

# Introduction, Overview, and Context

Utilization-Focused Reality Testing:  
Finding Out if What Js Hoped  
for Actually Happens

*A theory must be tempered with reality.*

*Jawaharlal Nehru (1889–1964)  
Prime Minister of India*

**K**iribati is an island nation in the central Pacific. The economy depends on producing coconut oil and fishing, but overfishing threatened to become an economic crisis. The Kiribati government created a program to subsidize the coconut oil industry with the goals of increasing incomes from coconut production and reducing overfishing. The idea was that if people spent more time growing coconuts, they would spend less time fishing.

On the face of it, this is a reasonable policy created by government officials who know their country and its people. The question is: Did it work? Answering that question is the job of evaluation. Lots of good ideas don't work in practice. Some do. Evaluation helps distinguish those ideas that work from those that don't. So what happened in Kiribati?

Sheila Walsh, a postdoctoral researcher at Brown University working with the Scripps Institution of Oceanography, interviewed Kiribati fishermen and helped monitor changes in fishing. She found that fishing increased by 33% and the reef fish population dropped by an estimated 17%, putting the whole ecosystem at risk. It turned out that paying people more to do coconut agriculture actually increased fishing. Why? How can that be? Walsh concluded that as people earned more money making coconut oil, they could work less to support themselves and spend more leisure time fishing. They didn't just fish for income. They fished because they liked to fish and so having more income from coconut production gave them more time to fish (Harris, 2009).

## 2 ■ ESSENTIALS OF UTILIZATION-FOCUSED EVALUATION

The program aimed at reducing overfishing actually made the situation worse, a classic example of an unintended consequence. Based on the evaluation findings, Walsh began working with the government to try other interventions to reduce overfishing, like creating new jobs for fishermen by hiring them to use their boats to patrol newly created ocean preserves. Those programs will also need to be evaluated to find out if they work.

This is an example of an evaluation used to support making decisions about what works and doesn't work. We can infer several factors that made the Kiribati evaluation useful. The government officials who implemented the program had to be willing to have it evaluated. The evaluation questions needed to be relevant and meaningful to those officials. The evaluator needed to be credible and produce credible findings based on credible methods. The intervention (the program) had to be sufficiently well conceptualized and implemented that the extent to which it attained the desired goals could be determined. There needed to be sufficient resources to undertake a credible evaluation. The fishermen had to cooperate with the evaluation, answering questions about their reactions to the program and how their practices and behaviors were affected. Credible statistics about ecosystem effects (overfishing) had to be collected with agreement among those involved about how to define "overfishing"; that is, *criteria for making judgments* had to be articulated and agreed on. The findings had to be actionable. When the results proved negative, the officials had to be motivated to engage with the evaluator about the implications of the findings and, using what was learned from the evaluation, look for alternative approaches. In so doing, the government officials and evaluator formed an ongoing relationship of inquiry, evaluation, and action.

A half century of research on evaluation use has validated the importance of these factors and how they relate to each other (Patton, 2008). This book draws on that research and organizes the factors into a framework and set of steps that, taken together, increase the likelihood that evaluations will be useful—and actually used. That framework is called *utilization-focused evaluation*.

### *What Is Evaluation?*

Before presenting the utilization-focused evaluation framework, let's begin by looking at these things called "evaluations" that we hope to see used. To evaluate something means determining its merit, worth, value, or significance. Program or project evaluations typically involve making the following kinds of judgments: How effective is the program? To what extent has the program been implemented as expected? What goals, outcomes, and results were achieved by the program? To what extent and in what ways did program participants benefit, if at all? What needs of participants were met? What unanticipated consequences resulted from the program? What are the strengths and weaknesses of the

program, and how can it be improved? What worked and what didn't work? What has been learned in this program that might be useful to other programs? To what extent do the benefits of the program provide sufficient value to justify the costs of the program? Should the program's funding be maintained as is, increased, or decreased? Evaluations, then, typically describe and assess what was intended (goals and objectives), what happened that was unintended, what was actually implemented, and what outcomes and results were achieved. The evaluator will then discuss the implications of these findings, sometimes including items for future action and recommendations.



