

## Chapter Nine

**Doing Evaluation Differently**

Sally Leiderman

*“We measure what we value.”*

This chapter focuses on evaluation as a part of community building work. People’s ideas about what matters in evaluation are part of their worldviews. And people’s worldviews are shaped by white privilege and internalized superiority and racism, by their training and life experiences, and by the credence they give to different ways of knowing. Some people trust data reported from government sources, in the media or by community organizations. Some people trust particular data from some of these sources but not all of them. And some people do not trust these sources at all.<sup>2</sup> Some people can make meaning from what they themselves observe and can generalize from that meaning to broader understandings; some people do not trust conclusions drawn in this way. Some people are more trusting of what they can “feel”; for some people that phrase has no meaning in terms of a way of gaining insight or information. All of these concepts are different ways of knowing.

Our ways of knowing and our trust in others’ ways of knowing depend, today at least, in large part on the way our professional and personal lives have been shaped by racism, other methods of oppression, white privilege and access to power. So racism, other methods of oppression, white privilege and access to power always influence evaluation. They influence the questions we choose to ask, the information we trust, which findings we decide are important or unimportant, how we make meaning of results, etc.

*We are using the term ‘evaluation’ broadly to describe the ways in which community residents, funders and others seek to understand what they are accomplishing, measure their results and hold themselves accountable for doing what they intend. In evaluation jargon, this could include: outcome, impact, process and implementation evaluations, self-assessment, participatory evaluation, learning circles and other ways people gather and share information to see how things are going. Methods can include formal and informal information-gathering and making meaning of information through reflection, qualitative and quantitative analysis, synthesis, storytelling, dialogue, etc.*

We also want to note that evaluation as practiced in most community building work today generally reflects Western ideas about cause and effect relationships. That is, most methods are based on a philosophy that change can be understood as a rational and somewhat linear process. For example, most evaluation methods

Our ways of knowing and our trust in others’ ways of knowing depend ... in large part on the way our professional and personal lives have been shaped by racism. ... So racism, other methods of oppression, white privilege and access to power always influence evaluation.

<sup>1</sup> Marc Miringoff, *The Social State of Connecticut*. (New York: Fordham Institute for Innovation in Social Policy, 2000).

<sup>2</sup> For example, readers may be familiar with the phrase “Figures don’t lie, but liars can figure,” which has always tempered my own reliance on quantitative data.

assume that one can predict what is likely to happen in the future based on what is happening now. Evaluators sometimes talk about the trajectory of change, and what we often do is look at short-term outcomes, hoping they will predict “success” against longer-term ones that can often not be measured in the studied time period.

This is particularly problematic for evaluations looking at eliminating white privilege or racism. We can look at behavioral change of individuals and institutions, policy-level change, and many other intermediate steps. But given that we have not yet eliminated white privilege anywhere in the United States, and we do not actually know what it will take to do that, our ability to know what it will take in the short-term to achieve certain results in the long-term is quite limited.

Thus, evaluation in this instance is a pretty speculative notion. So, the first lesson of this chapter is that we should take a long, hard look at what we think evaluation is, and whether its basic premises—that current methods enable us to track progress and measure results, and that it is important to do that—hold up to careful scrutiny.

Even with its limitations, evaluation can be a tool that helps communities and their partners do community building work. The steps people take to design an evaluation can clarify what they hope to accomplish. The upfront evaluation design conversations can also help set reasonable expectations among community building partners for how long and how deep a change is likely. Further, evaluation can be one of the processes that holds the partners in a community building effort accountable for doing what they say they will do.

Reflective kinds of evaluations can surface insights, lessons and opportunities to move work forward. Participatory evaluations and data-sharing can offer another avenue for people to engage with a community building effort. There are also job and career opportunities in evaluation, and many chances to share skills that people who are part of community building work, but not in formal evaluation roles, can incorporate into their work. For example, as they are participating in evaluation conversations, community building practitioners often get better at assessing whether their strategies are sufficient to produce the changes they intend.

Evaluations can provide additional evidence that helps people celebrate what they’ve accomplished. They can help people make the case for what they’ve done, support efforts to get more resources to the work and protect parts of the work from being eliminated. We’ve seen evaluations used in all those ways many times. They can also be a major tool for shifting power among various community building players and partners, based on who decides how success will be defined and measured, which processes are being evaluated, who controls the dissemination of information, and, especially, what gets evaluated and what kinds of evidence are given credence.

