From *Facilitating Evaluation*
by Michael Quinn Patton (2018, Sage)

5

Engage Through Options

The Second Operating Principle

"Nothing is more dangerous than an idea when it's the only one we have."

Émile-Auguste Chartier (1868–1951)
French philosopher and journalist

***The engage through options principle*:** Provide a process for generating and comparing options, exercising choice, and making decisions.

***Rationale**:* Facilitation should not be a rubberstamping process in which participants go through the motions of adopting predetermined results. The purpose of evaluation facilitation with key stakeholders is to generate possibilities, consider options, and establish priorities. The evaluation facilitator helps identify, clarify, and explain options so that informed choices can be made.

(Continued)

(Continued)

***The principle in practice*:** Deliberating on options and expressing preferences increases participants' understanding of the implications of making certain choices and deepens ownership of the decisions and recommendations that emerge from the process. Facilitation typically involves working with diverse groups of people. Human groups inevitably manifest power differentials, diverse patterns of interaction, variations in emotional engagement, and whatever is brought into the group from the larger society and culture: gender, race, political, social, cultural, and language issues, to name but a few. These issues frame and contextualize evaluations and therefore, inevitably, must be addressed in evaluation facilitation. Skilled evaluation facilitators anticipate and have ways of dealing with whatever emerges at the intersection of society and evaluation on the path to group success in determining what options are most appropriate for the situation at hand. Making informed decisions is empowering.

RESEARCH ON DECIDING AMONG OPTIONS

A great deal of research has been focused on how we choose among alternatives and make decisions about what to do (Kahneman, 2011; Kahneman & Tversky, 2000; Minnich, 2017; Silver, 2012). Facilitators help groups consider options, set priorities, and make decisions. Evaluation facilitators help stakeholders decide what purposes an evaluation will serve and choose among options for fulfilling those purposes. Exhibit 5.1 lays out twelve key findings from research about decision making that can inform facilitating choices among options.

Exhibit 5.1 Research on Deciding among Options

1. Making informed decisions engages and empowers.	Decision making is engaging and empowering when those who are making decisions understand the options before them, can make informed choices, and have both responsibility and authority to do so.
2. Framing matters.	How a decision is framed affects what decision is made.
3. Valuing undergirds choosing among options.	Options are compared using criteria, and criteria express value preferences.

4. Risk assessment is multidimensional.	Logic, cognition, emotions, psychological processes, past experiences, probability calculations, degree of uncertainty, and social–political context, among other factors, come into play as individuals and groups assess risks when comparing options.
5. Comparisons reveal how preferences are weighted.	Assessing a single option (rating how well it meets criteria) is quite different from comparing options (ranking options on multiple criteria).
6. Making choices and decisions involves trade-offs.	Most options pose both advantages and disadvantages, potentially positive outcomes versus potentially negative outcomes, and more of one desirable thing tied to less of another desirable thing.
7. Attributes and criteria vary in how hard they are to assess.	Not all criteria are created equal; how hard it is to decide is a function of how hard it is to asses and weigh criteria.
8. Irrational processes can trump rational ones.	Biases, selection perception, decision heuristics, cultural and political factors, and emotions are powerful influences on how choice is exercised and decisions are made.
9. Decisions are often made with little attention to their long-term consequences.	Decisions often have delayed consequences, so the short-term implications of a decision are given more weight.
10. People vary in how they exercise choice.	On all the preceding premises, people vary in how their behaviors, decision-making processes, and value preferences are expressed, manifest, and used in a group.
11. Decisions tend to satisfice.	Settling on a reasonably acceptable option (*satisficing*) is more practical and efficient than trying to optimize or maximize (Simon, 1978).
12. Decision rules (how a group decides) vary.	When decision rules are unclear, decision making gets muddled (Kaner, 2014, pp. 323–325).

It's a long but important list. Each item can do no more than suggest some aspect of the complex nature of human decision making. That complexity is what makes facilitation necessary to help groups decide among alternative possibilities. It may help to think about a group decision you've been part of and consider how each item below was evident, or not, in that decision process.

WHAT MAKES EVALUATION DECISION MAKING CHALLENGING FOR STAKEHOLDERS

The preceding knowledge about how people and groups make decisions is generic. It affects all kinds of decision making, including making decisions about an evaluation. More specifically, research on evaluation use, and evaluations of evaluations (meta-evaluations), have documented and elucidated the challenges stakeholders face in making decisions about evaluations. Guiding stakeholders through these challenges is the job of evaluation facilitators. Here are seven tips for facilitating evaluation options.

1. *Different evaluation purposes require adapting an evaluation to its primary intended uses.* Differentiating and understanding the implications of diverse evaluation purposes can be daunting for stakeholders. Formative, summative, or developmental purposes and knowledge-generating, accountability, and monitoring uses constitute quite varied options.

2. *Diverse evaluation situations and contexts pose different and often-difficult challenges.* Local, regional, national, and international initiatives vary in complexity; sectors vary in what indicators and issues are primary (education, health, community development, job training, environment, to name but a few).

3. The expansion of evaluation types, models, and approaches offers a wide and ever-increasing variety of options. I identified 80 such options in *Utilization-Focused Evaluation* (Patton, 2008, pp. 300–305).

4. *Evaluator roles vary*. Internal versus external, independent judges versus developmental coaches, critical friends versus accountability reporters, researchers versus designers, and evaluation teams versus individual practitioners are but a few of the role variations that exist.

5. *Programs and organizations display enormous variation*. Different sizes, varying histories, differences in types and nature of funding, leadership and staffing variations, diversity of missions,

roles in the community, and so forth can affect how evaluations are conducted.

6. *Facilitation situations vary*. The numbers, types, roles, and expertise of stakeholders in a group being facilitated affects facilitation options (see Principle 1, being guided by the personal factor).

7. *Process options add layers of complexity*. Process uses (involving stakeholders in learning from the evaluation process) and embedding evaluative thinking throughout the facilitation (Principle 4) mean attending to emergent relationship dynamics and dealing with multiple perspectives.

MYRIAD OF CHOICES

The preceding factors only hint at how many things can affect evaluation choices and make deciding among options challenging for stakeholders. When I'm framing the work that a group will undertake, I often set the stage by comparing designing an evaluation with playing a game of chess. Bruce Pandolfini (1998), a world-class chess master, developed a consulting niche aimed at major corporate leaders to teach them the mindset of a chess master so that they could become more skilled at strategic analysis and thinking. There are some 85 billion ways of playing the first four moves in a game of chess. Deciding what moves to make requires both strategy and tactics grounded in an analysis of the situation presented by a particular game and opponent within an overall framework of fundamental chess-playing principles, understanding what the different pieces do, how they can be moved, and how they relate to each other. Once the game starts, subsequent moves are contingent on and must be adapted to what one's opponent does and the unfolding situation. To become more sophisticated and intentional about situational analysis in evaluation, stakeholders need guidance concerning what to pay attention to because they can't consider everything.

Alkin (1985), in a classic and influential framework, identified some 50 factors associated with evaluation use. He organized them into four categories:

1. *evaluator characteristics*, such as expertise, interactive skills, political acumen, and credibility;

2. *user characteristics*, such as interest in the evaluation, willingness to commit time and energy, and position of influence;
3. *contextual characteristics*, such as size of organization, political climate, and existence of quality information; and
4. *evaluation characteristics*, such as nature and timing of an evaluation report, relevance of evaluation information, and rigor expectations of the evaluation.

Now, the practical problem: How do stakeholders analyze a real-world situation taking into account 50 factors? Those who study decision making say it can't be done, so let's try simplifying. Let's imagine 20 situational variables that can affect how an evaluation is designed and conducted, things such as the number of stakeholders to be dealt with, the evaluation's purpose, staff attitudes toward evaluation, the budget and timeline for evaluation, and the program's prior experience with evaluation. If we conceive of three points (or situations) on each of these dimensions—the two endpoints and a midpoint; for example, low budget, moderate budget, substantial budget—then the possible combinations of these 20 dimensions represent 8,000 unique situational configurations for evaluation. Nor are these static situations. The program you thought was new turns out to have been created out of and to be a continuation of another program; only the name has been changed. Now, in case 8,000 situations to analyze, be sensitive to, and design evaluations for doesn't seem challenging enough, add two more points to each dimension—a point between each endpoint and the midpoint. Now, combinations of the five points on all 20 dimensions yield 3,200,000 potentially different situations. Perhaps such complexity helps explain why the slogan that won the hearts of evaluators in attendance at the pioneering 1978 Evaluation Network conference in Aspen, Colorado, was the lament, "Evaluators do it under difficult circumstances."

Of course, one could make the same analysis for virtually any area of decision making. Life is complicated, so what's new? First, let's look at what's old. The evidence from social and behavioral science is that in other areas of decision making, when faced with complex choices and multiple situations, we fall back on a set of rules and standard operating procedures that predetermine what we will do, which effectively short-circuit situational adaptability. The evidence is that we are running most of the time on preprogrammed tapes. That has always been the function of rules of thumb and scientific paradigms. Faced with a new situation, stakeholders (often unconsciously) turn to old and comfortable patterns. Choosing among options and negotiating trade-offs can be challenging. Exhibit 5.7 at the end of this chapter provides a warm-up exercise to help groups start thinking in terms of trade-offs among options.

The point of this framing is to raise a fundamental question: How can evaluation facilitators prepare themselves to deal with a lot of different people and a huge variety of situations? The research on decision making says we can't systematically guide stakeholders through consideration of every possible variable, or 50 variables, or even 20 variables. What stakeholder groups need for making sense of situations and comparing options is guidance on how to set priorities. The overarching facilitation principle for managing situational complexity is to keep the group focused on use (see Chapter 3). For every issue that surfaces in evaluation facilitation and negotiations, for every design decision, for every budget allocation, and for every choice among alternatives, keep asking those involved to deliberate on the question, "How will each option likely affect evaluation credibility and use in this situation?"

With that overview of the challenges of facilitating deliberation of evaluation options and making decisions about which to choose for a particular evaluation situation, let's turn to the GUIDE framework for operational guidance.

