, and that it will help sharpen your thinking about how you believe leadership contributes to social change.
Implementing Evaluation for Leadership Development

Individual and Organizational Capacity for Leadership
Collective Capacity for Leadership
Social Change Levers
Leadership development programs are as varied as the funders that support them. No two will have the same structure, content, or aspirations, yet they are likely to share common elements. We hope you can see aspects of your investments and aspirations reflected in the theory of change introduced in Part 1.

While we believe that it is best to begin with a clear representation of a program’s theory of change when designing an evaluation, theory is often better suited to identifying what types of changes are meaningful rather than for selecting an appropriate design. Using the same building blocks and social change levers introduced earlier, this section dives deeper into measurement to help you customize an evaluation plan that fits your program. For each building block, we provide practical considerations for shaping a design as well as resources to help select the best measures and methods for a given context and purpose.

Individual and Organizational Capacity for Leadership

Because individuals are always a central focus of leadership development, evaluating changes in the leadership capacities of individuals as a direct result of leadership investments remains of central importance. Most theories of change include individual leadership, which suggests a focus on leadership skills and knowledge enhancement. If your theory of change operates through effective organizations, evaluation might also focus on changes in internal or external leadership practices and perceptions of the organization as a leader in a sector or on an issue. In both cases, thinking clearly about what changes we want to see as the result of our leadership development is a critical starting point.
**Key Considerations**

**Tie outcomes to stage of development**

Expected outcomes associated with a leadership development program, as well as information needs, may vary considerably at different points in time. While some changes might occur as the result of a single training or event, others may take time or repeated exposure. Evalulead (Grove et al, 2005) introduces a framework for thinking about change:

**Episodic.** Well-defined, time-bound results stimulated by actions of the program or its participants and graduates, such as knowledge gained from a training or organizational collaboration on specific projects.

**Developmental.** Changes that occur across time, such as sustained change in individual behavior and mindsets, or changes in organizational decision making, reflection, or actions.

**Transformative.** Fundamental shifts in values and perspectives that seed the emergence of fundamental shifts in behavior or performance, such as substantial shifts in viewpoint, vision, paradigms, or career shifts, or fundamental shifts in how an organization is perceived.

As training is being delivered, episodic measures might be the most appropriate. In contrast, at the end of a program, or five years after, measurement might focus on transformative change. When developing an evaluation approach over time, we recommend carefully considering when and how to look for change, and paying explicit attention to when it makes sense to collect different types of data.

**Use culturally responsive approaches**

Leadership development takes place in a social, political, and cultural context. It is important for both funders and evaluators to understand the different life experiences of participants, and the different histories of the communities in which they live and work, so that they can honor these differences and respect these histories.

Culturally responsive evaluation demands attention to how the evaluation is staffed, how power is held or exercised, whose values and interests are reflected in the evaluation, what methods are appropriate, how samples are selected, how to best engage with stakeholders throughout the evaluation, whose perspectives are included in sense-making, and who is given the power or responsibility to act on the results. (Public Policy Associates, 2015).
Organizations vary considerably in size, structure, and culture. Responsive evaluation demands attention to the context in which organizational leadership development is being supported.

Youth leadership development is a special case of leadership development. Outcomes for youth leadership need to be developmentally appropriate. A related consideration is level of experience; some investments target “thought leaders,” or those individuals who sit at the pinnacle of their chosen fields, while others may give opportunities to individuals who are still emerging in their leadership practice.

A diversity of measurement approaches is the norm in leadership development, and for good reason. Mixed methods bring a broader range of evidence to bear in answering evaluation questions. Mixing qualitative and quantitative methods can enrich understanding, support assessment of a broader set of outcomes, and strengthen findings. Being clear about the types of changes to evaluate and the purpose for evaluation (e.g., accountability, learning, impact) will help guide decisions about the best way to collect evidence.

Tables 4 and 5 provide additional guidance on the selection and use of the methods introduced in Part 1 (Table 1). Table 4 focuses on measurement at the individual level. Table 5 focuses on organizational outcomes, which are relevant if your theory of change includes organizations as change agents toward a social goal.
Table 4 Methods and Data Sources to Evaluate Leadership Outcomes at the Individual Level

<table>
<thead>
<tr>
<th>Method / Assessments of leadership style</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leadership style assessments can provide insights into an individual’s cognitive styles, personality, working habits, and preferences. Awareness of leadership style can be instrumental in improving an individual’s abilities to work effectively in groups or manage other people. They may be particularly relevant for younger and less experienced leaders.</td>
</tr>
</tbody>
</table>

Applicable when...

- Self-awareness is a major focus of leadership development.
- Program participants are younger, less experienced, or less established in their practice.

Implementation considerations

- Leadership style assessments are most commonly incorporated early in the leadership development process, when participants can act on this information during subsequent programming.
- Often, leadership style assessments include a facilitated session to allow leaders to reflect and learn from the assessment results.

Example data sources

- Myers Briggs Test
- Hermann Brain Dominance Instrument
- Kevin Fong’s The Five Elements of Organizational Success
- CliftonStrength
- DISC assessment

<table>
<thead>
<tr>
<th>Method / Reports of leadership program staff and trainers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Those involved in delivering leadership development training can provide valuable insights into the ways that individuals have grown or progressed based on program goals and aspirations.</td>
</tr>
</tbody>
</table>

Applicable when...

- Leadership program staff and trainers have a sufficient degree of touch to be able to provide valuable insights.

Implementation considerations

- Program staff and trainers might best provide input at the end of a training or series of trainings, although there may be a rationale for collection at multiple times to provide a point of comparison (e.g., pre- and post-training).

Example data sources

- Program and trainer reports might take the form of:
  - Program reports
  - Trainer reflections
  - Individual interviews
Self-reports are commonly used to assess changes over time in individuals. Self-reports can measure understanding or skills at a single point in time or explicitly ask respondents to compare two points in time. Traditional pre-post designs are common when using self-reports. Retrospective pre-test is an option. Advantages of this approach include single administration and avoiding the risk that participants will misunderstand terms or concepts. Disadvantages include insufficient recall and respondent desire to please trainers or evaluators.

Applicable when...