### Evaluation as Defined in the *Encyclopedia of Evaluation*

Evaluation is an applied inquiry process for collecting and synthesizing evidence that culminates in conclusions about the state of affairs, value, merit, worth, significance, or quality of a program, product, person, policy, proposal, or plan. Conclusions made in evaluations encompass both an empirical aspect (that something is the case) and a normative aspect (judgment about the value of something). It is the value feature that distinguishes evaluation from other types of inquiry, such as basic science research, clinical epidemiology, investigative journalism, or public polling (Fournier, 2005, p. 140).

In the simplest terms, evaluations answer three questions: What? So what? Now what?

**What?** What happens in the program? What services and experiences does the program offer? What activities and processes occur? What changes in attitudes, knowledge, skills, and/or behaviors, if any, occur in participants? What outcomes and impacts result from the program? What unanticipated outcomes emerged? What are the documented costs and benefits of the program?

**So what?** So what do the findings *mean*? Why did the results turn out as they did? What are the implications of the findings? What judgments can be made? To what degree and in what ways can the program be considered a success? A failure? A mixed bag of positives and negatives? How does this program compare to other programs? What sense can we make of the findings?

**Now what?** What recommendations flow from the findings? What improvements should be made? Should its funding be continued, expanded, reduced, or ended? Should others adopt the program? Or avoid it? In short, what actions flow from the findings and interpretations of the findings?

## 4 ■ ESSENTIALS OF UTILIZATION-FOCUSED EVALUATION

Let's apply this framework to the Kiribati case.

**What?** Increased coconut production did not reduce fishing. Overfishing actually increased during the intervention.

**So what?** The program did not work as hoped. Fishermen don't just fish for income. They fish because they like to fish. Getting money from other sources won't reduce their fishing.

**Now what?** Look for alternative approaches to reduce overfishing.

### *Evaluation Reports—and Beyond*

Often evaluation questions are answered in formal reports. Some evaluation reports are entirely internal to an organization for use by staff and administrators to support ongoing managerial decision making. Other evaluation reports are published or posted on the Internet to meet an obligation for public accountability or to share lessons learned. But producing an evaluation report is not the purpose of evaluation. Evaluation is not an end in itself. The purpose is to inform thought and action. Moving from *what*, to *so what*, to *now what* means moving from data to interpretation to action. Action flows from using evaluation findings. Getting evaluations used is what *utilization-focused evaluation* is all about.

### Utilization-Focused Evaluation: Overview

Utilization-focused evaluation (U-FE) begins with the premise that evaluations should be judged by their utility and actual use; therefore, evaluators should facilitate the evaluation process and design any evaluation with careful consideration of how everything that is done, *from beginning to end*, will affect use. Use concerns how real people in the real world apply evaluation findings and experience the evaluation process. Therefore, the *focus* in utilization-focused evaluation is on *intended use by intended users*. Since no evaluation can be value-free, utilization-focused evaluation answers the question of whose values will frame the evaluation by working with clearly identified, primary intended users who have responsibility to apply evaluation findings and implement recommendations.

Utilization-focused evaluation is highly personal and situational. The evaluation facilitator develops a working relationship with intended users to help them determine what kind of evaluation they need. This requires negotiation in which the evaluator offers a menu of possibilities within the framework of established evaluation standards and principles.



## Standards for Evaluation

The profession of evaluation has adopted standards to guide professional practice. Professional evaluators are challenged to take responsibility for use. If evaluations are ignored or misused, we have to look at where our own practices and processes may have been inadequate.

### Utility

The Utility Standards are intended to ensure that an evaluation will serve the information needs of intended users.

### Feasibility

The Feasibility Standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.

### Propriety

The Propriety Standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.

### Accuracy

The Accuracy Standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the feature that determine worth or merit of the program being evaluated.