GUIDE DIMENSIONS OF THE ENGAGE THROUGH OPTIONS PRINCIPLE

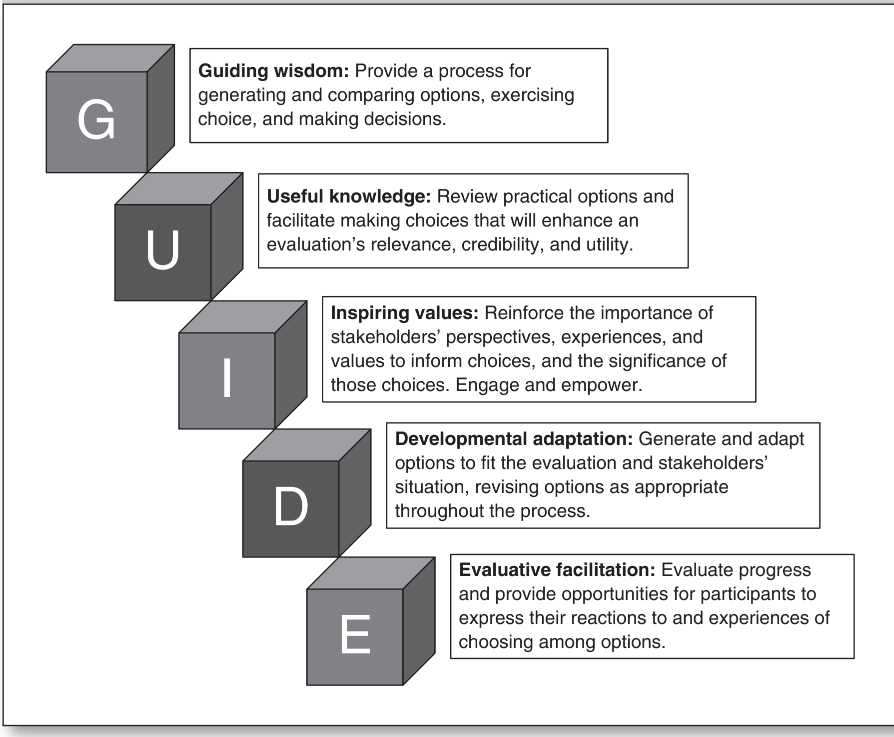
To review, a high-quality effectiveness principle is guiding (G), useful (U), inspirational (I), developmental (D), and evaluable (E). Exhibit 5.2 presents the GUIDE dimensions of the engage through options principle.

Guiding Wisdom

> Provide a process for generating and comparing options, exercising choice, and making decisions.

Evaluation facilitation typically involves helping stakeholders decide an evaluation's primary purposes; prioritize evaluation questions; determine appropriate, credible, and affordable methods; and make these decisions with attention to how and by whom the evaluation processes and findings will be used. This typically involves five steps. In Chapter 1, I introduced the five-step framework to describe

Exhibit 5.2 GUIDE Dimensions of the Engage through Options Principle



my facilitation of the American Evaluation Association (AEA) statement on *What Is Evaluation*? Now I want to apply those five steps to the principle of engaging through options.

Step 1: *Framing*. Ensure that the Group Understands Its Task and the Decision Rules It Will Be Following

The first step in facilitation involves ensuring that the roles, responsibilities, authority, and intended outcomes of the group are clarified and settled. How the group will make decisions must be agreed to by participants. The facilitator's role in guiding the group's work should be made explicit and agreed to by those involved. Exhibit 5.3 summarizes these evaluation facilitation decision-making options in the form of a checklist. It is critical that the evaluation facilitator clarify how the group will function. The answers to the seven questions in the *Checklist for Establishing Facilitation Parameters* determines the evaluation facilitator's role and responsibilities and how the group will engage in its work. The checklist presented in Exhibit 5.3 is more detailed than the questions framework introduced in Chapter 1.

Exhibit 5.3 Checklist for Establishing Facilitation Parameters: Framing and Clarifying Roles, Responsibilities, and Power

1. **Scope of work: What is the group to produce or accomplish?**	a. Full evaluation design from purpose and questions to methods and type of reporting b. Front-end work (questions and methods) c. Back-end work (interpreting findings and recommending action) d. Other scope of work parameters or some combination of a, b, and c, depending on the issue
2. **Stakes: How controversial, visible, or important is the group's work?**	a. High stakes: summative evaluation (rendering overall judgment of merit, worth, effectiveness, and significance); informing or making major budget, strategy, and/or policy decisions; assessing accountability publicly; high-visibility decisions; addressing highly controversial issues b. Medium stakes: advising on evaluation purpose, design, and utility but not rendering judgments or making decisions c. Low stakes: formative, developmental, knowledge-generating evaluation work; low visibility; low to no controversy about the group's work; results not public
3. **What is the group's authority?**	a. Consultative, deliberative role: generate and narrow options, analyze strengths and weaknesses, costs and benefits of alternatives but not make a recommendation b. Make recommendations (to whom?) c. Make decisions (to be acted on by whom?) d. Other authority or some combination of a, b, and c, depending on the issue
4. **Source of options to be considered**	a. Options given to the group by a funder or the person/group seeking input b. Options to be generated by the group c. Options to be generated by the facilitator and deliberated by the group d. Other approach to options or some combination of a, b, and c, depending on the issue

(Continued)

Exhibit 5.3 (Continued)

5. **Time and resource constraints**	a. What is the timeline for the group's work? How compressed or urgent is the timeline? b. What time commitment is expected of participants? c. What resources are needed and available to support the group's work? d. What other factors will affect the group's work?
6. **Group decision rule**	a. Offer multiple and diverse perspectives without agreement b. Majority votes with minority reports c. Majority rules with no minority reports d. Consensus (all must agree to make a decision) e. Some combination of a, b, c, and/or d, depending on the issue and stakes
7. **Who determines group decision rule?**	a. Predetermined by sponsor b. Determined (negotiated) by the group c. Determined by the facilitator

Framing Responsibilities and Clarifying Stakes: An Example. Let me share an example of negotiating the stakes for an evaluation. I facilitated a two-day session with 10 leaders of programs working on obesity. The federal and state funding agencies' program managers also participated. The purpose of the facilitation was to identify potentially common principles that cut across the diverse nonprofit programs being funded. They served different populations with different outcomes, but there was a sense that they were working from common principles. However, adopting and publicly sharing such principles could have political implications, so the funders were nervous about the group agreeing on principles and endorsing them without the federal and state project managers having an opportunity to review the principles with their superiors. The solution I negotiated was to treat the two-day session as exploratory, generative, and building capacity but not as a decision-making work session. Part of the facilitation would involve agreeing on what an endorsement and adoption process could look like, but the two-day session itself would be framed and facilitated as generating possibilities and building the capacity to understand, identify, and evaluate shared principles. This was a breakthrough in the facilitation negotiations with funders and sponsors that allowed the process to move forward.

Step 2: *Generative Engagement*. Create a Process for Generating Evaluation Options

The most basic options in evaluation involve what questions to ask. I find that when I begin facilitating an evaluation process, the people with whom I'm working typically expect me to tell them what the focus of the evaluation will be. They're passively waiting to be told by the evaluation expert—me—what questions the evaluation will answer. But I don't come with specific evaluation questions. I come with a process for determining what questions will be meaningful and what answers will be useful given the evaluation situation, priorities, and decision context. Taking them through the process of formulating questions and determining evaluation priorities is aimed at engendering their commitment to data-based evaluation and use. To illustrate, let me reprise a favorite example illustrating the engaging and empowering nature of facilitating question options (Patton, 2008, pp. 49–51). The context has changed, but the lessons remain relevant, I think.

The Frontier School Division in Manitoba, Canada, encompasses much of northern Manitoba—a geographically immense school district. The Deputy Minister of Education in Manitoba thought evaluation might be a way to shake things up in a district he considered stagnant, so he asked me to facilitate an evaluation process with district officials. The actual form and content of the evaluation was to be determined internally, by them. So I went to Winnipeg and met with key stakeholders: the division administrators, a representative from the parents' group, a representative from the principals' group, and a representative from the teachers' union. I had asked that all constituencies be represented in order to establish credibility with all the people who might be involved in using the evaluation.

In as much as I had been brought in from outside by a superordinate official (the deputy minister), it was not surprising that I encountered reactions ranging from defensiveness to outright hostility. They had not asked for the evaluation, and the whole idea sounded unsavory and threatening. I began by asking them to tell me what kinds of things they were interested in evaluating. The superintendent frowned and responded, "We'd like to see the evaluation instruments you've used in assessing other school districts." I replied that I would be happy to share such instruments if they should prove relevant, but it would be helpful to first determine the evaluation issues and priorities of Frontier School Division. They looked skeptical and, after a lingering silence, the superintendent tried again: "You don't need to show us all the instruments you intend to use. Just show us one so we have an idea of what's going to happen." I again replied that it was too early to talk about instruments. First, we had to identify their evaluation questions and concerns. Then we would talk about instruments.

However, their folded arms and scowling faces told me that what they interpreted as my evasiveness was only intensifying their initial suspicions and fears. I was deepening their resistance by what they perceived as my secretiveness about the content of *my* evaluation scheme. The superintendent tried again: "How about just showing us one part of the evaluation, say, the part that asks teachers about administrative effectiveness?" And he smiled sarcastically.

At that point, I was about to throw in the towel, give them some old instruments, and let them use what they wanted from other evaluations. But first, I made one more attempt to get at their issues. I said, "Look, maybe your questions will be the same as questions elsewhere. Maybe other districts' instrument will work here. Or maybe you don't need an evaluation. I certainly don't have any questions I need answered about your operations and effectiveness. Maybe you don't either. In which case, I'll tell the deputy minister that evaluation isn't the way to go. But before we decide to quit, let me ask you to participate in a simple little exercise. It's an old complete-the-blank exercise from grade school."

I then turned to the blackboard and wrote a sentence in capital letters: I WOULD REALLY LIKE TO KNOW _________ ABOUT FRONTIER SCHOOL DIVISION. I turned back to them and continued, "I want to ask each of you, individually, to complete the blank 10 times. What are 10 things about Frontier School Division that you'd like to know, things you aren't certain about, things that would make a difference in what you do if you had more information? Take a shot at it, without regard to methods, measurement, design, resources, precision—just 10 basic questions, real questions about this division."

