

From *Utilization-Focused Evaluation*, 4th ed
by Michael Quinn Patton (2008, Sage)

8

Evaluation Focus Options

Developmental Evaluation and Other Alternatives

*C*reative thinking may mean simply the realization that there's no particular virtue in doing things the way they always have been done.

—Rudolf Flesch

*J*f you can see in any given situation only what everybody else can see, you can be said to be so much a representative of your culture that you are a victim of it.

—S. I. Hayakawa

More Than One Way to Manage a Horse

Here is a story about the young Alexander, later to become *Alexander the Great*, as recorded by the ancient Greek historian and biographer Plutarch.

There came a day when Philoneicus the Thessalian brought King Philip a horse named Bucephalus, which he offered to sell for 13 talents. The king and his friends went down to the plain to watch the horse's trials and came to the conclusion that he was wild and quite unmanageable, for he would allow no one to mount him, nor would he endure the shouts of Philip's grooms, but reared up against anyone who approached. The king became angry at being offered such a vicious unbroke animal and ordered it led away. But Alexander, who was standing close by, remarked, "What a horse they are losing, and all because they don't know how to handle him, or dare not try!"

272 ■ FOCUSING EVALUATIONS: CHOICES, OPTIONS, AND DECISIONS

King Philip kept quiet at first, but when he heard Alexander repeat these words and saw that he was upset, he asked him: "Do you think you know more than your elders or can manage a horse better?"

"I could manage this one better," retorted Alexander.

"And if you cannot," said his father, "what penalty will you pay for being so impertinent?"

"I will pay the price of the horse," answered the boy. At this, the whole company burst out laughing. As soon as the father and son had settled the terms of the bet, Alexander went quickly up to Bucephalus, took off his bridle, and turned him towards the sun, for he had noticed that the horse was shying at the sight of his own shadow, as it fell in front of him and constantly moved whenever he did. He ran alongside the animal for a little way, calming him down by stroking him, and then, when he saw he was full of spirit and courage, he quietly threw aside his cloak and with a light spring vaulted safely onto his back. For a little while, he kept feeling the bit with the reins, without jarring or tearing his mouth, and got him collected. Finally, when he saw that the horse was free of his fears and impatient to show his speed, he gave him his head and urged him forward, using a commanding voice and touch of the foot.

King Philip held his breath in an agony of suspense until he saw Alexander reach the end of his gallop, turn in full control, and ride back triumphant, exulting in his success. Thereupon the rest of the company broke into loud applause, while his father, we are told, actually wept for joy. When Alexander had dismounted, he kissed him and said: "My boy, you must find a kingdom big enough for your ambitions. Macedonia is too small for you." (Adapted from Plutarch 2001:139–45)

More Than One Way to Focus an Evaluation

Young Alexander showed that there was more than one way to manage a horse. What I like most about this story, as a metaphor for managing an evaluation, is that he based his approach to the horse on careful observations of the horse and situation. He noticed that the horse was afraid of its shadow, so he turned him toward the sun. He established a relationship with the wild animal before mounting it. He was sensitive to the horse's response to the bit and reins. Alexander exemplified being active, reactive, interactive, and adaptive. Chapter 6 explored how these traits can serve an evaluator in being situational responsive. This chapter goes farther with situational responsiveness and contingency

thinking by presenting a broad range of evaluation options and identifying the factors that affect choosing a specific evaluation approach to match the priority information needs of primary intended users.

The last chapter focused on goals and outcomes as traditional ways to focus an evaluation. A program with clear, specific, and measurable goals is like a horse already trained for riding. Programs with multiple, conflicting, and still developing or ever-changing goals can feel wild and risky to an evaluator whose only experience is with seasoned and trained horses. Just as there's more than one way to manage a horse, depending on its characteristics, there's more than one way to manage evaluation of a program, depending on the program's characteristics and the environment in

which it operates. This chapter will examine why goals-based evaluation may not be as useful as some other options and offer alternatives for focusing an evaluation.

Problems with Goals-Based Evaluation

One can conduct useful evaluations without ever seeing an objective.