### Accountability

The Accountability Standards aim to ensure that evaluations are conducted in accordance with standards of quality.

*For the full set of detailed standards, see Joint Committee on Standards for Educational Evaluation, 2010.*

Specific standards have also been adapted to various international contexts (Russon & Russon, 2004) and reviewed through the lens of cultural diversity (AEA Diversity Committee, 2004). See also the *Guiding Principles for Evaluators* (AEA, 2004).

Utilization-focused evaluation does not advocate any particular evaluation content, model, method, theory, or even use. Rather, it is a process for helping primary intended users select the most appropriate content, model, methods, theory, and uses for their particular situation. *Situational responsiveness* guides the interactive process between evaluator and primary intended users. This means that the interactions between the evaluator and

the primary intended users focus on fitting the evaluation to the particular situation with special sensitivity to context. A utilization-focused evaluation can include any evaluative purpose (formative, summative, developmental), any kind of data (quantitative, qualitative, mixed), any kind of design (e.g., naturalistic, experimental), and any kind of focus (processes, outcomes, impacts, costs, and cost-benefit, among many possibilities). *Utilization-focused evaluation is a process for making decisions about these issues in collaboration with an identified group of primary users focusing on their intended uses of evaluation.*

A psychology of use undergirds and informs utilization-focused evaluation: Intended users are more likely to use evaluations if they understand and feel ownership of the evaluation process and findings; they are more likely to understand and feel ownership if they've been actively involved; by actively involving primary intended users, the evaluator is training users in use, preparing the groundwork for use, and reinforcing the intended utility of the evaluation every step along the way.

Before we begin systematically examining the utilization-focused evaluation framework in depth and detail, let's look at some of the challenges that arise when trying to conduct useful evaluations. The Kiribati overfishing situation offers a positive example of evaluation use. Let's take a moment and look at the shadow side, evaluations that aren't used. Why would people ignore and resist evaluation findings? Well, we have a lot of research on that as well. Consider the case of DARE as an example.

### *Resistance to Reality Testing: The Case of DARE*

DARE stands for *Drug Abuse Resistance Education*, a 17-week drug education program targeted at middle school children in which police officers went 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). Thousands of evaluations of DARE were conducted in local districts as well as several national evaluations (GAO, 2003). These evaluations have consistently shown that the program “was not effective in actually keeping young people from using drugs” (C. H. 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” (C. H. Weiss et al., 2005, p. 15). Indeed, DARE is still going strong (DARE, 2011), though with a revised curriculum and more comprehensive approach. Evaluation findings remain negative. And DARE remains popular. Why?

The program is popular with police, parents, school officials, and elected representatives who are quite sensitive to what constituents like. In 1997 the attorney general of Minnesota, Skip Humphrey (son of Minnesota's well-known U.S. senator and vice president, Hubert Humphrey) conducted a major review of DARE. He concluded that the program was ineffective but too popular to eliminate: "I don't think we should dump something that's got a positive aspect to it. The public feels this is an affirmation of community standards" (O'Connor, 1997, p. A1). To cut to the chase, politics trumped evidence. Parents liked the idea that at least something was being tried, even if ineffective. Police liked connecting with students to make a positive impression. School officials and elected officials showed they were being responsive to widespread concerns about drug use. So what if it doesn't work? The Minnesota Institute of Public Health conducted a statewide survey of school officials, parents, police, and student in all 80 Minnesota school districts:

- 94% agreed: "DARE is very popular in our community."
- 88% agreed: "Even if there is no scientific evidence that DARE works, I would still support it." (O'Connor, 1997, p. A5)

Economist Edward Shepard (2002) found that between \$1 billion and \$1.3 billion was spent annually on DARE in the United States long after negative evaluation findings were widely reported. DARE has an effective national organization that promotes DARE and attacks negative findings about outcomes by emphasizing DARE's popularity with parents and police (Hanson, 2007). The DARE case illustrates the power of belief over evidence, the sometimes-dominating role of politics in undermining science, and the resistance that can be created when powerful stakeholders with great sums of money at stake work to undermine the use of evaluation findings.