After about 10 minutes, I divided them into three groups of four people each and asked them to combine their lists together into a single list of 10 things that each group wanted to know, in effect, to establish each group's priority questions. Then we pulled back together, each group reported its list, and together, we generated a single list of 10 basic evaluation questions—answers to which, they agreed, could make a real difference to the operations of Frontier School Division.

The questions they generated were the kind an experienced evaluator could anticipate being asked in a districtwide educational evaluation because there are only so many things one can ask about a school division. But the questions were phrased in their terms, incorporating important local nuances of meaning and circumstance. Most. important, they had discovered that they had questions they cared about—not my questions but their questions, because during the course of the exercise, it had become their evaluation. The whole atmosphere had changed. This became most evident as I read aloud the final list of 10 items they had generated that morning. One question read, "How do teachers view the effectiveness of administrators and how often do they think administrators ought to come into classrooms?"

One of the administrators who had been most hostile at the outset said, "That would be dynamite information. We have no idea at all what teachers think about us and what we do. I have no idea if they want me in their classrooms or not, or how often they think I ought to visit. That could turn my job around. That would be great to know."

Another question concerned the relationship between the classroom and the community. Both the teacher and parent representatives said that nobody had ever thought about that in any real way: "We don't have any policy about that. We don't know what goes on in the different schools. That would be really helpful for us to know."

We spent the rest of the day refining questions, prioritizing, formalizing evaluation procedures, and establishing an agenda for the evaluation process. The hostility had vanished. By the end of the day, they were anxious to have me make a commitment to return. They had become excited about doing *their* evaluation. The evaluation had become meaningful, important, and credible because the questions were their questions.

A month later, they found out that budget shifts in the Ministry meant that the central government would not pay for the evaluation. The deputy minister told them that they could scrap the evaluation if they wanted to, but they decided to pay for it out of local division funds. The evaluation was completed in close cooperation with the task force at every step along the way. The results were disseminated to all principals, teachers, and parent leaders. The conclusions and recommendations formed the basis for staff development conferences and division policy revisions.

The evaluation process itself had an impact on the division. Over the next few years, Frontier School Division went through many changes. It became a very different place in direction, morale, and activity than it was on my first visit. Not all those changes were touched on in the evaluation nor were they all a direct consequence of the evaluation. But generating a list of real and meaningful evaluation questions played a critical part in getting things started.

So, let's review where we are in examining the principle of engaging through options. Step 1 under the guidance criterion is ensuring that the group understands its task and the decision rules it will be following. Step 2 is creating a process for generating options. We turn now to the third step: facilitating a systematic comparison of options. This step is informed by the utility criterion in the GUIDE framework for effectiveness principles.

Step 3: *Comparative Analysis*. Facilitate Systematic Comparison of Options

In setting the stage for comparing options and deciding priorities, I like to remind those I'm working with that no evaluation can look at everything. Priorities must be determined, questions focused. The larger context, I remind them, is that all of life involves prioritizing and focusing; for example, deciding how we spend time and money, what work tasks get done, what leisure activities to engage in, and what we eat. Despite evidence of a widespread attention deficient disorder in modern society, or perhaps because of it, knowing how to focus is a key to success. So, too, evaluations questions must be focused and methodological choices must be made.

This problem of focus is by no means unique to program evaluation. Management consultants find that a major problem for executives is focusing their energies on priorities. The trick in certain forms of meditation is learning to focus on a single mantra, koan, or image. Professors have trouble getting graduate students to analyze less than the whole of human experience in their dissertations. Time management specialists find that people have trouble setting and sticking with priorities in both their work and personal lives.

Depending on the time available and my assessment of the group's readiness to begin making comparisons among options, I may ask each participant to share a personal example of having considered alternative courses of action in making a decision about what to do. Here are examples shared by participants in one group:

- choosing between an Android versus an Apple smartphone
- deciding the menu for a special occasion
- comparing vacation possibilities, both place and activities
- picking a childcare center

Before moving on to Step 4 in the general facilitation framework, we will turn first to the second GUIDE criteria.

Useful Knowledge

Facilitate Choices among Options to Enhance Utility

In utilization-focused evaluation, focus flows from intended use by intended users, our overarching evaluation facilitation principle (Chapter 3). With inevitably limited time and resources, decisions must be made about what's worth finding out in an evaluation. Choosing to look at one issue in depth is also a decision not to look at something else in depth. The evaluation facilitator frames the work of the stakeholder group and lays out a process for reviewing options and making decisions. Making intended evaluation use the primary criterion for choosing among options enhances the evaluation's relevance. To build capacity for comparing options and using criteria to inform decisions about which option to choose, I often begin by posing two different evaluation questions and having participants compare and contrast their implications for use.

Exhibit 5.4 presents the alternative questions and provides a framework for facilitating group discussion of their implications.

Actionable Evaluation Questions

Using relevance, credibility, and potential utility as decision criteria for choosing among options leads to asking questions that will yield actionable answers. There's a deceptively simple logic involved in asking evaluation questions. It goes like this: Ask questions. Get answers. Act based on those answers.

Questions ⟶ Answers ⟶ Action

Exhibit 5.4 Alternative Evaluation Questions

Does the program work?
versus
What works for whom in what ways under what conditions with what results?

Five Facilitation Questions

1. What different evaluation purposes are represented by these two questions?
2. How would the answers be different?
3. What intended users would likely prefer each of these questions? Why?
4. Compare the likely uses of findings and answers to the two questions.
5. What assumptions are embedded in each question?

Would that it was that simple. When we inquire more deeply into these connections, as I do in framing an evaluation decision-making process for a stakeholder group, I start by generating criteria for what constitutes good questions, good answers, and appropriate actions. I find that a concrete example grounds the discussion. Evaluations of drug abuse resistance education offers a powerful cautionary tale that can engage and illuminate stakeholders regardless of their background and expertise. So, I use DARE as an example.

DARE stands for Drug Abuse Resistance Education, a 17-week drug education program targeted at middle school children in which police officers go into schools to teach the dangers of drug use and alcohol abuse. Launched in 1983 in Los Angeles, DARE spread rapidly, eventually being taught to fifth- or sixth-graders in 80% of the school districts across the United States and 54 other countries around the world, reaching an estimated 36 million students annually (Eyle, 2002; Hanson, 2007).

I start with a summative evaluation question: Will middle school students who participate in DARE eschew drug use in 20 years? Or even 10 years? I ask the group, "What's your reaction to the utility of that question?"

Quickly, the discussion concludes that that question, if it could be answered at all, certainly can't be answered within an actionable time frame. Decisions about whether to continue funding DARE must be made within a year or two, certainly well before a 10-year follow-up study has been completed. So what questions can be answered? What are the attitudes toward drug use of students who complete DARE? Do students who complete DARE in 7th and 8th grades use drugs less in high school than students without DARE experience?

In posing these questions, which have a more actionable time frame, I ask the group what findings might emerge and what actions would reasonably follow from those findings. What actions would follow from findings that DARE is effective? Ineffective? Mixed results?

After brief discussion, I share the history, scope, and actual results of such evaluations. Thousands of evaluations of DARE have been conducted in local school districts as well as several national evaluations (Government Accountability Office, 2003). These evaluations have consistently shown that the program "was not effective in actually keeping young people from using drugs" (Weiss, Murphy-Graham, & Birkeland, 2005, p. 15). However, the program did lead students to have more positive views of the police. But knowledge and attitude changes about drugs were not sustained and DARE did not lead to lower drug use (Eyle, 2002; Hanson, 2007).

I know of no program that has been so thoroughly evaluated with such consistently negative results—and yet remains widely popular. Distinguished evaluation pioneer and knowledge use scholar Carol Weiss of Harvard University has examined DARE in depth as an example of "the neglect of evaluation" and "an elegant case of nonutilization" (Weiss et al., 2005, p. 15). Indeed, DARE is still going strong, though with a revised curriculum and more comprehensive approach. Evaluation findings remain negative and DARE remains popular.

I then invite reactions to the DARE story. The discussion is typically quite intense. My purpose is not to attack DARE but to get the group thinking about the challenges and complexities of evaluation use. I use the DARE discussion to pave the way for a more general question: "What factors," I ask, "affect use, nonuse, and misuse in *your* experience?" This exercise is a warm-up. It takes about 15 minutes. Then we're ready to deal with the evaluation issues the group has been convened to take on.

Answerable Questions

Questions that can yield actionable answers must be answerable questions. This seems straightforward but, again, it's a distinction that may require some facilitation to appreciate. Many important questions are primarily philosophical or religious in nature; that is, matters of belief and opinion can't be answered with data. Consider evaluating a high school sex education program. I was facilitating an advisory task force for such an evaluation and one stakeholder participant said right away, "I have some questions I'd like us to ask. Is sex education moral? Is teenage sex moral?"

I responded, "What kind of data could answer those questions?" The group discussion concluded that those weren't empirical questions. Data couldn't answer those questions. Now, it is true that data could be gathered through questionnaires or interviews about what parents, teachers, students, and community members *believe* about the morality of sex education, but that's different from deciding a moral question. Whether sex education is moral can be discussed and debated, but it can't be answered empirically. When moral issues arise, the evaluation facilitator may have to diplomatically help people learn to distinguish between answerable and unanswerable questions.

Thus, part of the evaluation facilitation challenge is clarifying questions to make them clear, answerable, and actionable. I was working with a task force evaluating an employment training program. One person proposed the question, "Do dropouts stop coming to the program?"

I was confused. The definition of a *dropout* is someone who stops coming to the program. So I probed. "Say more. What are you trying to find out with that question?" It turned out that the issue was voluntary versus forced dropouts. How many dropped out by choice versus getting kicked out? The program reported the dropout rate quarterly but didn't make this distinction. What constituted a dropout and why participants were dropping out were important and answerable questions, especially because the program had a high dropout rate.