- Growth of individual leaders is a major focus of the investment.
- There is interest in linking specific program delivery elements to changes in participants.

Implementation considerations

- Leader self-reports might be carried out at multiple points in time. Pre-tests (baseline) can inform programming and serve as a basis for comparison, with post-tests conducted after key training events or at program completion. Limitations include the burden of multiple administration and the potential for introducing terms and concepts before participants are ready for them.
- Retrospective pre-test is an option. Advantages include single administration and avoiding the risk that participants will misunderstand terms or concepts. Disadvantages include insufficient recall and respondent desire to please trainers or evaluators.
- Individual interviews are more time intensive than surveys. A review of written program reports may be least burdensome for participants, but value is dependent on the quality of the reports.

Example data sources

Self-report might take many forms, including:

- Surveys
- Individual interviews
- Participant journals/stories
- Review of grantee or fellow reports
### Table 4  Methods and Data Sources to Evaluate Leadership Outcomes at the Individual Level (continued)

<table>
<thead>
<tr>
<th>Method</th>
<th>Multi-rater capacity assessments (360° assessments)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>These are often used to measure the extent to which individuals have gained desired knowledge or competence. Multi-rater assessments leverage the observations of peers, supervisors, co-workers, subordinates, or others whose knowledge of the individual leaders in question would provide important perspective on how leaders have changed their attitudes, ways of thinking, or behaviors.</td>
</tr>
</tbody>
</table>

**Applicable when...**

- Leadership development focuses on the acquisition of specific knowledge, skills, and competencies, either among individual leaders in a program or within an organization.
- Multiple people with deep knowledge of the individual leader can be identified to participate in the assessment.

**Implementation considerations**

- Post-assessment might be most useful at the end of a program, or at key program milestones at which there is a reasonable expectation that the program may have resulted in change in knowledge, skills, or competencies. To assess change, the results need to be compared against a baseline collected previously or retrospectively.

**Example data sources**

- Several multi-rater assessment tools exist. While items in these assessments might prove useful, a careful review of the content is merited to identify their fit for a specific program or set of competencies.

*Examples include:*

- Concord Leadership Assessment Survey
- Leadership Practices Inventory
- Benchmarks 360 Assessment Suite
  *(Center for Creative Leadership)*
Career position tracking

One outcome of interest in leadership development is the extent to which individual leaders advance in their careers within their organizational or field, or take on extracurricular assignments, such as serving on boards. Particular attention should be paid to opportunities for advancement of diverse leaders typically underrepresented in leadership positions.

**Method** / **Career position tracking**

**Applicable when...**

» Career advancement and/or movement into positions of influence is a major goal of the program.

**Implementation considerations**

» Often, career growth takes time, potentially months or years after training or participation in formal program activities has ended. It is therefore important that the program remain in contact with participants to be able to follow up with them at later intervals.

**Example data sources**

*Career position tracking might occur via:*

» Surveys

» As part of an individual interview

---

Participatory methods

Participatory methods are particularly useful in meeting leaders on their own terms and identifying how leaders perceive and frame their own growth. It can also provide an opportunity to actively engage stakeholders in the evaluation process.

**Method** / **Participatory methods**

**Applicable when...**

» There is interest in engaging leaders and stakeholders more directly in the evaluation.

**Implementation considerations**

» If stakeholders engage in data collection, training and skills development may be needed, and evaluators will need to adopt the role of coach, facilitator or critical friend.

**Example data sources**

*Appropriate techniques include:*

» Most Significant Change

» Appreciative Inquiry

» Oral histories and stories
### Table 5: Methods and Data Sources to Evaluate Leadership Outcomes at the Organizational Level

<table>
<thead>
<tr>
<th>Method /</th>
<th>Organizational capacity assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>These assessments focus on evaluating the extent to which organizations have specific skills, organizational structures, and practices necessary to operate effectively, including leadership capacity.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Applicable when...</th>
<th>» Leadership development focuses on developing the leadership capacity of an entire organization.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation considerations</td>
<td>» Assessment often occurs at two points (pre- and post) to evaluate whether organizational leaders have gained desired skills and whether the organization itself has assumed leadership on an issue. The two points are generally chosen based on the schedule of specific development activities that are undertaken.</td>
</tr>
<tr>
<td>Example data sources</td>
<td>Numerous organizational capacity assessments exist, including:</td>
</tr>
<tr>
<td></td>
<td>» Organizational Capacity Assessment Tool 2.0 (OCAT) by McKinsey &amp; Company</td>
</tr>
<tr>
<td></td>
<td>» Marguerite Casey Fdn. Organizational Capacity Assessment Tool</td>
</tr>
<tr>
<td></td>
<td>» Race Matters Organizational Self-assessment</td>
</tr>
<tr>
<td></td>
<td>» The Table Group Team Assessment</td>
</tr>
<tr>
<td></td>
<td>» Western States Center organizational assessment tools</td>
</tr>
<tr>
<td></td>
<td>» My Healthy Organization</td>
</tr>
<tr>
<td></td>
<td>» Power Analysis</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method /</th>
<th>Self-report of organizational leaders and staff</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reports of organizational leaders and staff can provide information on the extent to which organizations have made operational or positional changes and the extent to which a robust leadership pipeline has been established.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Applicable when...</th>
<th>» Efforts focus on fostering leadership growth and effectiveness within an entire organization.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>» Organization has previously set leadership performance objectives or outcomes that can be assessed over time.</td>
</tr>
<tr>
<td>Implementation considerations</td>
<td>» Self-reports might be carried out at multiple points during leadership development programs. At baseline and at the conclusion of the training are common times, although interim assessments can be useful to trainers or to assess change at key program milestones.</td>
</tr>
<tr>
<td>Example data sources</td>
<td>Self-report might take many forms, including:</td>
</tr>
<tr>
<td></td>
<td>» Surveys</td>
</tr>
<tr>
<td></td>
<td>» Individual interviews</td>
</tr>
</tbody>
</table>
**Method / Review of organizational internal documents, budgets, reports, or media coverage**

Internal organizational documents can provide insights into resource allocation, changes in organizational policies and practices, organizational performance, leadership succession planning, and external perception of organizational leadership.