## Questions, Responses and Suggestions

The ideas in this chapter come primarily from three sources. One is a series of lengthy conversations with Barbara Major about white privilege and cultural and institutional racism in evaluation. Another is work done to develop Evaluation Tools for Racial Equity,<sup>3</sup> a website designed to help communities assess their own progress toward anti-racism and inclusion goals. It is important to note that the website was developed with the help of many advisors and four community-based groups doing anti-racism work: Teaching for Change, IMPACT Silver Spring, ERASE Racism and the South Orange/Maplewood Coalition on Race. The third source is the experience of the author doing evaluations of community building efforts for the past 30 years.

All of these sources inform this chapter. In addition, we considered this information through the lens of the following three questions:

- What are some of the ways in which white privilege and racism are embedded in evaluation as we do it today?
- What would anti-racist evaluation look like? Or, even more strongly, what would evaluation be like if we considered evaluation one of the tools of community building work that could help to dismantle white privilege?
- How would evaluation need to be different from the way it is today to serve that purpose?

The rest of this chapter is divided into two parts. The first lays out some responses to the questions above. We hope these will stimulate thinking among evaluators, community residents, funders and others involved in evaluations about what we can do differently and better. The second part of this chapter provides some specific suggestions that apply these responses to evaluation work.

## Responses

*What are some of the ways in which white privilege and racism are embedded in evaluation as we do it today?*

As Barbara Major points out, evaluations are processes that judge, and in so doing, they explicitly privilege some worldviews over others. She also notes that few evaluations in community building work are reciprocal among communities and the other partners in the work (funders, technical assistance providers, evaluators, state and local systems, government entities, etc.) Even when community residents participate in evaluation, the evaluation is still generally focused on learning about a community's competence, capacity and the fidelity with which community-level participants implement the efforts' plans. Very few evaluations spend equal resources to learn about the competence, capacity and

[Evaluations] can also be a major tool for shifting power among various community building players and partners, based on who decides how success will be defined and measured, which processes are being evaluated, who controls the dissemination of information, and, especially, what gets evaluated and what kinds of evidence are given credence.

<sup>3</sup> Please see [www.evaluationtoolsforracialequity.org](http://www.evaluationtoolsforracialequity.org). The tipsheets from which sections of this chapter were drawn were cowritten by Sam Stephens.

fidelity of other partners—particularly funders. Lack of reciprocal accountability is one marker of white privilege.

There are many other ways in which white privilege is embedded in the underlying (and often unarticulated) assumptions of typical community building evaluations, their methods and in the ways findings are used. Some examples are noted in Figure A (see p. 101).

*What would anti-racist evaluation look like? Or, even more strongly, what would evaluation be like if we considered evaluation one of the tools of community building work that could help to dismantle white privilege?*

Below is a series of ideas about what evaluations of community building efforts might look like, or include, if we applied the insights that are implied by the discussion in Section II of this monograph to community building evaluations. Some of these ideas are already being applied in a few evaluations. However, many are just at the idea stage. We hope that readers will expand this list, refine it, put many of the ideas into practice, and share with us the extent to which they are successful.

Even when community residents participate in evaluation, the evaluation is still generally focused on learning about a community's competence, capacity and the fidelity with which community-level participants implement the efforts' plans.

*Evaluations can more directly track and measure progress toward elimination of white privilege by:*

- Tracking structural and institutional changes, wherever possible, at the community level;
- Designing evaluations that can look at the separate effects of race and class, or by raising the importance of doing so even if they cannot;
- Maintaining a focus on community-level processes and relationships, in addition to individual-level processes and relationships; and
- Including power analyses and tracking changes in the flow of power and its consequences for various groups.

*Evaluations can also help community building partners maintain their focus on the consequences of white privilege and racism by:*

- Tracking changes in community-level outcomes for different racial groups over the long-term (report cards)—not just in the short-term and not just for all community residents as a whole;
- Tracking the extent to which race becomes a less powerful predictor, in a statistical sense, of how people fare as a way of looking at changes in racial equity;
- Tracking a community's relationship to itself (per B. Major);
- Tracking progress toward a community's understanding of white privilege and oppression (per B. Major);