From a utilization-focused evaluation perspective, good questions are those that the primary intended users want answered and would use; that is, they can specify the relevance of answers for future action. Some questions are interesting but not particularly actionable. In an evaluation of an adult literacy program, a question arose about what proportion of participants were men. It's almost automatic to ask gender as a background question. But why in this case? How would gender data be used? Was the program supposed to be gender-balanced? Were there targets for a certain number of men versus women? Was there reason to believe that men and women had different success rates or dropout rates? And suppose that turned out to be true. What would the program do about gender differences? What would make exploring gender differences a priority question? Discussing the potential utility of gender data opened concerns about how males and females were recruited and treated that were both important and actionable. Finding that out meant probing into the purpose of collecting and analyzing data by gender.

The point is to get beyond typical questions that people are used to seeing on questionnaires, such as gender and age, and ask, what would we do with that information? What's the actionable issue? Why is gender or age important?

You see that what drives the question formulation process is ongoing attention to utility. As an evaluation facilitator, I establish a pattern of probing into potential utility and making potential use of findings a priority.

Matching Questions to the Evaluation's Purpose

The process of working with primary intended users to identify priority questions varies depending on the evaluation's overall purpose, any decisions pending about the program's future, and a myriad of situational factors. Exhibit 5.8 in the Tools and Resources section at the end of this chapter presents questions that flow from and are matched to different evaluation purposes.

We're ready now for Facilitation Step 4. Step 1 was ensuring that the group understood its assignment and decision rules. Step 2 involved generating options. These options were sorted, compared, and contrasted in Step 3. Step 4 moves the group to choosing among options.

Step 4. Guide the Group in Converging on Its Preferred Option and Make Design Decisions

Convergence involves choosing from among many options which ones to act on. The facilitator guides the group in deciding. In beginning to facilitate convergence decision making, I like to introduce the Latin root for the word *decide*. The root *cide* is shared with suicide and homicide, which means to cut off or kill. *Homicide* refers to killing another person, *suicide* to killing one's self. To *decide* is to kill options.

A related metaphor that nicely makes the point involves contrasting the medieval Latin words *incidere* and *decidere. Incidere* is related to incisors (teeth) and chewing. *Decidere* connotes "stop chewing." The colloquial version is "Crap or get off the pot." *Inscisours* was the Latin name of a cutting tool. To decide is to cut until you have produced what you want. Facilitation toward convergence (eliminating and prioritizing options) is the cutting tool that takes a group to a point of strategic focus. Evaluation facilitation cutting tools include rankings and ratings.

Rankings and Ratings

Facilitating a compare-and-contrast discussion (Step 3) often eliminates most options. In some cases, clear preferences emerge and the choices become obvious. Step 3 flows naturally and seamlessly into Step 4. But the evaluation facilitator must be prepared to guide a group through a rating and ranking process when several competing options remain and the group is divided on which to choose. If further discussion isn't leading to a consensus choice, formally rating options gives everyone a voice, an equal say in expressing preferences. If the choice is among 3 to 5 options, having each participant rank order the choices can reveal how close the group is to consensus.

Data Collection Options

Option A. Survey key informants

Option B. Telephone interviews with key informants

Option C. Face-to-face interviews with key informants

Option D. Focus group with key informants

Option E. No key informant data collection

After discussion of the pros and cons of each choice, ask participants to individually rank order their preferences.

1. Your top data preference: A B C D E other ________
2. Your 2nd data preference: A B C D E other ________
3. Your 3rd data preference: A B C D E other ________
4. Your 4th data preference: A B C D E other ________
5. Your 5th data preference: A B C D E other ________

People can have difficulty ranking more than three things; five is the most that can reasonably be ranked. With five or more choices, ratings supersede rankings as a systematic way of expressing preferences. I facilitated a group of eight philanthropic foundation program officers working on a questionnaire to send to grantees about their evaluation experiences, concerns, and preferences. The group generated 35 possible items that we needed to reduce to 15. The participants rated each item on a four-point scale:

1. Very important
2. Fairly important
3. Somewhat important
4. Not a priority item

The results made it easy to pick the 15 highest rated items.

With a different philanthropic group, the issue was agreeing on a set of priority criteria to apply in selecting among grant proposals. The group generated 16 possible criteria. After discussion, they agreed that half that many would be manageable. The rating scale I generated reflected the tone and language of the discussion:

1. This criterion resonates; I'd support this
2. Needs some wordsmithing but is in the right direction
3. Doesn't resonate
4. I oppose this criterion

Six items emerged as consensus criteria, so those became the ones adopted.

Evaluation Options Inventory. The engage through options principle applies to all aspects of evaluation facilitation. The preceding examples have highlighted options for comparing alternative evaluation questions. As a group prioritizes questions, the next options to be considered are methodological. Consider the following suggestive inventory of fifteen methods, design, and reporting options that an evaluation facilitator may pose at some point for group deliberation and decision making.

1. *Type of data to collect*. Mail questionnaires, telephone interviews, or personal face-to-face interviews?
2. *Interview options*. Individual interviews or focus groups?
3. *Survey design options*. Even-numbered or odd-numbered scales on survey items?
4. *Survey questions to ask*. Opinion, knowledge, and/or behavioral questions? All closed questions or are some open-ended? If some are open-ended, how many?
5. *Test options*. Norm-referenced or criterion-referenced tests?
6. *Instrumentation choices*. Develop new customized instruments or adopt instruments already available?
7. *Design choices*. Experimental design, quasi-experimental design, or case studies? Fixed or emergent design?
8. *Observation options*. Participant observation or spectator observation? A few in-depth observations or many shorter observations? Single or multiple observers? Standardized or individualized protocols?
9. *Follow-up data collection*. Follow up after two weeks, three months, six months, or a year? Follow up everyone or a sample?
10. *Sampling options*. What kind of sample: simple random, stratified, and/or purposeful? What size sample?
11. *Comparisons*. What comparisons to make: Past performance? Intended goals? Hoped-for goals? Other programs?
12. *Conceptual options*. Logic model or theory of change? Or both? Or neither?

13. *Focus options*. Measure performance on stated goals or also search for unintended consequences?
14. *Reporting options*. Write a comprehensive report? Disseminate only an executive summary? Oral briefings? Disseminate findings to whom, in what forms, and for what purposes?
15. *Use of findings*. Will recommendations be generated? What follow-up plan should be made to ensure use of findings and recommendations?

These fifteen evaluation issues are far from exhaustive of the decisions entailed in designing an evaluation. Nor are they all dealt with at once or even in every evaluation facilitation. I have listed them as examples of what an evaluation facilitator may need to help a group deliberate and decide. To facilitate methodological options can involve not only posing the alternatives and asking for group discussion but may also involve building the group's capacity to make methods decisions based on utility considerations. Let me share with you my favorite way of introducing methods options to a group. I like to have them engage with the odd–even question in survey design.

The Odd–Even Question in Survey Design. Should response scales be even-numbered (e.g., four or six response choices) or odd-numbered (e.g., three or five choices)? It doesn't seem like such a big deal, actually, but I've seen evaluators on both sides of the question go at each other with the vehemence of Marxists versus capitalists, osteopaths versus chiropractors, or cat lovers versus dog lovers. What's all the ruckus about? It's about the value and validity of a midpoint on questionnaire items. If I'm going to be facilitating methodological and measurement decisions with a group, I like to warm up by asking, "Should a survey give people a midpoint?"

An even-numbered scale has no midpoint:

Should the workshop be expanded from one day to two days?			
Strongly Agree	Agree	Disagree	Strongly Disagree

An odd-numbered scale has a midpoint.

Should the workshop be expanded from one day to two days?				
Strongly Agree	Agree	No Opinion	Disagree	Strongly Disagree

Even-numbered scales force respondents to lean in one direction or the other (although a few will circle the two middle responses, creating their own midpoint if not provided with one on the survey). Even-numbered scales allow the respondent to hedge, to be undecided, or, in less kind terms, to cop out of deciding one way or the other, or yet again, to be genuinely in the middle.

One thing about surveys is clear: If given a midpoint, many respondents will use it. If not given a midpoint, most respondents will answer leaning one way or the other (though some ornery respondents will create their own midpoint).

Which one is best? Having carefully considered the arguments on both sides of the issue, having analyzed large number of questionnaires with both kinds of items, and having meditated on the problem at great length, I find that I'm forced to come down firmly and unwaveringly right smack in the middle. *It depends*. Sometimes odd-numbered scales are best and sometimes even-numbered scales are best. How to decide?

The issue is not technical, statistical, or methodological. The issue is one of utility. What do intended users want to find out? Will the findings be more useful if respondents are forced to lean in one direction or the other? Or is it more useful to find out how many people are undecided or don't know. As an evaluation facilitator, I guide the group through this deliberative process to consider the value and implications of offering a midpoint. Do they believe that down deep inside, everyone really leans one way or the other on an issue, or do they believe that some people are genuinely in the middle on the issue and they want to know how many have no opinion?

Not only can non-researchers make this choice, but they often enjoy doing so; engaging them in thinking about such alternatives and their implications is empowering. To facilitate consideration of methodological options can involve more than simply posing the alternatives, facilitating group discussion, and building the group's capacity to make methods decisions based on utility considerations. The evaluation facilitator may also be called on to inspire the group to take on methods decisions and enhance their confidence to do so. Let's turn, then, to the inspiring criterion in the GUIDE framework for principles.

Inspiring Values

Reinforce the importance of stakeholders' perspectives, experience, and values to inform choices, and the significance of those choices. I have found that, given the opportunity, non-researchers appreciate the opportunity to consider methods options. My consistent message to those I am facilitating is that their expertise, experience, and insights are essential to ensure an evaluation's relevance, credibility, and utility. That includes bringing their expertise, experience, and insights to bear in selecting methods. Not everyone sees it this way.

For the most part, evaluation professionals have come to accept that use can be enhanced by actively involving stakeholders in decisions about the evaluation's purpose, scope, and focus to ensure relevance and buy-in. In other words, they can accept playing a consultative and collaborative role during the conceptual phase of the evaluation. Where my colleagues often part company with me is in the role to be played by non-researcher stakeholders and intended users in making measurement and design decisions. "The evaluator is nothing," they argue, "if not an expert in methods and statistics. Clearly, social scientists ought to be left with full responsibility for operationalizing program goals and determining data collection procedures." In evaluation facilitation, I take a different path based on the principle of engaging through options. That principle applies to methods decisions as well as all other aspects of evaluation.