—Smith 1980:39

Alternatives to goals-based evaluation have emerged because of the problems evaluators routinely experience in attempting to focus on goals. In addition to fuzzy goals and conflicts over goals—problems addressed in the previous chapter—a longstanding concern has been that too much attention to measurable goals can distort a program's priorities. Lee J. Cronbach and Associates (1980) at the Stanford Evaluation Consortium, in their classic treatise on reforming evaluation, warned

It is unwise for evaluation to focus on whether a project has "attained its goals." Goals are a necessary part of political rhetoric, but all social programs, even supposedly targeted ones, have broad aims. Legislators who have sophisticated reasons for keeping goal statements lofty and nebulous unblushingly ask program administrators to state explicit goals. Unfortunately, whatever the evaluator decides to measure tends to become a primary goal of program operators. (P. 5)

In other words, what gets measured gets done. An example is when teachers focus on whether students can pass a reading test rather than on whether they learn to read. The result can be students who pass mandated competency tests but are still functionally illiterate. There are, then, two sides

to the goals sword: (1) a powerful focusing purpose (what gets measured gets done) and (2) a potentially distorting consequence (doing only what can be quantitatively measured, which is dependent on the state of the art of measurement and limited by the complexities of the real world).

Reification: Are Goals Real?

Another critique of goals is that they're not real. Since evaluation is grounded in reality testing, it behooves us to examine the reality of goals. To "reify" is to treat an abstraction as if it is real. Goals have long been a special target of social scientists concerned with concept reification. For example, Cyert and March (1963:28) asserted that *individual people have goals, collectiveness of people do not*. They likewise asserted that only individuals can act; organizations or programs, as such, cannot be said to take action, a matter of ongoing debate among sociologists (Fuchs 2007; Greenwood 2007). The future state desired by an organization (its goals) is nothing but a function of individual aspirations. In brief, *social scientists who study program goals are not quite sure what they are studying*. Organizational goals analysis is controversial and confusing. In the end, most researchers follow the traditionally pragmatic logic of pioneer organizational sociologist Charles Perrow (1970):

For our purposes we shall use the concept of an organizational goal as if there were no question concerning its legitimacy, even though we recognize that there are legitimate objections to doing so. Our present state of conceptual development, linguistic practices, and ontology (knowing whether something exists or not) offers us no alternative. (P. 134)

Like Perrow, funders, program staff, and evaluators are likely to come down on the

274 ■ FOCUSING EVALUATIONS: CHOICES, OPTIONS, AND DECISIONS

side of practicality. The language of goals will continue to dominate evaluation. By introducing the issue of goals reification, I have hoped merely to induce a modicum of caution and compassion among evaluators before they impose goals clarification exercises on program staff. Given the way organizational sociologists have gotten themselves tangled up in the question of

whether program-level goals actually exist, it is just possible that *difficulties in clarifying a program's goals may be due to problems inherent in the notion of goals rather than staff incompetence, intransigence, or opposition to evaluation*. Failure to appreciate these difficulties and proceed with sensitivity and patience can create staff resistance that is detrimental to the entire evaluation process.

Turbulent Environments and Changing Goals

A half-century ago, organizational sociologists discovered that the clarity and stability of goals are contingent on the degree of stability or turbulence in an organization's environment (Emery and Trist 1965). Evaluators, having traditionally defined their task as measuring goal attainment, have been slow to incorporate this understanding by adapting what we do to different conditions. Uncertainty includes things like funding instability, changes in governmental rules and regulations, mobility and transience of clients and suppliers, technological innovation, and political, economic, or social turbulence. What is important about classic works in organizational sociology (e.g., Azumi and Hage 1972; Hage and Aiken 1970) from an evaluation perspective is the finding that the degree of uncertainty facing an organization directly affects *the degree to which goals and strategies for attaining goals can be made concrete and stable*. The less certain the environment, the less stable and concrete will be the organization's goals. Effective organizations in turbulent environments adapt their goals to changing demands and conditions.