- Having communities evaluate foundations (per B. Major);
- Shifting the power dynamics in evaluation away from privileged institutions and “experts” in evaluation and toward groups of color and individual communities. These shifts could encompass:
  - Establishing what success will mean;
  - Establishing the timetable for looking at various results;
  - Establishing the theory of change that will essentially define the kinds of short-term and long-term outcomes that the evaluation will measure;
  - Deciding what parts of a community building effort to measure and what questions to ask;
  - Deciding what evidence is credible;
  - When evaluation resources are limited, deciding where to use them to get the most benefit from evaluation;
  - Interpreting results;
  - Framing results (for example, showing how institutional policies and practices influence individual outcomes, by presenting the differences in resource allocation across schools—and residential segregation patterns around schools—on the same page as student outcomes by racial groupings);
  - Sharing results.
- Providing back-up information about statistics and how they are derived. This includes clearly laying out the way various statistics are calculated (including data sources, populations that have been included and excluded from the data, and the specific algorithms used to create the statistic being shown);
- Reviewing findings by racially, ethnically or culturally classified subgroups within major racial categories to get more accurate (and less stereotypical) understandings of what different groups may be experiencing (e.g., how Southeast Asians in a community are faring in terms of health, income, stable and affordable housing and educational measures);
- Using evaluation for learning rather than for setting consequences;
- Not using evaluation at all if we cannot figure out ways to address the issues raised above (and other issues of privilege and racism we haven’t listed here).

*How would evaluation need to be different from the way it is today to serve that purpose?*

There are several ways in which evaluation as a practice (and a field) would need to be different if we were to reduce or eliminate its embedded racism, and use it

as a tool to help eliminate white privilege in our community building work. At the very minimum, more of the people involved would have to have a deep understanding of what white privilege is and how it works, a similarly deep understanding of the various mechanics of racism, and a willingness to bring those understandings more fully to evaluation work.

In addition, we would have to change many aspects of our practice. For example, we would need to understand, and probably change, the extent to which we are “privileging” certain ways of knowing and devaluing others. In my work, I have learned to be very skeptical of government and administrative data, particularly as they apply to people of color. People who are undocumented are almost always undercounted in these kinds of data. Community partners have shared many stories of the ways in which census data, high school completion data, special education designations and domestic violence information are misleading, because of the way seemingly objective definitions are applied differently to people of color than to white people. As a result, I have become much more cautious about using these kinds of data in evaluation, and much more careful to understand exactly how they were collected and what judgments go into the summary information that is most often reported. One practice change that flows from this is to always gather stories about the same issue, and to see where the stories and the data align and where they do not.

Another change would be to require that people involved in deciding what to measure, how, and what the consequences are spend time understanding each other’s worldviews, theories of change and analyses of white privilege and structural, institutional and cultural racism. A related change is to make sure that a wide range of perspectives about these things is represented in that discussion, and that the group works hard together to reconcile or agree to live with differences among these perspectives in evaluation design, implementation and the way meaning is made of the results. Two ways to make it more likely that those changes can happen are to broaden the range of people who are considered evaluators (privileging experience and insight as much as academic credentials) and working to bring more people of color into the “professional” evaluation world.

Finally, evaluators would have to decline to participate in evaluations that they believe are structured, intentionally or unintentionally, in ways that maintain white privilege.

## **Suggestions**

### *Using evaluation to help structure different theories of change*

As noted above, people’s ideas about how change happens are based on their analyses, judgments and observations about how the world works. Data and other kinds of evidence help people make these judgments. But people often choose the data and evidence they examine, interpret data differently and draw different conclusions about what data from one situation imply about another. These differences will affect people’s theories of change about, for example, whether or not anti-discriminatory employment legislation translates into more

open access to high-paying jobs for which people are qualified, or if diversity training of banking executives and loan officers results in a re-examination of credit granting policies of lending institutions.

In developing theories of change for evaluation purposes, it is important to look at all the available evidence about how change happens and for whom. It is also important for people to be open to new or alternative theories of change. This is particularly the case for white people who may be very steeped in dominant culture analyses and explanations that tend to assign most of the responsibility for poor outcomes to individual actions. For example, whites may say that families of color should do more upkeep on their homes to improve community housing stock—rather than look for institutional or structural explanations of outcomes, as people in many communities rent their homes and major upkeep is usually the responsibility of the owner, not the renter. People will also vary in the extent to which they trust systems or leaders to do what they say they will do. It is also important to listen to the voices of people who can talk about whether or not strategies designed to benefit them have benefited or harmed them in the past.

In addition, community change of any type is almost never a straightforward and linear process. Changes in community norms, institutions and patterns not only take a long time to unfold, but within any given time period there are likely to be steps taken toward the desired goal and those in the opposite direction. At any given time, if we took a "snapshot" of the community, we might see progress or retrenchment or stalemate or all three at the same time.