Beyond Technical Expertise

The common perception of methods decisions among non-researchers is that such decisions are primarily technical in nature. Sample size, for example, is determined by a mathematical formula. The evaluation methodologist enters the values of certain variables, makes calculations, and out pops the right sample size to achieve the desired level of statistical robustness, significance, power, validity, reliability, generalizability, and so on—all technical terms that dazzle, impress, and intimidate practitioners and non-researchers. Evaluation researchers have a vested interest in maintaining this technical image of scientific expertise, for it gives us prestige, inspires respect, and, not incidentally, it leads non-researchers to defer to us, essentially giving us the power to make crucial methods decisions and then interpret the meaning of the resulting data. It is not in our interest, from the perspective of maintaining prestige and power, to reveal to intended users that methods decisions are far from purely technical. But, contrary to public perception, evaluators know that methods decisions are never purely technical. *Never.* Ways of measuring complex phenomena involve simplifications that are inherently somewhat arbitrary, always constrained by limited resources and time, inevitably involve competing and conflicting priorities, and rest on a foundation of values preferences that are typically resolved by pragmatic considerations, disciplinary biases, and measurement traditions.

The reason to debunk the myth that methods and measurement decisions are primarily technical is to enhance use. We know that use is enhanced when practitioners, decision makers, and other users fully understand the strengths and weaknesses of evaluation data and that such understanding is increased by being involved in making methods decisions. We know that use is enhanced when intended users participate in making sure that, when trade-offs are considered, as they inevitably are because of limited resources and time, the path chosen is informed by relevance. We know that use is enhanced when users buy into the design and find it credible and valid within the scope of its intended purposes as determined by them. And we know that when evaluation findings are presented, the substance is less likely to be undercut by debates about methods if users have been involved in those debates prior to data collection.

As in all other aspects of the evaluation, the utilization-focused evaluation facilitator advises stakeholders about options; points out the consequences of various choices; offers creative possibilities; engages with users actively, reactively, interactively, and adaptively to consider alternatives; and facilitates their involvement in methods decisions through a process of negotiation. At the stage of choosing methods, the evaluation facilitator engages as a technical adviser, consultant, teacher, and advocate for quality. The evaluation facilitator has a clear stake in the quality of the evaluation because, if nothing else, the evaluator's reputation is at stake in every evaluation. And the primary intended users have a stake in ensuring a credible evaluation that provides useful information. The evaluation facilitator guides the process of making design, methods, and measurement decisions.

Competence to Compare Methods

Christie (2007) found that stakeholders could distinguish among the merits and uses of different kinds of designs. Using a set of scenarios derived from actual evaluation studies, she conducted a simulation to examine what decision makers reported as evaluation design preferences and likely influences. Each scenario described a setting in which results from one of three types of evaluation designs would be available: large-scale study data, case study data, or anecdotal accounts. The simulation then specified a particular decision that needed to be made. Decision makers were asked to indicate which type of design would influence their decision making. Results from 131 participants indicated that participants were influenced by all types of information, yet large-scale and case study data were more influential relative to anecdotal accounts, certain types of evaluation data were more influential among certain groups of decision makers, and choosing to use one type of evaluation data over the other two depended on the independent influence of other types of evaluation data on the decision maker as well as prior beliefs about program efficacy. These decision makers had varying design preferences and were quite capable of distinguishing the credibility and utility of various types of evaluation studies—or measurement options.

Courtesy Fresh Spectrum

Examples of Methodological Issues That Inspire Thoughtful Deliberation and Decision

To lay the scaffolding for methods decisions, I like to ask, "What are some statistics that commonly determine if a program is successful?" I write their answers on a flipchart:

- ✓ Test scores, percentage that pass
- ✓ High school graduation rate
- ✓ Percentage of program graduates who get and keep jobs
- ✓ Percentage of babies born healthy
- ✓ Reduced air pollution

In three minutes, the group can usually generate 10 or more success indicators. Then I invite them to step inside the numbers with me. Consider a program's dropout rate, the opposite of its completion rate. In most programs, the dropout rate is an important indicator of how participants are reacting to a program. But when has someone dropped out? After they miss three sessions? Five sessions? Ten sessions? Who decides? Dropout rates typically involve some arbitrary cutoff. For example, school districts vary widely in how they define, count, and report high school dropouts, as do chemical dependency, adult literacy, parent education, and all kinds of other programs. These variations exist because there is no universal right answer for when someone has dropped out of a program.

It seems simple. Participants have dropped out when they stop coming. So, a parent misses three consecutive parent education sessions. The program decides that missing three consecutive sessions means they've dropped out. There are only 30 sessions, so missing three means missing 10% of the program. Then the parent returns for the next session after missing three in a row. Does the program allow reentry? Did the parent drop out and reenter, or is this treated as continuation? Does missing three sessions trigger a staff intervention to encourage potential dropouts to reengage? Or if someone stops coming, are they left alone? What is the data system for monitoring attendance? Are potential dropouts noted in real time (when they've missed the third session) or only tabulated at the end of the program? These are just a few of things that have to be determined and understood to calculate and interpret a dropout rate. What is a high rate? A low rate?

No less vague and difficult are concepts such as *in the program* and *finished the program.* Many programs lack clear beginning and ending points. For example, a job training program aimed at chronically unemployed minority men has a month-long assessment process, including testing for drug use and observing a potential participant's persistence in staying with the process. During this time, the participant, with staff support and coaching, develops a plan. The participant is on probation until he completes enough of the program to show seriousness and commitment, but the program is highly individualized so different people are involved in the early assessment and probation processes over very different time periods. There is no clear criterion for when a person has begun probation or completed probation and officially entered the program. These processes, in aggregate, will determine key dropout, completion, and acceptance rates. Making sure that such categories are meaningful and valid, so that the numbers are credible and useful, involves far more than statistics. Careful thought must be given to how the numbers and reported rates will be calculated and used, including whether they can be used for comparisons with similar programs.

Dosage is another tricky issue. In a 30-session Early Childhood Parent Education program, how many sessions must a parent attend to have been exposed to enough content to expect the program to show impact? Few, if any, parents get to all sessions. Is attending 20 sessions enough to expect positive outcomes? Non-researchers can grasp the significance of these measurement issues. Rather than breeding skepticism about how statistics are gathered and reported, I find that non-researchers take seriously the challenge of determining success criteria and sufficient dosage levels. They can understand the importance of understanding how target dosage rates are established and the consequences of whatever targets are chosen to indicate success.

Facilitating Major Decisions

An evaluation can involve many methods, measurement, and design decisions. How many such decisions to bring before a group, and which ones they should deliberate, involves judgment about which options will be appropriately decided by the group. I find the 80/20 rule instructive in this regard. The 80/20 rule (also known as the Pareto principle or the law of the vital few) states that for most interventions, roughly 80% of the outcomes stem from 20% of the effort. Correspondingly, 80% of what matters flows from 20% of the decisions made. Well-documented examples of the 80/20 rule include 20% of employees generate 80% of a company's output; 20% of customers are responsible for 80% of a business's revenues; 80% of the clothes you wear are from 20% of what you own; you spend 80% of your time in 20% of your house; you make 80% of your phone calls to 20% of the people in your phone list; you spend most of your money on a few major items; and so forth (Koch, 1999). The 80/20 ratio is not precise; it's a framework of proportionality rather than a precise calculus. But its implications are powerful in many arenas of prioritizing, including facilitation. It is the job of the evaluation facilitator to engage the group in the most influential, understandable, and utility-enhancing trade-offs and methods decisions. Examples include interviews versus questionnaires, a mix of quantitative and qualitative methods, target levels in key indicators that constitute success or failure, length of time for follow-up to determine sustained impact on key outcomes, and what background variables to include in analyzing and comparing differential outcomes among subgroups (e.g., gender, race, age).

And, of course, as discussed earlier, the huge decision of whether survey opinion items should include a midpoint!

Developmental Adaptation

In developmental adaptation, you should generate and adapt options to fit the evaluation and stakeholders' situation, revising options as appropriate throughout the process. I want to begin consideration of the developmental criterion with an example of a fluid evaluation facilitation situation. A program director commissioned me to facilitate staff involvement in designing a new evaluation system. At our first session, the director, who was much admired and trusted, announced that she was retiring after 15 years of service. The new director would be hired after a national search, although I learned that there were two internal candidates for the position. The evaluation process I was to facilitate was rescheduled to begin after the new director had been on the job for a month. Much would be in flux. How much I couldn't know in advance. I'd have to be flexible and responsive.

The engage through options principle means that the options to be facilitated may change and evolve as the situation changes and evolves. Let's look at a real-time facilitation adaptation that involved changing the framing of options.

Cross-Cultural Facilitation Example: The Emergence of Two-Eyed Seeing

A philanthropic foundation asked me to work with a group of Native American leaders representing tribes around the U.S. The purpose was to identify strategic priorities that the foundation might support. These would be shared strategic priorities across tribes. I opened with the framework of *two-eyed seeing,* which recognizes that indigenous perspectives and Western cultural perspectives are different and that indigenous people often work, live, walk, and engage in both worlds, often simultaneously. Seeing through two eyes, indigenous and Western, is the guiding principle brought into a colearning journey by *Mi'kmaw* Elder Albert Marshall. *Etuaptmumk* is the *Mi'kmaw* word for *two-eyed seeing*. I had learned about the *two-eyed seeing* framework when facilitating a rights-based evaluation workshop with Elder Marshall during the 2016 Canadian Evaluation Society annual meeting in Newfoundland.

> We often explain *Etuaptmumk*—Two-Eyed Seeing by saying it refers to learning to see from one eye with the strengths of

> Indigenous knowledges and ways of knowing, and from the other eye with the strengths of Western knowledges and ways of knowing . . . and learning to use both these eyes together, for the benefit of all. (Marshall, 2016, p. 1)

For the opening exercise, I asked each of the 20 Native American participants to share a recent experience that portrayed their reality of two-eyed seeing. One participant shared an example of working on a community college set of courses for Native American students that was based on Native American perspectives, ways of interacting, values, and knowledge while meeting the traditional community college curriculum requirements and expectations about teaching. Another shared an example of the conflict between her tribe's deliberation process in governance and the required procedures imposed by a federal government agency from whom they receive funding.