**Applicable when...**

- Changes sought are focused on internal organizational operations (resource allocation, partnerships, leadership development, and financial sustainability) or on the position of the organization relative to an issue or field.
- Organization has previously set leadership performance objectives or outcomes.

**Implementation considerations**

- Review of internal documents might occur at multiple points, but often aligns with financial, grant reporting, or other organizational planning or reporting cycles.

**Example data sources**

*Organizational documents might include:*

- Budgets
- Organizational charts
- Leadership succession plans
- Internal policy manuals
- Evaluations or audits
- Annual reports or strategy documents
- Traditional media
- Social media
Collective Capacity for Leadership

If your theory of change rests on investing in leadership to build robust local or national networks to support collective action, evaluation might focus on characteristics of networks, as well as on the ability of collaborators to align goals, coordinate resources, or organize around shared goals.

Key Considerations

Be clear about network purpose

Networks can serve a variety of functions. Clarifying a network’s purpose is an essential precursor to gauging its progress in achieving its goals and for using evaluation findings to support the intentional development of desirable network properties. Common purposes of networks include the following (Plastrik and Taylor, 2014).

- **Innovation.** Generation of novelty (new knowledge, products)
- **Diffusion.** Rapid spread of ideas and products.
- **Combination.** Assembling of new capacities
- **Alignment.** Formation of new identity / brand
- **Mobilization.** Reaching and activating many people
- **Exchange.** Sharing of information widely
- **Assessment.** Provision of diverse feedback/evaluation
- **Advocacy.** Influencing existing decision-making structures
- **Delivery.** Bringing resources and assistance to increase capacity

Consider optimal relationships as they relate to equity and purpose

While the benefits of collaboration are widely recognized, so are the costs. It follows that optimal levels of interaction are contingent on the purpose of a network or collective. For example, a network that is designed to support the diffusion of information might not require as strong ties between members as a network that seeks joint action to deliver a result. Table 6 provides additional guidance on the selection and use of the collective leadership methods introduced in Part 1 (Table 2).
<table>
<thead>
<tr>
<th>Method</th>
<th>Measurement of network composition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Network composition is about ensuring diversity of perspective and positions relevant to framing problems, identifying solutions, and carrying out actions.</td>
</tr>
<tr>
<td>Applicable when...</td>
<td>» Diverse perspectives can enrich the work of the network.</td>
</tr>
<tr>
<td>Implementation considerations</td>
<td>» Might occur at a single point in time or over a period of time to understand whether progress is being made (e.g., toward increasing diversity). » Can be low intensity and require few resources.</td>
</tr>
<tr>
<td>Example data sources</td>
<td>Data might come from: » Meeting attendance sheets » Application forms » Surveys » Meeting observations » Interviews » Community Forums</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Self-report of participants in collective efforts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Participants can provide valuable insights into the ways in which collaborators are sharing information and working together, and what progress has been made toward overall goals.</td>
</tr>
<tr>
<td>Applicable when...</td>
<td>» The focus is on how those involved in a collective effort are working together, strengthening connections, and contributing to desired impacts.</td>
</tr>
<tr>
<td>Implementation considerations</td>
<td>» Self-reports can be carried out at multiple points during leadership development programs as part of ongoing monitoring or to assess changes at key points in time. Official meetings of an initiative or network can provide opportunities to collect member information in specific areas of interest.</td>
</tr>
<tr>
<td>Example data sources</td>
<td>Reports may take the form of: » Surveys » Network scorecard (e.g. Network Health Scorecard available at <a href="http://www.networkimpact.org">www.networkimpact.org</a>) » Interviews » Meeting minutes</td>
</tr>
<tr>
<td>Method /</td>
<td>Network analysis</td>
</tr>
<tr>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td><strong>Network analysis</strong> allows for quantitative assessment of how networks are structured and function. It can provide insights into the extent to which those involved in collected efforts are connected and interacting in desired ways toward the ultimate goal as well as the distribution of power and authority within the network.</td>
<td></td>
</tr>
<tr>
<td><strong>Applicable when...</strong></td>
<td></td>
</tr>
<tr>
<td>» Evaluation seeks to understand, quantify, and visualize how a collective effort functions, including ways of working and depth and quality of relationships.</td>
<td></td>
</tr>
<tr>
<td>» Relationship building is a key outcome of leadership development.</td>
<td></td>
</tr>
<tr>
<td><strong>Implementation considerations</strong></td>
<td></td>
</tr>
<tr>
<td>» Network analysis can occur at different times. If the goal is to understand how leadership development activities may have contributed to network growth, a pre-post design might be warranted. However, if analysis is focused on whether actors are interacting in desired ways, a single point in time design may be sufficient.</td>
<td></td>
</tr>
<tr>
<td>» Network mapping requires the use of software programs, such as Kumu, Gephi or Node XL, to analyze network metrics and create visualizations of the network composition.</td>
<td></td>
</tr>
<tr>
<td><strong>Example data sources</strong></td>
<td></td>
</tr>
<tr>
<td>» Data collection often takes the form of:</td>
<td></td>
</tr>
<tr>
<td>» In-person network mapping</td>
<td></td>
</tr>
<tr>
<td>» Network surveys</td>
<td></td>
</tr>
</tbody>
</table>
Social Change Levers

What changes do you want to see as the result of leadership development? If your theory of change rests on the social change levers that are believed to lead to population-level changes, evaluation might focus on the extent to which barriers to community engagement or decision making have been dismantled, or how leaders have built public and political will to sustain movements and enact policies. Table 7 provides additional guidance on the selection and use of the social change methods introduced in Part 1 (Table 3).