Utilization-focused evaluation, then, is informed by studies of positive evaluation use, like the Kiribati example, as well as negative examples of evaluation resistance like DARE.

And before harshly condemning DARE supporters for failing to act on evaluation evidence, ask yourselves these questions: Are you consistently following evidence-based health advice? Eating nutritious foods? Exercising regularly? Engaging in safe sex? Are you following the evidence about how to nurture positive interpersonal relationships? Are you managing your finances according to well-established principles? Are you doing the things in your life that you know you should be doing based on evidence about what works? Do you even stay up to date about such evidence and use it to guide your decisions?

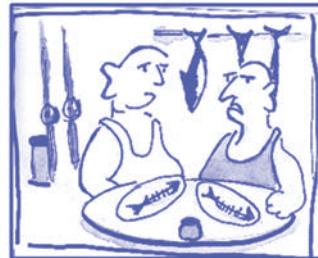
You see, evaluation use is not just about making program and policy decisions. It has to do with the larger issue of how we, as human beings, make decisions and engage in reality testing in all facets of our lives. Before looking systematically at the details of and steps in conducting utilization-focused evaluation, let's look a little more deeply into the complexities of reality testing. For a commitment to evaluation use ultimately depends on a commitment to reality testing.



*"If coconut subsidies don't work we can educate people to resist overfishing."*



*"I can visualize the campaign now:  
**JUST SAY NO TO  
OVERFISHING."***



*"Have you ever heard of DARE?"*

### Getting Ready for Evaluation: Engendering Commitment to Engage in Reality Testing

Farmers till the soil before planting seeds. They don't scatter seeds on hard, untilled soil, for that would mean the germinated seeds wouldn't take root and grow into healthy plants. Likewise, evaluators have to prepare people in programs and organizations to undertake evaluation. Providing evaluation findings to people who are not prepared to hear and use them means that the findings will not take root and grow into action.

Evaluation aims to be logical, rational, and empirical. Evaluation tests reality. Evaluation involves asking: Is what program funders and staff hope and think is happening really happening? Are desired results being attained? But people aren't automatically inclined to welcome such questions—or reality testing generally. Quite the contrary. We are inclined to believe what we want to believe and treat our beliefs as reality.

Perception is treated as reality. Sociologist W. I. Thomas posited that **what is perceived as real is real in its consequences**. This has become known as the *Thomas Theorem*. It captures succinctly a great deal of what social and behavioral science has discovered about how we function as human beings. We routinely treat our perceptions as if they are reality. We act on those perceptions and our actions have consequences. In the DARE program, many parents, educators, and public officials perceived that putting a drug resistance program into schools would reduce drug use. They wanted to believe this. They did believe this. And when the evidence showed the program did not reduce drug use, they preferred their perceptions and beliefs to reality and said to themselves and each other, “It’s better to do something than nothing, so DARE is good.” They didn’t look at negative consequences and opportunity costs (meaning that money spent on DARE was not available to support other potentially more effective interventions).

But this happens in many areas of our personal lives. Studies regularly show that people perceive themselves as being good drivers, above-average drivers, even excellent drivers when they are at best quite ordinary. People driving cars, even at high speeds, believe that they can safely talk on mobile phones or send text messages. Interestingly, most people don’t think that other drivers should do these things, but they view themselves as more competent and in control. Put that together with studies that show that incompetent people don’t realize how incompetent they are. Poor communicators perceive themselves as excellent communicators. Tyrannical bosses perceive themselves as excellent managers. Married men perceive themselves as sharing equally in household tasks while their wives perceive them as making only token contributions.