We then reviewed what it means to think strategically from a tribal perspective, given that *strategy*, the word itself, is a military term and that the genocide of American Indians was a result of military strategy. We then discussed what kind of process the group wanted to use to respond to the foundation's request for strategic priorities. We agreed to abandon the word *strategic* and the word *priorities* because the group did not feel that they could represent their own tribes, much less the entire national Native American community. It was a very rich and honest discussion about the burdens and misconceptions that often arise for them in being asked to be a part of some process or group in which their role was to represent the Native American perspective, as if such a perspective exists as a single, monolithic way of seeing and knowing.

The language that emerged as comfortable was for focusing the group's work was *identifying areas of need and potential action*. Once identified, the group insisted that these areas of need would not be prioritized nor would everyone have to agree with all of them and that how many in the group saw a particular need as a concern not be reported. All perspectives would be represented equally without a vote or counting how many supported any specific need. Indeed, the analytical framework that resonated for the group was *two-eyed seeing*, which I had used to frame introductions at the start of the session. I suggested that we try two-eyed seeing as a template for identifying and discussing areas of need for potential action. Exhibit 5.5 shows the two-eyed seeing framework that emerged.

The purpose of this facilitation was not explicitly evaluative. But evaluative thinking undergirds my facilitation in asking questions about

Exhibit 5.5 Two-Eyed Seeing Framework for Identifying Native American Needs and Opportunities

Issue	Native American Perspectives on the Issue	Dominant White Culture Perspectives on the Issue
1. Name (label or categorize) the perceived need		
2. Describe the need or opportunity: history, scope, nature, characteristics, range, and so on.		
3. What kind of evidence is available about the need or opportunity?		
4. What makes it significant?		
5. Potential actions, responses, interventions, collaborations		
6. Comments, observations, caveats, cautions, concerns		
7. Who is already engaged on this issue?		

evidence, definitions, criteria, and multiple perspectives. My client was the philanthropic foundation that hired me. The program officer who commissioned the process and selected the participants was uncertain how to interpret the result and asked if I could convert the results into strategic priorities. The foundation's board had adopted a strategic plan that included a section devoted to identifying and funding Native American strategic priorities. I suggested that the report to the board use the two-eyed seeing language that emerged from the group and include reference to the foundation's statement of values, which include the following:

- Work with people where they are, not where we want them to be.
- Listen and respect the people we work with.
- Support dialogue.
- Value diversity.

I explained that I was guided by those values when facilitating the group and I felt confident that the board would appreciate my adherence to their values rather than to the narrow language in their strategic plan. I understood my role and responsibility to be facilitating expression of the participants' perspectives in capturing, synthesizing, and reporting the knowledge of the participants. The end result was well-received. Offering the group an alternative option and adapting to what emerged proved engaging, meaningful, credible, and useful.

Evaluable Facilitation

Evaluable facilitation involves evaluating progress and providing opportunities for participants to express their reactions to and experiences of choosing among options. Real-time adaptation when facilitating is a form of improvisation. Principle 4 on observation will examine in depth how to read groups to get immediate, in-the-moment, as-it-happens feedback. Feedback about the engage through options principle includes both formal and informal feedback.

Evaluation facilitation inherently involves evaluation. In the five-step facilitation process we've been following, two forms of evaluation must be attended to by an evaluation facilitator: (1) evaluating the group's process and readiness to transition from one step to the next step (or the next item on the agenda) and (2) providing opportunities for those involved to express their reactions to the options considered, the decisions made, and the processes provided for generating, comparing, and choosing among options. Embedded and inherent in each of the five steps for facilitating options are progress questions and criteria. The time to ask those questions and apply those criteria is in the transitions between steps. As a way of bringing closure to each step, the evaluation facilitator should do five things. Exhibit 5.6 lists those things.

An Extended Stepwise Example

Earlier in this chapter, I described facilitating a two-day session with 10 leaders of programs working on obesity. The federal and state funding agencies' program managers also participated. The purpose of the facilitation was to identify potentially common principles that cut across the diverse nonprofit programs being funded. They served different populations with different outcomes, but there was a sense that they were working from common core principles. I have described how the evaluation facilitation was framed but not what subsequently happened. I want to take you through that process now to illustrate the five facilitator responsibilities for guiding a group through the five steps for

Exhibit 5.6 Facilitating Transitions between Steps: Five Responsibilities

To help bring closure in moving from one step to the next:

1. Determine that it is time to move to the next step (or move onto the next item on the agenda) because the necessary work at that stage has gone as far as needed or possible, according to the facilitator's evaluation of progress and possibility.
2. Announce that the group has reached a transition point, having accomplished what it needed to accomplish for a particular step or agenda item.
3. Summarize what has been accomplished, including anything that has been set aside to deal with later or set aside permanently.
4. Check in with the group about both what has been accomplished and the process used for accomplishing it; provide an opportunity for reflection and evaluation.
5. Present the next step, making a smooth transition from what has just been accomplished to what needs to be accomplished next.

facilitating options. In so doing, I'll place particular emphasis on the evaluable criterion: providing opportunities for participants to express their reactions to and experiences of choosing among options.

Step 1. Framing: Ensure That the Group Understands Its Task and the Decision Rules It Will Be Following. Because adopting and publicly sharing such principles could have political implications, the funders were nervous about the group agreeing on and endorsing principles without the federal and state project managers having opportunity to review them with their superiors. The solution was to treat the two-day session as exploratory and capacity building but not as a decision-making assignment.

Facilitator's Responsibilities for the Transition from Step 1 to Step 2

Facilitating the Transition from Step 1 to step 2	**Evaluation Facilitator's Judgment and Guidance to the Group**
1. Determine that it is time to move to the next step, or the next item on the agenda (the evaluation facilitator's evaluation of progress).	I judged that the introductions made by the organizers, and my own framing of what we would be doing, were sufficiently clear to move forward.

(Continued)

(Continued)

Facilitating the Transition from Step 1 to step 2	**Evaluation Facilitator's Judgment and Guidance to the Group**
2. Announce that the group has reached a transition point, having accomplished what it needed to accomplish for a particular step or part of an agenda.	"Every process begins with framing what will happen, what the group expects to accomplish, and how the group will make decisions. Having done that, we are ready to get to work. But before proceeding, let's make sure that we're all on the same page."
3. Summarize what has been accomplished, including anything that has been set aside to deal with later or set aside permanently.	"We've completed introductions, agreed on how we will work together, and decided that the purpose of our work is to generate a draft list of overarching principles that provide coherence across your diverse programs. We will not attempt to finalize the principles or officially adopt them. This is an exploratory process to learn about principles, learn about each other, and provide an opportunity for cross-program collaboration and interaction."
4. Check in with the group about both what has been accomplished and the process for accomplishing it; provide an opportunity for reflection and evaluation.	"How does that sound (referring to summary above)? Any clarifying questions? Anything else we need to take up before digging into the work?" There followed questions on dinner arrangements, the ending time the next day, and how the two days should be officially referred to in required reports to the funders. It was important to get these logistical concerns clarified so the people could concentrate on the work at hand.
5. Present the next step, making a smooth transition from what has just been accomplished to what needs to be accomplished next.	"Okay, we're going to start working on principles. We'll start by reviewing examples of practitioner principles for other kinds of programs, discuss them, and then start generating possible principles for your work."

These five responsibilities of the evaluation facilitator should be attended to in the transition from each step to the next. I won't provide a detailed script illustrating each of the five responsibilities, but I will illustrate the evaluation check-in (Responsibility 4) for each transition.

Using the principles facilitation for obesity programs as the example, I'll offer a sample of what each check-in yielded that helped guide my ongoing facilitation.

Evaluation Check-Ins for Each Step-to-Step Transition: Illustrations

Transitions	Evaluation Check-In	Examples of Responses in the Obesity Principles Facilitation Example
Step 2 (generating options) to Step 3 (comparing options): content check-in	"We've generated a number of options. Before we begin reviewing and comparing options, is there anything you want to clarify, add, or otherwise comment on?"	Four small groups of four participants had meant to generate possible principles. Two reported that they each still had a topic they still wanted to cover but had not had time to articulate a principle for that topic. I wrote the topics on a "parking lot" flipchart page and promised to give all groups time to add principles after we got some experience reviewing and comparing those already generated.
Step 2 to Step 3: process check-in	"What was the process of generating options like for you? What are your reactions to what you just did?"	We started getting caught up in wordsmithing, even though I had told them to avoid that temptation at this stage. (This generated laughter all around.) What a principle is was becoming clearer, but the participants were still not sure. (I reassured them that they were right where they needed to be on the learning curve and were doing well.) Several participants expressed excitement about what was emerging and its potential importance.
		There was a question from a late arrival about what was going to happen with the principles generated. (I had to briefly restate the framing and purpose of our work together agreed to in Step 1. I asked a volunteer to provide more details to that person during the upcoming break.)

(Continued)

(Continued)

Transitions	Evaluation Check-In	Examples of Responses in the Obesity Principles Facilitation Example
Step 3 (comparing options) to Step 4 (choosing among options): content check-in	"We've been comparing options and, in so doing, have developed and applied criteria for choosing among options. The next step is to make decisions about which options to pursue. Before doing so, are there any final thoughts about the comparisons we've done and the discussions we've just had?"	"Since our choices are really constrained by limited resources, I think we ought to consider the option of trying to get more resources for evaluation." "I realized there many more options and we were able to consider, but I'm comfortable that we have discussed the right ones. I'm ready to move on." Question: "What happens if we think of new options as we try to narrow down?" I assured the group that the process was nonlinear. If new options and ideas emerge, we would decide how to engage with them.
Step 3 to Step 4: process check-in	"Before we move on to narrowing options, I'd like to get your reaction to the process we use for making comparisons."	The work has become easier as participants have gotten to know each other. It has been helpful to keep focused on what would be useful and to be practical about what is possible.
Step 4 (choosing among options) to Step 5 (determining next steps): content check-in	"We've reached another transition stage. We've narrowed the options, focused the evaluation, and decided on the major evaluation questions and methods. We now need to finalize those decisions and decide how to communicate them to the people affected. Any final comments or reflections on the decisions you've taken?"	"We wish we could do more, but it was good to be forced to choose." "We learned a lot about evaluation, especially how to use criteria to make choices." "We need to be sure to report why we decided what we decided not just what we decided." "We need to report some of the different perspectives that informed our choices." "We should note what more could be done with more resources."