Key Considerations

<table>
<thead>
<tr>
<th>Consider the pace at which social change happens</th>
<th>Identify drivers of change external to the leadership program</th>
<th>Select approaches that look at contribution and attribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social change is often slow. Leaders may collectively work on an issue for years before a movement takes hold or a policy is advanced. Outcomes should be appropriate to the time frame and include early markers of change. For example, if policy advancement is the goal, it may be appropriate early on to measure the extent to which leaders have created the necessary conditions for changes in policy. Interim measures (ORS Impact, 2007) will tell you if you are on the right path and help predict longer-term outcomes.</td>
<td>Social change outcomes typically result from complex processes involving many actors and influences. The complexity of social systems makes causality difficult to determine. By explicitly considering the external forces or drivers of change that have influenced an observed change, the contribution of the leadership program can be placed in a larger context.</td>
<td>To return to the theme raised at the beginning of this guide, looking beyond the immediate outcomes associated with individual leaders to the larger contribution to the issues, sectors, and communities that leaders are trying to affect can help the field learn which short-term and interim outcomes best predict longer-term goals. In evaluation language, this means looking not only at attribution but also at contribution. This has implications for the questions that evaluation seeks to answer as well as the measures and methods selected.</td>
</tr>
</tbody>
</table>
### Table 7  Methods and Data Sources to Evaluate Contribution of Leadership Development to Social Change

<table>
<thead>
<tr>
<th>Method /</th>
<th>Bellwether interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>The bellwether methodology determines where an issue is positioned in the policy agenda queue, how lawmakers and other influentials are thinking and talking about it, and how likely they are to act on it. By definition, bellwethers are gauges of future trends or predictors of future events.</td>
<td></td>
</tr>
<tr>
<td>Applicable when...</td>
<td>Evaluation is focused on determining how key influential people view an issue.</td>
</tr>
<tr>
<td>Implementation considerations</td>
<td>Generally useful in a sector that is more mature. Resource intensive because it involves identifying bellwethers, design of structured interview protocol, and multiple bellwether interviews. May also require use of social capital to gain access to important individuals.</td>
</tr>
</tbody>
</table>
| Example data sources | **Potential data sources:**  
Lawmakers  
Key decision makers  
Industry or sector leaders  
Researchers in the field  
Influencers |

<table>
<thead>
<tr>
<th>Method /</th>
<th>Case studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case study research develops an in-depth understanding of a single or multiple “cases,” set in their real-world contexts. The process aims to produce a deep understanding—that is, an insightful appreciation of the cases—hopefully resulting in new learning about real-world behavior and its meaning. Often, cases can reveal changes that have occurred across multiple domains, and what has enabled or stymied progress in a given context.</td>
<td></td>
</tr>
<tr>
<td>Applicable when...</td>
<td>Evaluation is focused on complex or transformative changes that take place over time and in a specific context.</td>
</tr>
<tr>
<td>Implementation considerations</td>
<td>Resource intensive, involving case identification, development of a sampling strategy, multiple sources of data, analysis of data to draw out themes, and creation of the written product.</td>
</tr>
</tbody>
</table>
| Example data sources | **Potential data sources:**  
Program reports  
Trainer reflections  
Individual interviews  
Oral histories |
Contribution analysis

Contribution analysis is a step-by-step process for assessing causal questions in program evaluations. The value of contribution analysis is in reducing uncertainty about the contribution an intervention is making through an increased understanding of why the observed results have occurred and the roles played by the intervention and other internal and external factors.

Applicable when...

» Contribution analysis is less suitable for understanding whether a result has occurred as a direct result of the effort. It is more useful when thinking about whether an effort has contributed to an observed result, and understanding the nature of the contribution.

Implementation considerations

» Most suited for leadership development efforts that have a clearly articulated theory of change and a clear outcome of interest to focus on.

Example data sources

.Primary and secondary data sources that may be relevant include:

» Previous evaluations
» Program documents
» Individual interviews
» Workshops with key informants

Key informant interviews

These are qualitative in-depth interviews with people who know what is going on in the area of inquiry. The purpose of key informant interviews is to collect information from a wide range of people with firsthand knowledge about the community, sector, or issue that can provide insight on the nature of problems and understand what progress has been made. These can be with grassroots or grasstops.

Applicable when...

» Changes occur in a complex environment, and change is sought at multiple levels.

Implementation considerations

» Generally, more applicable after a leadership development initiative is more mature.
» Resource intensive as it involves identifying key informants, multiple in-depth interviews, and analysis of data to identify key themes.

Example data sources

Interview samples may come from any of the following sources:

» Individuals representing advocates in an issue area
» Individuals representing a sector
» Individuals encompassing different parts of a movement
» Community engagement sessions
### Table 7: Methods and Data Sources to Evaluate Contribution of Leadership Development to Social Change (continued)

<table>
<thead>
<tr>
<th>Method /</th>
<th>Media analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Media analyses can be used to identify messages, examine how those messages are framed, and see how existing coverage of an issue could be improved. These analyses entail systematically taking a “slice” of media coverage from a set timeframe.</td>
</tr>
</tbody>
</table>

**Applicable when...**
- Increasing the public’s awareness, or perception, of an issue is a major goal.
- Leadership initiatives have sufficient scale and reach to influence media coverage.

**Implementation considerations**
- More applicable for mature leadership development efforts.
- May require substantial resources depending on the time period covered and the number of sources reviewed.

**Example data sources**
- Traditional and earned media, print and online
- Social media and blogs

<table>
<thead>
<tr>
<th>Method /</th>
<th>Monitoring secondary data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Policy changes, resources, and program and policy implementation information (e.g. budgets, resources, policies, regulations) may be available in publicly accessible documents. Monitoring of policy changes, budget allocations, and philanthropic investments can help to identify larger population-level changes that have occurred.</td>
</tr>
</tbody>
</table>

**Applicable when...**
- Leaders, organizations, or networks have contributed to a policy win, gained new resources for their efforts, or contributed to community impact.

**Implementation considerations**
- More applicable for mature leadership development efforts or those that have greater reach and/or influence.
- May be most useful at the end of a legislative session.

**Example data sources**
- Examples include:
  - Legislation passed
  - Budgets
  - Philanthropic investment
### Method / Evaluation Tools

#### Movement Capacity Assessment Tool

New tools are being developed to systematically assess the strength of social movements. One example is the Movement Capacity Assessment Tool currently being piloted by the Global Fund for Women. Another example is Our Health Alliance, a tool developed by RoadMap and Movement Strategy Center designed specifically for alliances engaged in social change and movement building.

#### Applicable when...
- The goal of leadership development is to support movement building.

#### Implementation considerations
- Both a tool for assessment progress toward movement building, and for guiding strategy.
- The tool is set to be released in 2018.