We like to perceive ourselves as rational and thoughtful about money, but the new field of behavior finance consistently documents that our decisions about money are irrational and highly emotional. People take pride in making their own decisions, but the evidence shows that people are heavily influenced by those around them. Following the herd and *groupthink* are common. For example, Salganik and Watts (2008), in an article titled “Leading the Herd Astray: An Experimental Study of Self-Fulfilling Prophecies in an Artificial Cultural Market,” reported results of a devious experiment in an online music market in which more than 12,207 participants listened to, downloaded, and ranked songs by unknown bands. Participants could see each others’ ratings, but what they didn’t know was that the researchers inverted the ratings, reporting as most popular those that were actually least popular. As new listeners joined the experiment, they followed the false ratings and rated highest those songs that had initially been rated lowest. Good news for quality: Over time, a few of the songs initially rated high returned to high ratings despite the false low ratings fed to new participants, but even that process was affected by social norms about what constitutes quality.

We think that we’re making autonomous decisions when, in fact, we’re actually like sheep, thus the “following the herd” metaphor. It happens on Wall Street where money

managers' stock recommendations and economists' market predictions notoriously manifest following the herd and groupthink patterns. It happens in mainstream media reports as when media of all kinds reported positively on the U.S. invasion of Iraq and only gradually did negative reporting emerge as the war became a quagmire and public opinion turned.

The advocates of DARE became more adamant in their belief in DARE as they shared their views with each other and took comfort from finding their perceptions reinforced by others, both other people like themselves (other parents and teachers) and people in authority (police officers and elected officials). *Groupthink*.

We are all socialized to believe that certain things are true and are discouraged from questioning those beliefs. We tend to belong to groups of people who share the same beliefs we do, so we hear the same beliefs repeated as truth and accept them as truth. This is certainly true of religious cults and fringe political groups, we all know that. But it's also true for all of us, for as social beings we seek out affinity groups of people who share our worldviews. In programs, then, staff tell each other that they're doing all they can, working hard, and making a difference—and believe that and want to believe that—and are not automatically open to some evaluator coming in and asking: What's your evidence that you're making a difference? This can be a threatening question, even a terrifying one. This is part of what evaluators face as we try to involve people in serious reality testing.

By the way, scientists are as prone to reality distortion as are ordinary people. Kevin Dunbar, director of the Laboratory for Complex Thinking and Reasoning at the University of Toronto Scarborough, has studied extensively how scientists interpret laboratory findings. He found that when lab experiments turned out differently from what scientists hypothesized, their first instinct was to explain away the results. When the same results were replicated, they would just ignore the findings and deal with cognitive dissonance by blaming imperfect methods and measures, and set about correcting and improving their methods and measures in search of results that would confirm their hypotheses. Dunbar is among those who has documented with social experiments of his own that scientists, like nonscientists, see what they want to see and interpret research findings through selective perception and layers of preconception and strong beliefs. Scientists operate with a strong "confirmation bias," expecting to confirm their theories and beliefs. Belief, especially scientific theory-based belief, can make us blind to what is actually happening (Dunbar & Fugelsang, 2005). The good news is that over time, with many scientific experiments and many scientists struggling in a domain of inquiry, the truth gradually emerges. But beliefs and theories die hard.

The fact is, we carefully edit our reality, searching for evidence that confirms what we already believe. Although we pretend we're empiricists—our views are dictated by nothing but the facts—we're actually blinkered, especially when it comes to information that contradicts our theories. The problem with science isn't that most experiments fail—it's that most failures are ignored (Lehrer, 2010a, reporting on Dunbar's studies of scientists).



## Of Paradigms, Reality Testing, and Decision Making

Substantial research has documented the limits of human rationality in decision making and reality testing. Daniel Kahneman was awarded the Nobel Prize in Economics for his pioneering work showing that how we decide what to do is far from rational. Our rationality is “bounded” by our beliefs about what is possible and our limited cognitive and emotional capacity to consider alternatives (Simon, 1957, 1978). This applies no less to well-educated professionals than to common folk. We all act out of deeply embedded heuristics, rules of thumb, standard operating procedures, long-practiced behaviors, socialized beliefs, and selective perception. We operate within and see the world through *paradigms* built on implicit assumptions, comfortable habits, values defended as truths, and beliefs projected as reality. Our paradigms tell us what is important, legitimate, and reasonable. They tell us how to think and what to do (Kuhn, 1970).