Transitions	Evaluation Check-In	Examples of Responses in the Obesity Principles Facilitation Example
Step 4 to Step 5: process check-in	"Once again, as we have done before, let's do a process check-in. You made some tough choices. Have you feel about the process you've just experienced?"	"We appreciated the different perspectives and need to compromise when we disagree." "It was hard at first, but then we got a rhythm and it worked." "There is no way we can communicate the depth and richness of our discussion; whatever report we produce, a lot will be left out, which is too bad." "We should, however, explain and describe the process, even though we can't communicate its fullness."

Step 5. Articulate and Communicate the Decision and the Rationale for the Decision to Those Who Must Implement the Evaluation.

The final step is one that is often overlooked: How will the group's decisions be communicated to those not a part of the facilitated process? Sometimes—indeed, often—part of the facilitator's job is to write the summary of the group's work. That may be the default option, but it is not the only possibility. I put the issue to the group. I am always watching for participants who have been especially energetic and enthusiastic and pose the possibility of writing the draft summary for the group. Or participants may each write a part of the summary document to spread the workload and deepen the shared ownership. How and when the draft will be reviewed by participants before being finalized must also be decided. Everyone should leave a facilitated session knowing what has been decided and what will happen next.

At this point, the group's work is hopefully done and its goals have been achieved. Part of the facilitator's job is to review what has been accomplished, anything not accomplished, and to conclude by inviting reflections on both the content of what was generated and the process for generating the group's conclusions, products, and recommendations.

FACILITATION LESSONS

1. ***Respect and apply the 80/20 rule.*** Evaluation facilitators have a responsibility to help stakeholders identify those vital few questions among the trivial many that are high in payoff and add real value to the evaluation's utility. The 80/20 rule expresses the importance of focusing on the right information: In general, 20% of the information obtained accounts for 80% of what's worth knowing (Koch, 2014). The trick is identifying and focusing on that critical 20%. Many questions are interesting, but which are crucial?

2. ***Go slow to go fast later.*** Whereas an evaluation facilitator may feel pressure to move the group toward making decisions fast, it is accepted wisdom among experienced facilitators that diverse stakeholders are unlikely to be able to compare options and prioritize strategies without a full exchange of perspectives, building trust, generating mutual understanding, and negotiating inclusive agreements. "When engaging across diverse perspectives, go slow at first to build common ground and a shared based of information. Invest in fostering trust-based relationships from which innovative and supported solutions can emerge and be implemented" (Laberge, 2013, p. 1). Taking time to thoughtfully compare options ultimately pays off by speeding the process of deciding among them. Go slow to go faster later is a facilitation mantra intertwined with the proverb:

   If you want to go fast, go alone.

   If you want to go far, go together.

3. ***Be explicit about trade-offs.*** Since no evaluation can examine everything, designing an evaluation often involves trade-offs. An example is the trade-off between breadth versus depth. Getting more data usually takes longer and costs more, but getting less data usually reduces confidence in the findings. Studying a narrow question or very specific problem in great depth may produce clear results but leave other important issues and problems unexamined. On the other hand, gathering information on a large variety of issues and problems may leave the evaluation unfocused and result in knowing a little about a lot of things but not knowing a lot about anything.

   Once stakeholders get turned on to learning from evaluation, they want to know everything. The evaluation facilitator's role is to help them move from a rather extensive list of potential questions to a much shorter list of realistic questions and finally to a focused list of essential and necessary questions. This process moves from generating many possibilities (divergence) to focusing on a few worthwhile priorities (convergence).

4. ***Balance competing demands.*** Stakeholders want accurate information; they apply "truth tests" (Weiss & Bucuvalas, 1980) in deciding how seriously to pay attention to an evaluation. They also want useful and relevant information. The ideal, then, is both truth and utility. But these may be in conflict and choices must be made about which to emphasize more. A simple example of such a choice is time. The time lines for evaluation are often ridiculously short. A decision maker may need whatever information can be obtained in three months, even though researchers insist that a year is necessary to get data of reasonably quality and accuracy. This involves a trade-off between truth and utility. Highly accurate data in a year will be useless to this decision maker. Utility demands getting the best data possible in three months. Facilitating evaluation means facilitating trade-offs and competing demands.

5. ***Watch for and assess threats to data quality.*** Evaluator facilitators have an obligation to think about, anticipate, and provide guidance about how threats to data quality will affect interpreting and using results. However, it is impossible to anticipate all potential threats to data quality. Even when faced with the reality of particular circumstances and specific evaluation problems, it is impossible to know in advance precisely how a design or measurement approach will affect results. For example, having program staff conduct client interviews in an outcomes evaluation could (1) seriously reduce the validity and reliability of the data, (2) substantially increase the validity and reliability of the data, or (3) have no measurable effect on data quality. The nature and degree of effect would depend on staff relationships with clients, how staff members were assigned to clients for interviewing, the kinds of questions being asked, the training of the staff interviewers, attitudes of clients toward the program, and so on. Program staff might make better or worse interviewers than external evaluation researchers, depending on these and other factors. Evaluation facilitators can help stakeholders grapple with these kinds of data quality questions. No automatic rules apply. There is no substitute for thoughtful analysis based on the specific circumstances and information needs of a particular evaluation, both initially and as the evaluation unfolds.

6. ***Watch for and assess threats to utility.*** Guide the group in staying utilization focused as options are generated, compared, and prioritized. Generating actionable, answerable, relevant, and timely questions enhance the potential for utility. Credible and meaningful answers to those questions seal the deal.

7. ***Good enough is good enough.*** Decision makers regularly face the need to act with imperfect information. They prefer more accurate information to less accurate information, but they also prefer some information to no information. The effects of methodological quality on use must be understood in the full context of a study, its political environment, the degree of uncertainty decision makers face, resources available for evaluation, and their relative need for whatever clarifying information they can get. If information is scarce, then new, timely information, even of less-than-ideal quality, may be somewhat helpful. Facilitating evaluation means guiding those involved toward what is doable and practical, not a textbook ideal. Good enough is not bad.

8. ***Facilitate utilization-focused methods decisions***. The scope and importance of an evaluation, which is another way of saying intended use by intended users, greatly affects the emphasis that should be placed on technical quality. Eleanor Chelimsky, former president of the AEA and founding director of the Program Evaluation and Methodology Division of the U.S. Government Accountability Office (GAO), has insisted that technical quality is paramount in policy evaluations to Congress. The technical quality of national policy research matters not only in the short term, when findings first come out, but over the long term as policy battles unfold and evaluators are called on to explain and defend important findings (Chelimsky, 2007).

   On the other hand, debates about technical quality are likely to be less rancorous in local efforts to improve programs at the street level, where the policy rubber hits the day-to-day programming road. Experienced local decision makers often apply less rigorous standards than academics and, if they find the evaluation effort credible and serious, they're more interested in discussing the substance of findings than in debating methods. Credibility involves more than technical quality, as important as technical quality is. Credibility, and therefore utility, is affected by balance, transparency, and open discussion of both strengths and weaknesses.

In facilitating methods choices, then, the utilization-focused facilitator guides the group in considering how attention to technical quality is tied to and balanced by concern for relevance and timeliness. As no study is ever methodologically perfect, it is important when facilitating evaluation to help those involved in design decisions to be thoughtful about which imperfections they are willing to live with in making the inevitable leaps from limited data to decisions and action.

9. ***Connect methods, credibility, and use.*** Methods and measurement choices affect the credibility of findings. That's basic. And credibility affects use. When facilitating methods negotiations among primary intended users, stay aware of and help intended users stay attentive to these fundamental connections among methods, credibility, and use.

Quality Methods → Credible Findings → Enhanced Utility

10. ***Guide the group in anticipating criticism without becoming defensive.*** Given that that quality and excellence are situational and that different researchers, evaluators, and stakeholders emphasize varying and diverse criteria for what constitutes credible evidence, it is futile to attempt to design studies that are immune from methodological criticism. There simply is no such immunity. When facilitating an evaluation design group, I consider it part of my responsibility to prepare participants for criticism regardless of what choices they make. People applying different criteria will reach different conclusions and make different judgments. Surfacing, examining, and negotiating criteria is part of the evaluation facilitator's role and function. Helping those involved anticipate how the criteria they prioritize may be criticized is also part of the job.

SUMMARY

We've covered a lot of facilitation territory in this chapter on the principle of engaging through options. The chapter opened with 12 research findings about how groups deliberate on and decide among options. I then reviewed what makes evaluation decision making challenging for stakeholders, including the myriad of approaches, models, and methods to choose from and the impossibility of systematically considering all possibilities. That set the stage for introducing the GUIDE criteria for engaging through options.

Guiding wisdom: Provide a process for generating and comparing options, exercising choice, and making decisions.

Useful knowledge: Review practical options and facilitate making choices that will enhance an evaluation's relevance, credibility, and utility.

Inspiring values: Reinforce the importance of stakeholders' perspectives, experiences, and values to inform choices and the significance of those choices. Engage and empower.

Developmental adaptation: Generate and adapt options to fit the evaluation and stakeholders' situation, revising options as appropriate throughout the process.

Evaluable facilitation: Evaluate progress and provide opportunities for participants to express their reactions to and experiences of choosing among options.

I also used the GUIDE criteria to provide illustrations of how to facilitate the five steps for facilitating options. I closed with 10 facilitation lessons relevant to the engaging through options principle.

Practice Exercise

Imagine that you are facilitating a group of diverse stakeholders to design an evaluation in your area of expertise. What example of methodological options would you offer the group to illustrate why their perspective was needed to inform methods choices? For example, in this chapter, I noted that I like to use the example of even-numbered versus odd-numbered survey scales as an example that requires a decision about what would be the most useful approach for a particular target population and evaluation purpose. What example would you offer a group? Why? How would present the choices? What guidance would you offer in facilitating the group's deliberations of alternatives?