#### Example data sources
- Data will come from individual leaders, or organizations involved in the movement.

---

#### Policymaker ratings

Developed by the Harvard Family Research Project to evaluate a state-level policy change campaign, this method involves having evaluators facilitate a rating process with a group of advocates to gauge three items for each decision maker in a decision-making body:
- Policymaker level of support
- Policymaker level of influence
- Rater level of confidence

#### Applicable when...
- The goal of leadership development is ultimately to influence decision makers to adopt and implement improved policies.
- Generally, more applicable when collective efforts have gained influence.

#### Implementation considerations
- Resource intensive because it involves an intensive design process and a facilitated session with advocates.

#### Example data sources
- Data comes from a carefully selected group of advocates with deep knowledge of the policy environment.
Table 7  Methods and Data Sources to Evaluate Contribution of Leadership Development to Social Change (continued)

<table>
<thead>
<tr>
<th>Method</th>
<th>Population-level surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Population-level surveys can be valuable for assessing whether change has occurred within an entire sector, community, or region, or among a group of intended beneficiaries.</td>
</tr>
</tbody>
</table>

| Applicable when... | · There is reason to believe that large-scale impact has been achieved.  
|                   | · Publicly available survey does not address the change of interest |

| Implementation considerations | · Likely to be resource intensive depending on the size of populations involved.  
|                               | · Surveys may suffer from low response rates if linguistic, cultural, and technological considerations are not incorporated into survey design. |

| Example data sources | Example populations include:  
|                     | · Sectors of interest (e.g., nonprofits working on food security in East Africa)  
|                     | · Communities of interest (e.g., low-income youth in Detroit)  
|                     | · Advocates working on state-level funding of early learning |

<table>
<thead>
<tr>
<th>Method</th>
<th>Process tracing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Process tracing is a case-based approach to causal inference which focuses on the use of clues within a case (causal-process observations) to adjudicate between alternative possible explanations.</td>
</tr>
</tbody>
</table>

| Applicable when... | · Multiple factors might have contributed to success. It can help to rule out alternative explanations for observed results and make a case that there is a causal relationship between an intervention and the desired result. |

| Implementation considerations | · Can be resource intensive because it may necessitate the gathering of substantial diagnostic evidence that will then be the subject of testing to confirm or disconfirm the causal hypotheses. |

| Example data sources | Potential data sources generally come from “prior knowledge,” including:  
|                     | · Conceptual frameworks  
|                     | · Recurring empirical regularities  
|                     | · Theories that weave together empirical regularities |
### Rapid Outcome Assessment (ROA)

This methodology assesses and maps the contribution of a project’s actions on a particular change in policy or the policy environment. It is a flexible and visual tool that can be used in conjunction with other evaluation tools and options.

<table>
<thead>
<tr>
<th>Applicable when...</th>
<th>Most applicable for assessing changes in policy or the policy environment.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation</td>
<td>Requires a facilitated workshop to identify key factors in the policy</td>
</tr>
<tr>
<td>considerations</td>
<td>environment, changes that have occurred, and a timeline for changes. The</td>
</tr>
<tr>
<td></td>
<td>workshop is used to target in-depth interviews to further refine</td>
</tr>
<tr>
<td></td>
<td>hypotheses and assess contribution to change.</td>
</tr>
<tr>
<td>Example data sources</td>
<td><strong>Potential data sources include:</strong></td>
</tr>
<tr>
<td></td>
<td>Background research on the policy context</td>
</tr>
<tr>
<td></td>
<td>Report of participants with knowledge of policy change and the policy</td>
</tr>
<tr>
<td></td>
<td>environment</td>
</tr>
<tr>
<td></td>
<td>Rapid Outcome Assessment Toolkit, Overseas Development Institute</td>
</tr>
</tbody>
</table>
3

Putting It All Together

Three Real-life Case Examples
Hypothetical Case Example
Evaluation of leadership in the social sector is complicated. Assessing the contribution of leaders to the issues, sectors, and communities they are trying to affect involves making hard choices about what questions to ask and what approaches to use. In this section, we provide concrete case examples to assist in achieving the following objectives:

- Help select a design and methods that align with program goals, the beliefs about how change occurs, and the program context.
- Help the field advance its understanding of how leadership development produces social change outcomes.

Our goal is to illustrate how the guidance provided in Parts 1 and 2 can be used to support evaluation. First, we briefly describe three fellowship programs with distinct approaches to evaluation as a direct result of their divergent approaches (and theories of change) to building leaders for social change. Then we present a hypothetical leadership development initiative, and demonstrate how variations in the program design and context affect evaluation design choices.

**Three Real-Life Case Examples**

The first two leadership development programs described on the next page share common features: both are leadership fellowships that strive to make lasting social change in priority communities and both use a cohort model, explicitly building networks of leaders as a strategy for social change. The third example focuses on a specific sector while also building local capacity through a train-the-trainer model. The differences across these three examples are important for evaluation design.
Leadership program description

The W.K. Kellogg Foundation’s Community Leadership Network explicitly seeks to build a network of community leaders who can improve outcomes for vulnerable children. The premise is that investing in the individual development of a diverse set of community leaders—building skills and personal awareness, instilling a systems thinking mindset, intentionally bringing a racial equity and healing lens to the work—and mobilizing networks among fellows and community partners will result in a proactive network of leaders who bring together diverse communities to advance public policies, actions, and practices that change conditions for children and their families.5

Sample Evaluation Questions

» How did fellows grow as leaders?
» Are leaders building networks that can amplify impact?
» What changed for communities and children as a result of CLN?

Approach to evaluation

The evaluation design recognizes the centrality of both individual growth and strong networks. Deliberately absent from this design is a focus on organizational-level measurement, because the leadership model is centered on fostering a network of individual leaders rather than organizations. In other words, the theory of change for the program is that more effective individual leaders who form strong, active networks in their communities and nationally will increase community self-determination, strengthen work on behalf of vulnerable children, and lead to more equitable outcomes for children. Evaluation of the first class of fellows used the following methods: program document review; interviews with fellows, staff, and mentors; social network analysis; retrospective pre-post survey; and case studies.