That is why it is important to look for overall trends, to use more than one way of observing and recording changes, and to evaluate strategies at several points in time. Strategies addressing white privilege or racism directly are especially likely to generate resistance in the community, particularly as they become more visible and more challenging to the status quo. For example, school district officials and school board members who initially support sharing of school achievement data may withdraw easy access to that information when it is analyzed to reveal a systematic pattern of assigning more substitutes or fewer credentialed teachers to schools with higher numbers of students of color, and the association of these differences with disparities in student outcomes. Collaborations working on improved access to health care may falter when strategies to insure undocumented workers become crucial to further progress.

A theory of change to support evaluation of anti-racism work should include strategies for anticipating, identifying and addressing resistance in the community. That will not only improve the work's overall effectiveness, but also allow the evaluation to include these strategies in its assessment.

*How can we avoid “blaming the victim” when presenting information on poor outcomes for different racial, ethnic, language or immigrant groups in community building work?*

Groups working to eliminate or reduce differences (disparities) in how racial or ethnic groups fare compared to other groups on important outcomes (education, wealth accumulation, health, etc.) report these differences to make their case for

A theory of change to support evaluation of anti-racism work should include strategies for anticipating, identifying and addressing resistance in the community.

change and to track the progress of their work. For example, they may need to show the different rates of graduation from high school for white, African American, Latino/a, Asian and Pacific Islander and Native American students (and they may also need to show differences in rates of graduation within these groups—by school, gender, language primarily spoken in the home, etc.). Groups use these kinds of data to raise awareness and concern, mobilize supporters, call officials and institutions to account and to provide baseline (starting) and follow-up information for evaluation.

This is the challenge of making sure people who view the data can see it in the context of an analysis of institutional (or structural) racism, or within the context of white privilege. The reason this is so important is that, without a context for viewing the data, people will create their own explanations.

In addition to the difficulty in finding accurate and comparable information about outcomes for different groups, there is another major challenge. This is the challenge of making sure people who view the data can see it in the context of an analysis of institutional (or structural) racism, or within the context of white privilege. The reason this is so important is that, without a context for viewing the data, people will create their own explanations. And people without an understanding of the cumulative effects of white privilege, institutional and structural racism will tend to look for individual—rather than institutional or structural—explanations that end up “blaming the victim” for poor group outcomes. That is, people will view persistent and large group differences as being solely the result of attitudes, actions and inherent abilities of the individuals in the group or of a group “culture,” and they will tend to discount or ignore the role of government policies (e.g., redlining), mechanics of resource allocation (e.g., basing school resources on local property taxes), intergenerational opportunities for wealth accumulation (linked to educational opportunity) and cultural norms that reinforce disparate outcomes by racial group (e.g., national ideas such as meritocracy and individuality).

When presenting data that demonstrate differences in outcomes among groups, particularly those that illustrate poor outcomes (such as school dropout rates or business failure rates), it is important to put this information into context. Specifically, it is helpful to provide data supporting an understanding of differences as a result of policies, practices and decisions that are the target for change—consistent with an analysis of white privilege and institutional or structural racism.

For example, high school graduation rate data for white, African American and Latino/a students could be accompanied by information on the number of substitute teachers or credentialed teachers in schools serving high proportions of students of color vs. schools serving high proportions of white students, on the availability (or lack) of opportunities for students to pursue their studies on an alternative schedule that accommodates work, or other structural factors that have been identified as key to student outcomes. Information on small business start-ups and failures for Latino/a-owned enterprises could be accompanied by summaries of the policies of lending organizations on how potential assets and costs are considered in making loans, if those policies are thought to influence the outcomes.

The Project Change groups (and others, we assume) also found it important to “test-market” the presentation of data (report cards, evaluation reports, summary data tables, etc.) to understand the conclusions that key audiences are likely to draw from the data itself, and from how different data are displayed or



grouped on a page (for example, showing rates of graduation by school and resource allocation to schools on the same page) and from the surrounding text. To be most effective, it can help to test materials with people who are likely to agree with our analysis and people who are likely to disagree. They should also be tested with people from different racial, ethnic, language and cultural groups, to see how they respond to the way the information is framed and what they think it implies about what needs to happen next.