Programs are inevitably based on paradigms of belief about what constitutes a problem, how to change human behavior, what works and doesn’t work, and generally how to make the world a better place. Evaluation invites those working within a program paradigm to test the assumptions of the paradigm against the realities of how the program actually works in practice. This is no small thing. It is a lot to ask. It is even more to expect that when results raise fundamental doubts about the program’s effectiveness, adherents of the program will simply accept the evaluation’s findings. That expectation is based on a paradigm of logic and rationality that is inconsistent with how people take in and use data. Evaluators need to understand how decision making actually occurs to work with decision makers and get them ready to undertake reality testing evaluation. Committing to serious evaluation can involve a paradigm shift among those involved, a shift from acting on belief and hope to acting on data and feedback.

### *Reality testing and philosophy of science*

While we are discussing paradigms and reality testing, I should perhaps add that I am using the term “reality testing” in its common and ordinary connotation of finding out what is happening. While philosophers of science will rightly point out that the whole notion of “reality” is an epistemological and ontological quagmire, I find that the people I work with in the “real world”—*their* phrase—resonate to the notion of “reality testing.” It is their own sense of reality I want to help them test, not some absolute, positivist construction of reality. The notion that reality is socially constructed doesn’t mean it can’t be tested and understood.

Evaluators are no less susceptible to treating their perceptions as reality. Evaluators, on the whole, want to have their findings used. Evaluators prefer being useful to not being useful. And when I ask evaluators whether what they do is useful, they generally respond positively, affirming their perception and belief. And when I ask what evidence they have of their utility, they report perceptions, not evidence. Just as program people do. When I ask

program people how they know they're effective, they often say, "I just know." Evaluators are susceptible to the same illusion. The final step in a utilization-focused evaluation is to follow up our own evaluations and find out how they were actually used, to engage in our own reality testing and use what we learn to improve our practice.

Taken as a whole, social and behavioral sciences research shows that humans are more often reality distorters than reality testers. We operate with selective perception, social and cognitive biases, dominating worldviews, and within paradigms of perception and belief that make it hard to even ask questions about reality much less test our perceptions and beliefs against reality. When we ask people in programs to undertake evaluation, we are asking them to subject their perceptions and beliefs to empirical test: Is what they hope for and believe in actually happening? It takes some preparation to help people embrace this question. It requires tilling the metaphoric program soil so that those involved can receive the seeds of evaluation and those seeds can grow into useful findings that actually get used. That's why utilization-focused evaluation begins by assessing program and organizational readiness for evaluation—and engaging stakeholders in getting ready for evaluation, the subject of the next chapter. The principle that guides this readiness assessment, as in all of utilization-focused evaluation, is *listening before telling*. (For an insightful and provocative exploration of this principle applied to development, see Quarry and Ramírez, 2009.)

### The Utilization-Focused Evaluation Checklist: Overview of the Book

Pilots use checklists for every stage of flying: preparing to fly, taking off, navigating, landing, and taxiing to the gate. Checklists are used by even the most experienced pilots to assure that everything that needs to be done is done to avoid a crash. Checklists save lives in hospital emergency, intensive care, and surgical units (Gawande, 2007, 2009). There are checklists for virtually any human activity and a website devoted to monitoring and linking people to specialized checklists (<http://checklists.com>).

Under the leadership of Dan Stufflebeam, the Center for Evaluation at Western Michigan University has developed checklists for evaluation on evaluation methods, management, models, capacity building, institutionalization, values, criteria, standards, and metaevaluation (evaluation of evaluations).

A sound evaluation checklist clarifies the criteria that at least should be considered when evaluating something in a particular area; aids the evaluator not to forget important criteria; and enhances the assessment's objectivity, credibility, and reproducibility. Moreover, such a checklist is useful in planning an enterprise, monitoring and guiding its operation, and assessing its outcomes (Stufflebeam, 2000, p. 1).