<table>
<thead>
<tr>
<th>Barr Foundation</th>
<th>David and Lucile Packard Foundation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barr Fellowship</td>
<td>Leadership Development in Fiji</td>
</tr>
</tbody>
</table>

**Leadership program description**

In contrast, the Barr Foundation fellowship explicitly focuses on leaders within nonprofit organizations, selected for their leadership and organizational contributions. The Barr Fellowship recognizes exceptional leaders; supports their professional development and rejuvenation; invests in their organization’s capacity to distribute leadership, design and implement innovative projects, and work collaboratively across sectors; and cultivates a network of leaders committed to advancing civic leadership in the Greater Boston area.

**Sample Evaluation Questions**

- Are fellows staying longer with their organizations?
- Are organizations increasing capacity to work together to create more community benefit?
- Are fellows bridging across sectors and neighborhoods?

**Approach to evaluation**

The longitudinal evaluation of this fellowship included an examination of individual growth, fellow tenure within their organizations, the strengths of social networks, and the emergence of cross-sector collaboration (Lanfer, Brandes, and Reinelt, 2013). Methods included interviews with fellows, interviews with board and staff members from fellows’ organizations, and formal social network analyses. These choices reflect an underlying theory of leadership that says investing in leaders within organizations and building strong networks will yield an increase in civic leadership in a specific community (Boston) to the benefit of the entire city. Because this program has a theory of change that builds on organizational capacity and leadership, the evaluation includes measures at the organizational level (e.g., organizational tenure), which is not relevant to the Community Leadership Network evaluation.

**Leadership program description**

This leadership development program (2011–2019) aims to create leadership for durable sustainable fisheries management in Fiji.

**The overall goals of the program include:**

1. Developing the leadership and management capabilities of leaders engaged in the fisheries conservation sector in Fiji.
2. Strengthening the capacity of institutions, organizations, and groups to support sustainable development in their communities.
3. Helping establish an enabling environment for leadership development in Fiji.

The current program design was modified in response to a 2014 evaluation. It includes a cross-sector approach, a train the trainer program, a peer mentoring structure, and integration of applied action learning projects. It targets a cross-sector group of leaders (20 each year) that represent the full spectrum of the field engaged in fisheries conservation in Fiji including: non-governmental organizations, government agencies, the private sector, academia, and community leaders.

**Sample Evaluation Questions**

- How did leaders respond and change to specific program elements?
- Is that learning leading to new ideas and solutions?
- What is being done differently as a result of the program?

**Approach to evaluation**

Building on the 2014 evaluation, the evaluation team will be using the EvaluLead framework for ongoing monitoring of the program. This approach will capture episodic changes in response to specific training components as well as developmental and transformational changes that result from the integration of all the design elements over time.
Hypothetical Case Example

Evaluation should be aligned with beliefs about what leadership competencies are important for social change and how those competencies contribute to social change. The design should also reflect community context, the types of change that might be observed at any given point in time, and the many other considerations introduced in this guide. The following case example draws on a fictional leadership program of a fictional foundation based in the Seattle area. We offer this case example to illustrate how evaluation methods and needs may shift over time based on evolving external conditions, level of program development, and funder informational needs. We have found that scenarios can be used as part of an exercise to think through expectations and assumptions early on while there is time to surface important perspectives and shape the evaluation design.

Overview and context
The central mission of The Northwest Foundation is to promote improved educational outcomes for low-income elementary school children in Seattle, WA.

The Foundation historically funds:

1. Educational advocacy organizations that work to secure increased resources for education.
2. Local community-based organizations that provide after-school programs and wrap-around services to low-income students.
3. Public schools with a high proportion of students eligible for free and reduced lunch, providing resources for enrichment programs and to hire additional teachers, instructional assistants, and translators.

To build the strength of its core organizational partners, the Foundation has decided to invest in leaders of its core partners through a two-year Seattle Educational Leadership Fellowship.

Leadership program theory of change
The theory of change for the Seattle Educational Leadership Fellowship is shown on the right. This theory of change builds on the building blocks introduced on pages 20-21. Using this theory of change, the following three scenarios show the leadership program at different stages of time and at varying scale. The evaluation design changes with each scenario.
Theory of Change: Seattle Educational Leadership Fellowship

Outcomes of fellowship program

- Increased awareness of leadership style
- Increased understanding of leadership skills & mindsets
- Changed behavior
- Career advancement
- Increased organizational capacity
- Increased learning and innovation
- Increased growth in quality partnerships
- Increased alignment of goals and resources
- Increased innovation
- Amplification of best practices
- Increased mobilization and influence on decision-making

GOAL: Improved educational outcomes for low-income elementary students
Scenario 1

Program Initiation

Leadership program description
The inaugural cohort of the Seattle Educational Leadership Fellowship is comprised of 50 mid-career professionals representing partners from all three of the Foundation’s core strategies. Over a two-year period, the 50 fellows will meet twice monthly for four hours. During the first year, the sessions will focus on capacity building and leadership development, with a specific focus on building up fellows’ understanding of systems thinking, structural racism, effective communications, and their own leadership style. During the second year, fellows will work in groups to develop a pilot project proposal related to elementary education. At the end of the fellowship, each group will pitch their idea to a group of philanthropic funders. Successful pitches will receive up to $100,000 to implement their project.

Direct outcomes of interest
While the projects themselves are a primary outcome of the fellowship, the Foundation hopes that the benefits of the fellowship will go beyond any individual project. First, they hope to increase the capacity of leaders involved in the program, thereby increasing the effectiveness of their core organizational partners. The Foundation also hopes to foster connections between organizations and sectors, leading to improved collaboration and alignment among partners.

Social change levers
The end goal of this work is to build diverse leadership in the education sector to support system reform that benefits elementary school students in such a way that income is no longer predictive of academic outcomes.

Evaluation questions
The evaluation partner for the two-year fellowship will assess outcomes at the end of the first year and again at the end of two years. The primary evaluation questions of interest during this timeframe are:

» How and to what extent are fellows building leadership capacity individually?
» How are fellows connecting with the other fellows in the cohort?
» To what extent do group pilot projects represent important innovations in the education sector? How responsive are they to community needs?