*Using evaluation to help set reasonable expectations for what can be accomplished*

One of the most important things evaluation can do is to help lay out for others the amount of change to be expected and how long it may take to accomplish. These expectations can be reflected in the questions, outcomes, findings and in the context in which the findings are presented. Putting these expectations on the table for discussion and negotiation, as part of the evaluation process, makes the evaluation more transparent and more of a tool to address privilege and racism. Doing this gives the players in community building an opening to begin to negotiate the consequences for meeting or not meeting these expectations. Groups use their own experience, their analyses of how power, privilege, oppression and racism work and the experiences of other communities and groups working on the same issues to help set these expectations. Two ways to help people do this are to:

1. Ask all partners in the community building effort to comment on their beliefs about the sufficiency of the proposed strategies to lead directly or indirectly to the proposed outcomes before the evaluation even starts. (For example, if the community building effort is designed to improve the family economic status of many people in a community over five years, ask the partners to share their own perspectives, experience or evidence that the range of proposed strategies is deep enough, broad enough and complete enough to get there.) Consider especially whether or not the activities are powerful enough to overcome resistance and strategies put in place by those working to preserve the status quo. Evidence can include common sense, the community's own experiences in the past, research from a wide range of sources and the experiences of other communities.
2. Ask all partners to clarify their beliefs about which outcomes can be affected by strategies under the direct or indirect control of the people doing the work. For example, suppose one goal is to increase the number of a community's high school graduates of color who attend and succeed in college. The group may be able to ensure (with difficulty) that many teachers are engaged in a process to reflect on the white privilege and racism built into their expectations of students of different backgrounds. They may also organize the community to put pressure on the School District to put Advanced Placement (AP) classes in all of the District's high schools, thus opening up access to eligible students regardless of the school they attend.

These activities could reasonably be expected to lead to the short-term outcome of increasing the number of achieving students of color who

graduate from high school with the expectation that they can attend and succeed in college: a short-term outcome of the strategy to increase the number of students of color in a community who attend and succeed in college.

However, as we all know, many other factors come into play, for example:

- The ability of a community or individual students or their families to finance a student's college education;
- The expectations of the family, the school and the community about whether or not the student is making a good choice to go to college rather than enter the military, take full-time employment or pursue other goals;
- Whether or not the student is accepted into a college in which he or she is likely to succeed;
- National policy regarding college financing, and other policies that affect the opportunities and costs of options for young adults;
- Home responsibilities;
- The health, motivation and resilience of the student; and
- The supports in the community to overcome barriers for individual students.

Expectations about the amount of change and timing of change to reach the desired outcome have to take into account which of these opportunities and challenges a given set of strategies can control or influence, and which they cannot.

## **Conclusion**

It is easy for community builders to want to over-promise what can be accomplished in a given timeframe, especially to key constituencies, because so much needs to be done and the consequences of the status quo are so damaging. In addition, funders who need to justify their support for this work may push communities and grantees to over-promise what they can do—directly or by transmitting their own anxieties. But if all involved in racial equity and inclusion work—groups doing the work, evaluators, funders, constituencies—are willing to put out real and honest expectations about what can be accomplished, and by when, those expectations will eventually become the standard by which our work can be judged.

**FIGURE A: Examples of White Privilege and Racism embedded in typical evaluations of community building efforts**

***Assumptions***

- Initiative-centric evaluations reify the notion that community building is a project, not a process, by assuming that it is possible and appropriate to accurately measure a piece of what happens in communities without taking into account context (per Major).
- Evaluation, as it is currently practiced, tends to reinforce dominant cultural values about who is an expert and what is right or best.
- Evaluations tend to measure goals other than the elimination of white privilege—serving as a distraction that helps maintain white privilege.
- Evaluations tend to judge communities as failures (per Major) if they do not meet timetables for change established outside the community, based on incomplete analyses of, for example, potential resistance to change that genuinely threatens the status quo or how power operates in a community.

***Methods***

- Evaluations often place more weight on findings generated through quantitative data than through qualitative data, without taking into account the validity of different ways of knowing, community assessment of the quality of quantitative data, and what we know about some data that are presented as objective but that are often computed using racialized and subjective algorithms (e.g., how children are classified into special education categories, how credit-worthiness is calculated, which children's test scores are included in high stakes testing results, etc.)
- Evaluations often fail to use methods that would allow us to separate effects of racial categorization from effects of class (or income or wealth accumulation). This makes it harder to identify institutional and other forms of racism as causes of poor outcomes, or to track which changes in institutional policies and practices are most successful at reducing the effects of institutional racism.

***How findings are used***

- Evaluation results are one of the tools that privileged systems and institutions use to reward certain kinds of behaviors and punish others. One example is when funding decisions reward tangible short-term changes that are not likely to lead to longer-term or more powerful changes.