TOOLS AND RESOURCES FOR FACILITATING ENGAGEMENT WITH OPTIONS

1. **The Designer's Holy Triangle.** An illuminating practice exercise to help a group deal with trade-offs in choosing among options. Exhibit 5.7 presents the exercise.

Exhibit 5.7 The Designer's Holy Triangle

Imagine that you are designing an evaluation. You are presented with three options, but because of time and resource limitations, you are limited to choosing two of the three options. What are the implications of these choices in commissioning an evaluation?

Options: Fast, Good, or Cheap: Pick 2

Option 1. Good + Fast = Expensive

Commissioning an evaluation that is good (high-quality design) and done quickly (short timelines) by prestigious evaluators will be expensive. Why?

Option 2. Good + Cheap = Slow

Commissioning an evaluation that is good (high-quality design) and done as prestigious evaluators are available will be slow to complete. Why?

Option 3. Fast + Cheap = Inferior Quality

Commissioning an evaluation that is fast and done on the cheap can lead to a poor-quality evaluation that is not useful to the organization. Why**?**

Source: © iStockphoto.com/vladwel

This is capacity-building exercise to set the stage for facilitating choices. Ask participants to share their own non-evaluation examples of these trade-offs. Lots of laughs ensue as examples are shared and the group begins to think in terms of comparing options and calculating trade-offs.

2. **Matching purposes with evaluation questions.** Exhibit 5.8 can be used when facilitating potential evaluation questions for different evaluation purposes.
3. *A Practical Guide for Engaging Stakeholders in Developing Evaluation Questions*, Preskill and Jones (2009).
4. *The Rainbow Framework*, Better Evaluation: http://www.betterevaluation.org/en/plan
5. *Decide Which Evaluation Method to Use*, Better Evaluation:

   http://www.betterevaluation.org/en/start_here/decide_which_method
6. *Sixty Methodological Potholes*, David Huron (2000):

   http://csml.som.ohio-state.edu/Music829C/methodological.potholes.html

Exhibit 5.8 Evaluation Questions Matched to Alternative Evaluation Purposes

Summative evaluation questions that lead to overall judgments of merit, worth and significance:

- To what extent and in what ways does the program meet participants' needs?
- To what extent have intended outcomes been attained?
- What have been the larger impacts of the program?
- To what extent can documented outcomes be attributed to the intervention?
- Is this an especially effective practice that should be funded and disseminated elsewhere as a model program?
- How do outcomes and the costs of achieving them compare to other intervention or program options? Does the program add value for money?
- What are unanticipated outcomes and impacts, if any?

Formative program improvement-oriented questions:

- What are the program's strengths and weaknesses? What works and what doesn't?
- What implementation processes need to be improved, if any?
- How are participants reacting to the program? What do they like and dislike? What do they find valuable? What do they resist? What factors seem to be affecting program completion?
- How do different subgroups in the program respond; that is, what works for whom in what ways and under what conditions?
- What are program staff reactions? What are their perceptions of what could be improved?
- Where are opportunities for improvement? How can outcomes and impacts be increased? How can costs be reduced? How can quality be enhanced?

Accountability questions:

- Are funds being used for intended purposes?
- Are resources being efficiently allocated?
- Is implementation following the approved plan?
- Are staff qualified? Are only eligible participants being accepted into the program? Are partners performing as promised and expected?
- Are quality control mechanisms in place and being used?
- Are goals, targets, and key performance indicators being met?
- Are problems being handled?

Ongoing monitoring questions:

- To what extent are inputs and processes flowing smoothly and according to plan?
- What are entry, participation, completion, and dropout rates? Are these changing?
- Are outputs being produced as anticipated and scheduled?
- Where are bottlenecks occurring, if at all?
- What are variations across subgroups or sites?
- Are funds coming in and expenditures going out in accordance with the approved budget?
- What, if anything, is changing in the program's context that is affecting or could affect operations?

Developmental evaluation questions for innovative interventions in complex systems:

- What is being developed?
- How is the intervention adapting to complex, emergent, and dynamic conditions?
- What innovations are emerging and developing? With what effects?
- What's happening at the interface between what the program is doing and accomplishing and what's going on the larger world around it?
- What complex systems changes are occurring?

Knowledge-generating questions:

- What are general patterns and principles of effectiveness across programs, projects, and sites?
- What lessons are being learned?
- How do evaluation findings triangulate with research results, social science theory, expert opinion, practitioner wisdom, and participant feedback?
- What principles can be extracted across results to inform practices and models in new settings?

7. *CDC Evaluation Resources*, Center for Disease Control (2016): https://www.cdc.gov/eval/resources/index.htm
8. *Choosing Appropriate Evaluation Methods Tool*, Bond (2016): https://www.bond.org.uk/resources/evaluation-methods-tool

9. *Evaluation Focus Options*, Patton (2008, Chapter 7).
10. *Section 1. A Framework for Program Evaluation: A Gateway to Tools*, Community Tool Box (2016): http://ctb.ku.edu/en/table-of-contents/evaluate/evaluation/framework-for-evaluation/main

BETWEEN-CHAPTERS PORTAGE

MQP introduction: Leah Moses founded The Improve Group in 2000. In working with Leah and her staff, I've been impressed by their creative capacity to adapt methods through facilitated engagement. This requires building trust and mutual understanding with special attention to strengthening relationships through skillful facilitation that is culturally sensitive and respectful. I invited the Improve Group to share an example that illustrates methods-focused evaluation facilitation.

Photo © Dan Goldstein, The Improve Group.

Facilitating an Evaluation Process to Adapt Methods and Thereby Enhance an Intervention

by Kate Noble (left in photo), Kayla Mueller (right), Leah Goldstein Moses (center), and Jessica Hallstrom (not pictured)

How do you facilitate an evaluation when you know the intervention you are evaluating isn't reaching a large portion of the population?

That's the case for many Latina women and folic acid. Identified during the 1960s as an important dietary supplement that can prevent birth defects, public health practitioners succeeded in getting folic acid added to all fortified grain products starting in 1998.[1] Today, you can find folic acid in most cereal, bread, and other similar packaged products. But many women, including Latinas, who are more likely to eat rice, tortillas, or other grains instead of cereal or bread, may not get enough of this important supplement. As a result, Latina women are at greater risk of having babies with birth defects.

To address the shortcomings of this intervention, the Minnesota Department of Health (MDH) partnered with the Improve Group to explore new strategies

1. Crider, K. S., Bailey, L. B., & Berry, R. J. (2011, March). Folic acid food fortification—Its history, effect, concerns, and future directions. *Nutrients*, *3*(3), 370–384. Available online at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3257747/

to increase folic acid intake among Latina women. They envisioned a project where Latina women would engage in focus group discussions to hear about and react to a few different program and service models. In our initial conversations with MDH, we learned that they hoped to find strategies effective for the entire Latina population instead of distinct programs targeted at specific age, geographic, or cultural groups. To capture these overarching insights about Latina women and folic acid, we shifted the focus group from having participants respond to specific program ideas to engaging them in more creative problem solving.

It became clear early in the project that this evaluation would engage a population without an existing connection to MDH—and that focus group participants may have limited knowledge of folic acid and be reluctant to talk about supplements, eating habits, or health. We knew almost immediately that we had to leverage the power of relationships to facilitate this process.

First, we identified some key advisors and engaged them in the design of our focus groups. These advisors worked closely with Latina women during and after pregnancy and included a nurse, a social worker, and a public health educator. Our advisors helped us to restructure our focus group questions, to think more critically about who we wanted to engage in the focus groups (including intentionally recruiting recent immigrants as well as women who had spent much or all of their lives in the United States), and to consider the limitations of using focus groups versus surveys.

Second, we selected communities that had substantial Latina populations and hired recruiters from these communities. Recruiters used their own social and professional networks to identify women to participate in the focus group. These recruiters helped us engage much more diverse focus group participants than we would have had using more standard random-selection phone calls, e-mails, or letters. As trusted members of their communities, our advisors helped to allay fears participants' about being targeted or belittled during the focus groups.

Third, we used very creative strategies to engage women during the focus groups to help set them at ease and enable them to think creatively about a new or unfamiliar issue. We brought many Spanish-language magazines to the focus groups, all sourced from supermarkets recommended by our advisors and recruiters as places where Latina women would shop. After placing these magazines out on the table, we asked participants to create a collage of clippings illustrating what a healthy lifestyle looked like to them. Then, women worked in pairs to identify commonalities and differences in their collages. This was followed by a group discussion about the themes that emerged about what defines a healthy lifestyle and how that might connect to having a healthy pregnancy. Finally, after participants reflected on themes related to a healthy pregnancy, they were given some information about folic acid and asked to design a plan that would ensure more Latina women got the folic acid they need. Participants were enthusiastic and engaged during this part of the focus group—and our MDH colleagues were pleasantly surprised at the level of thoughtfulness, detail, and nuance that participants brought to their plans.

Our focus groups were designed intentionally to support relationship-building—progressing from individual reflection to working in pairs to a full group discussion. As a result, by the time participants were designing their own folic acid strategies, they felt at ease and shared many great and creative ideas. At the very end of the focus group, however, we changed tactics: Some of our MDH colleagues were concerned that the collage and design sections may not yield specific, actionable information, and so we asked participants to respond to existing messages and materials that MDH wanted us to explore. The structure of these questions ran counter to the relationship-focused approach we used initially—and participants' responses to them were much less rich and detailed.

As we analyzed data following the focus group, we turned again to our key advisors for insight. They helped us to take a broader systems lens to our findings. For example, participants initially identified laziness as one barrier to folic acid. Our key advisors helped us to reframe that finding by explaining that Latina women are often constrained by busy schedules, exhausted from work and personal commitments, and uncertain in the face of a constant information barrage. They helped us understand that these factors, and not laziness, prevent many Latinas from recognizing the importance of folic acid and acting on it by increasing their intake through foods or supplements.

Throughout the project, relationships were critical to facilitating the evaluation—from clarifying questions to recruiting participants to gathering data to understanding results.