Key considerations
» Fellows are mid-career professionals.
» The purpose of networking is to support innovation, alignment, and exchange.
» Early in the program, episodic outcomes associated with training might be most relevant.
## Evaluation Design

<table>
<thead>
<tr>
<th>LEVEL</th>
<th>Evaluation Questions</th>
<th>Measures</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individual</strong></td>
<td><strong>YEAR 1</strong>&lt;br&gt;To what extent are fellows building capacity and leadership skills?</td>
<td>Increased understanding among fellows of:&lt;br&gt;» Systems thinking&lt;br&gt;» Effective communications&lt;br&gt;» Their leadership strength and weaknesses</td>
<td>» Assessments of leadership style&lt;br&gt;» Fellow self-reports (pre-and post-survey)&lt;br&gt;» Multi-rater capacity assessments (360⁰ review)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>YEAR 2</strong>&lt;br&gt;To what extent do fellows demonstrate sustained growth as a result of their participation in the fellowship?</td>
<td>Increased feeling of self-efficacy&lt;br&gt;» Improved managerial ability&lt;br&gt;» Increased social capital</td>
<td>» Interviews with fellows after their group pitch meeting</td>
</tr>
<tr>
<td></td>
<td>To what extent are fellows forming new relationships? Are they being sustained?</td>
<td>Increased number of relationships among fellows&lt;br&gt;» Deepened relationship among fellows</td>
<td>» Pre- and post-network analysis&lt;br&gt;» Level of collaboration rubric</td>
</tr>
<tr>
<td><strong>Collective</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Social Change Levers</strong></td>
<td>To what extent do the group pilot project ideas represent valuable innovations for the education field?&lt;br&gt;How responsive are the innovations to community needs, especially the families whose children need the most support?</td>
<td>Amount of external funding leveraged for pilot projects&lt;br&gt;» Increased number of new service models, programs, and innovations resulting from new projects&lt;br&gt;» Alignment of new innovations with community needs</td>
<td>» Secondary data&lt;br&gt;» Interviews with fellows&lt;br&gt;» Focus groups with funders participating in pitch meetings&lt;br&gt;» Focus groups with parents and community leaders</td>
</tr>
</tbody>
</table>
Key changes in the leadership development program
Two cohorts (100 fellows) have graduated from the program. Ten pilot projects have been completed; two have gained traction and achieved scale. The first, an innovative collaboration between a public elementary school and a group of community-based organizations providing academic support to students within their own homes, has been able to show positive outcomes for students and is currently in the process of being scaled up in an additional three schools. The second, a teacher training model that helps teachers meet the needs of children who have suffered trauma, has been widely lauded and has been implemented in 15 schools across the district.

Changes in the evaluation context
At this point in time, the Foundation is interested in understanding
> how pilot projects, both those that have been scaled and those that have not, are contributing to innovation in the education sector,
> how parents (and students) are responding to the innovations,
> the career trajectory of individuals from the first class of the fellowship (one of the fellowship goals was to create a bench of leaders ready to step into high-level organizational roles), and
> how fellows are contributing to increased organizational effectiveness within their organizations (one of the fellowship goals is to increase organizational capacity).

Key considerations
> Transformative changes among individual fellows are now of interest.
> Culturally responsive approaches are relevant for assessing alignment of innovations with community needs.
> A mix of methods overall and as part of the case studies is well suited to answering the complex set of questions.
## Evaluation Design

<table>
<thead>
<tr>
<th>LEVEL</th>
<th>Evaluation Questions</th>
<th>Measures</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individual</strong></td>
<td>To what extent has the fellowship led to career advancements?</td>
<td>» Increased number of fellows who have advanced in their careers</td>
<td>» Interviews with fellows</td>
</tr>
<tr>
<td></td>
<td></td>
<td>» Increased number of fellows seen as leaders by their peers</td>
<td>» Review of websites of fellows’ organizations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>» Survey of fellows’ peers</td>
</tr>
<tr>
<td><strong>Collective</strong></td>
<td>To what extent have core partner organizations become stronger as a result of the program?</td>
<td>» Increased infrastructure and capacity</td>
<td>» Self-report of organizational staff</td>
</tr>
<tr>
<td></td>
<td></td>
<td>» Increased learning and innovation</td>
<td>» Organizational capacity assessment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>» Increased number and quality of partnerships</td>
<td></td>
</tr>
<tr>
<td><strong>Social Change Levers</strong></td>
<td>To what extent are pilot projects contributing to innovation in the education sector?</td>
<td>» Increased innovation</td>
<td>» Case studies: two scaled pilot projects and two that were not scaled</td>
</tr>
<tr>
<td></td>
<td>How do parents (and students) of all backgrounds experience the innovation?</td>
<td>» Increased adoption of beneficial practices</td>
<td>» Interviews with:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>» Improved alignment of programs with community needs</td>
<td>» Partner organizations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>» Parents</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>» Bellwethers</td>
</tr>
</tbody>
</table>
Scenario 3
Ten Years Later and Broader Scale

Over the preceding 10 years, the Seattle Educational Leadership Fellowship has grown into The Northwest Foundation’s flagship program.

Key changes in the leadership development program
With the support of additional funders, the model has been scaled to three other populous cities in Washington State and currently includes a total membership of 400 fellows. As the program has continued, the Foundation has made a concerted effort to facilitate connections between different cohorts. By forming relationships across cohorts and geographic areas, the fellows now collectively represent an active coalition of leaders committed to achieving educational policy reforms that are responsive to community needs. In fact, the fellowship network has become something of a go-to group for policymakers at the city, state, and county levels. The network has participated in both the development of responsive legislation and successful campaigns to bring them to fruition.

Changes in the evaluation context
The network’s rise to prominence coincides with an array of educational reforms at both the state and local levels. While the Foundation knows it cannot take sole credit for these major systems changes, they believe they have made a significant impact. The goal of the evaluation is to explore the extent to which the network has contributed to or accelerated the changes that have been achieved.

The evaluation is aimed at two central questions:

- How has the Foundation contributed to the emergence of a strong network?
- To what extent and how has the network contributed to policies and sector reforms linked to improved educational outcomes for low-income elementary school children?