This book is organized as a checklist of utilization-focused evaluation essentials. Following this introductory chapter, each subsequent chapter is an item or step in the

checklist. The concluding chapter includes a complete U-FE checklist summary of primary tasks and major facilitation challenges for each step. Periodically we will pause between steps to do a *complex dynamic systems interconnections review*. Portraying the utilization-focused evaluation process as a series of steps is driven by the necessarily linear and sequential nature of writing and, to some extent, the human thinking processes involved in figuring out how to do something. But the process is not neatly linear and sequential, and certainly not mechanical. There are interconnections among the steps, feedback loops, and the complex dynamics that affect any open and emergent system (Patton, 2011). To reflect these real-world system dynamics, I will periodically offer an interlude between “steps” to call attention to the importance of examining the iterative and nonlinear dynamics of a utilization-focused evaluation process. At the end of the book and online you will find the full utilization-focused evaluation checklist (Patton, 2010). Here, then, as an overview of the book, are the major steps in the utilization-focused evaluation checklist.

### *Utilization-Focused Evaluation Checklist*

A chapter is devoted to each of the 17 steps in the checklist below.

Step 1.	Assess and build program and organizational readiness for utilization-focused evaluation
Step 2.	Assess and enhance evaluator readiness and competence to undertake a utilization-focused evaluation
Step 3.	Identify, organize, and engage primary intended users
Step 4.	Situation analysis conducted jointly with primary intended users
Step 5.	Identify and prioritize primary intended uses by determining priority purposes
Step 6.	Consider and build in process uses if and as appropriate
Step 7.	Focus priority evaluation questions
Step 8.	Check that fundamental areas for evaluation inquiry are being adequately addressed
Step 9.	Determine what intervention model or theory of change is being evaluated
Step 10.	Negotiate appropriate methods to generate credible findings that support intended use by intended users
Step 11.	Make sure intended users understand potential methods controversies and their implications
Step 12.	Simulate use of findings
Step 13.	Gather data with ongoing attention to use

(Continued)

(Continued)

Step 14. Organize and present the data for interpretation and use by primary intended users
Step 15. Prepare an evaluation report to facilitate use and disseminate significant findings to expand influence
Step 16. Follow up with primary intended users to facilitate and enhance use
Step 17. Metaevaluation of use: Be accountable, learn, and improve



### PRACTICE EXERCISES

1. This chapter includes an example of a program evaluation that was used to change the intervention (the Kiribati overfishing initiative) and an example of resistance to evaluation findings (DARE). Find your own examples, an example of evaluation use and an example of resistance to evaluation use. You'll find examples reported in the news and posted on the Internet. It's worth having your own examples that are meaningful to you.
2. Develop an "elevator speech" for both evaluation in general and utilization-focused evaluation in particular. An *elevator speech* is a short description of something presented in the time it takes an elevator to go from the top floor to the first floor, or vice versa, in a tall building. The idea is that sometimes we meet important people in our lives in elevators, waiting to board an airplane, in coffee shops, at meeting receptions, and other places where, without notice, we are asked to explain what we do or something we care about. The idea of an *elevator speech* is be ready with a prepared presentation that grabs attention and says a lot in a few words. You never know what it might lead to.
3. Search the Internet for *recent research* on the human barriers to rational decision making and reality testing. Find examples of experiments and research that are relevant to your own field of interest. For instance, if you work in health, find examples of research on how people in hospitals make mistakes; classic examples include Atul Gawande (2002) and Jerome Groopman (2007). For samples of a half-century of research on the nonrational nature of human decision making, see Gigerenzer, Todd, and the ABC Research Group, 1999; Inbar, 1979; Kahneman and Tversky, 2000a, 2000b; Kuhn, 1970; Simon, 1957, 1978; Tversky and Fox, 2000; Tversky and Kahneman, 2000.