Key considerations
- The network of fellows is now of prime interest, rather than individuals and organizations, and methods should be aligned to the network purpose, which has expanded to include joint action.
- Methods that look at contribution are relevant now that enough time has elapsed to see changes in the social levers and outcomes.
Evaluation Design

<table>
<thead>
<tr>
<th>LEVEL</th>
<th>Evaluation Questions</th>
<th>Measures</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collective</td>
<td>How has the Foundation supported the strengthening of the network?</td>
<td>✷ Facilitators of improved collaboration among partners</td>
<td>✷ Social network analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>✷ Increased alignment of goals and resources</td>
<td>✷ Interviews with Foundation staff and selected network members</td>
</tr>
<tr>
<td></td>
<td></td>
<td>✷ Amplification of best practices</td>
<td></td>
</tr>
<tr>
<td>Social Change Levers</td>
<td>How has the fellowship network contributed to educational reform at the state and local levels?</td>
<td>✷ Increased number of policies implemented</td>
<td>✷ Contribution analysis</td>
</tr>
<tr>
<td></td>
<td>How have policies related to educational outcomes been advanced?</td>
<td>✷ Increased community voice in sector and policy reforms</td>
<td>✷ Survey of educational leaders, bellwethers</td>
</tr>
<tr>
<td></td>
<td>How have community needs been represented in the reforms and policies?</td>
<td></td>
<td>✷ Interviews with fellows/network members</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>✷ Document review (legislation, meeting minutes)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>✷ Interviews with community leaders and parents</td>
</tr>
<tr>
<td>Large-scale Social Change</td>
<td>Do students have better academic outcomes? Which students?</td>
<td>✷ Improved educational outcomes for students</td>
<td>✷ Student survey</td>
</tr>
<tr>
<td></td>
<td>To what extent have inequities been addressed?</td>
<td>✷ Improved post-secondary outcomes for students</td>
<td>✷ Document review (student performance data, post-secondary education and employment data)</td>
</tr>
<tr>
<td></td>
<td>Are there any unintended outcomes that adversely affect students, families or communities?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conclusion

What these three scenarios demonstrate is that the questions and methods appropriate for evaluating a given leadership program should be guided by the theory of change as well as purpose and context.

The leadership development field is not static. Beliefs about who leaders are, how they lead, and how their leadership affects outcomes continue to evolve. Likewise, the role of philanthropic funders in leadership development is undergoing self-examination, with some funders openly addressing questions about their role in supporting leaders, especially those that are working to change systems that are resistant to change. Some funders will likely continue to see leadership development as an important strategy for supporting the achievement of their social change goals. Other funders may increasingly attempt to break down this traditional power dynamic, and place more responsibility and decision-making power with individual leaders and within the communities and sectors they wish to support.

Good evaluation practice demands that evaluators pay attention to these dynamics and their consequences, collecting data and engaging in sense-making in ways that are both sensitive to these dynamics and that critically question values and assumptions. Evaluators must simultaneously attend to the key attributes of the leadership development effort, the social context in which it takes place, and the broader purpose that leadership is intended to address. By bringing a critical eye to leadership development, evaluation can support learning and advance the field. This guide was prepared in the spirit of promoting shared learning. We invite others to use the framework and the resources provided, and adapt and build on them in their own work.
Bibliography


Building Movement Project. Race to Lead: Confronting the Nonprofit Racial Leadership Gap. 2017


CliftonStrengths available at www.gallupstrengthscenter.com


Concord Leadership Assessment Survey. Available at https://www.concordconsulting.com/tools/360-degree-feedback-tool/


Bibliography

Hernez-Broome, Gina; Hughes, Richard L. Leadership Development Past, Present, and Future. Center for Creative Leadership.


Keleher, Terry; Leiderman, Sally; Meehan, Deborah; Perry, Elissa; Potapchuk, Maggie; Powell, John; Yu, Hanh Cao. Leadership and Race: How to Develop and Support Leadership that Contributes to Racial Justice. Leadership for a New Era Series, Leadership Learning Community, July 2010.

Kirkpatrick, Donald; Winfrey, Elaine C. Kirkpatrick’s Four Levels of Evaluation.
Bibliography


Leadership Practices Inventory: Available at http://www.leadershipchallenge.com/professionals-section-lpi-about.aspx


Bibliography


Meehan, Deborah; Reinelt, Claire; Leiderman, Sally. Leadership and Large-Scale Change: How to Accelerate Learning and Deepen Impact. Leadership Learning Community, 2015. Downloaded from http://leadershiplearning.org/leadership-and-large-scale-change-how-to-accelerate-learning-and-deepen-impact


My Healthy Organization available at http://myhealthyorganization.roadmapconsulting.org

Organizational Capacity Assessment Tool. Available at http://caseygrants.org/resources/org-capacity-assessment/


Our Healthy Alliance available at http://ourhealthyalliance.roadmapconsulting.org


The Table Group Team Assessment. https://www.tablegroup.com/teamwork

Torres, Roselinde; Reeves, Martin; and Love, Claire. Adaptive Leadership. Boston Consulting Group.
Bibliography


The primary authors of this brief are Carlyn Orians, Alex Chew, and Paula Rowland from ORS Impact.

The authors gratefully thank the staff of the David and Lucile Packard Foundation—in particular Jamaica Maxwell, Linda Baker, and Meredith Blair Pearlman—for seeing the need for this brief and placing their trust in us. We also owe a debt of gratitude to the staff at the W.K. Kellogg Foundation who allowed us to serve as their evaluation partner on the most recent chapter of their significant investment in leadership development. We would be remiss in not recognizing another debt of gratitude to the excellent work of members of the Leadership Learning Community and their partners, in particular Claire Reinelt, for thoughtful writings on leadership development and their tireless commitment to sharing learnings for the betterment of all of us who are committed to leadership for social good. Special thanks go to Sarah Stachowiak, Analissa Iversen and Tracy Hilliard of ORS Impact who provided invaluable review comments. Finally, we would like to thank Tracy Wilson for lending us her editing and proofreading skills.

Financial support has been provided by the David and Lucile Packard Foundation.