I have also hoped that reviewing the conceptual and operational problems with goals would illuminate why utilization-focused evaluation does not depend on clear, specific, and measurable objectives as the *sine qua non* of evaluation. Clarifying goals is neither necessary nor appropriate in every evaluation. Nowhere is this premise clearer than in goal-free evaluation.

Goal-Free Evaluation

Philosopher-evaluator Michael Scriven, a strong critic of goals-based evaluation, has offered an alternative: *goal-free evaluation*. Goal-free evaluation involves gathering data on a broad array of *actual effects* and evaluating the importance of

these effects in meeting demonstrated needs. The evaluator makes a deliberate attempt to avoid all rhetoric related to program goals. No discussion about goals is held with staff and no program brochures or proposals are read; only the program's actual outcomes and measurable effects are studied, and these are judged on the extent to which they meet *demonstrated participant needs*.

Scriven (1972b) offered four reasons for doing goal-free/needs-based evaluation:

1. To avoid the risk of narrowly studying the stated program objectives and thereby missing important unanticipated outcomes
2. To remove the negative connotations attached to the discovery of unanticipated

effects, because “the whole language of ‘side-effect’ or ‘secondary effect’ or even ‘unanticipated effect’ tended to be a put-down of what might well be the crucial achievement, especially in terms of new priorities” (pp. 1–2)

3. To eliminate the perceptual biases and tunnel vision introduced into an evaluation by knowledge of goals

4. To maintain evaluator objectivity and independence through goal-free conditions

In Scriven’s (1972b) own words,

It seemed to me, in short, that consideration and evaluation of goals was an unnecessary but also a possibly contaminating step. . . . The less the external evaluator hears about the goals of the project, the less tunnel vision will develop, the more attention will be paid to *looking for actual effects* (rather than checking on *alleged effects*). (P. 2)

Scriven (1972b) distrusted the grandiose goals of most projects. Such great and grandiose proposals “assume that a gallant try at Everest will be perceived more favorably than successful mounting of molehills. That may or may not be so, but it’s an unnecessary noise source for the evaluator” (p. 3). He saw no reason to get caught up in distinguishing alleged goals from real goals: “Why should the evaluator get into the messy job of trying to disentangle that knot?” He would also avoid goals conflict and goals war: “Why try to decide which goal should supervene?” He even countered the goals clarification shuffle:

Since almost all projects either fall short of their goals or overachieve them, why waste time rating the goals, which usually aren’t what is achieved? Goal-free evaluation is unaffected by—and hence does not legislate against—the shifting of goals midway in a project. (P. 3)

Scriven (1991b) also dealt with the fuzziness problem: “Goals are often stated so vaguely as to cover both desirable and undesirable activities, by almost anyone’s standards. Why try to find out what was really intended—if anything?” Finally, he has argued that “if the program *is* achieving its stated goals and objectives, then these will show up” in the goal-free interviews with and observations of program participants done to determine actual impacts (p. 180).

Goal-Free Evaluation: Bizarre Idea?

Contrary to a common argument that Goal-Free Evaluation (GFE) is a bizarre idea, most consumer product evaluation is done in GF mode. No one buying a car, for example, asks for a statement of the design team’s goals; the buyer usually has the family climb into each one under consideration, drives it around a bit, and haggles for the best price. In other words, one is driven by one’s own (perceived) needs assessment, not by matching goals to performance. GFE simply brought us back from the world of managerial values to that of consumer product evaluation (Scriven, personal communication).

For all its virtues, goal-free evaluation carries the danger of substituting the evaluator’s goals for those of the project, as evaluation theorist Marvin Alkin (1972) has posited.

This term “Goal-Free Evaluation” is not to be taken literally. The Goal-Free Evaluation *does* recognize goals (and not just idiosyncratic ones), but they are to be wider context goals rather than the specific *objectives* of a program. . . . By “goal-free” Scriven simply means that the evaluator is free to choose a wide context of goals. By his description, he implies that a goal-free

276 ■ FOCUSING EVALUATIONS: CHOICES, OPTIONS, AND DECISIONS

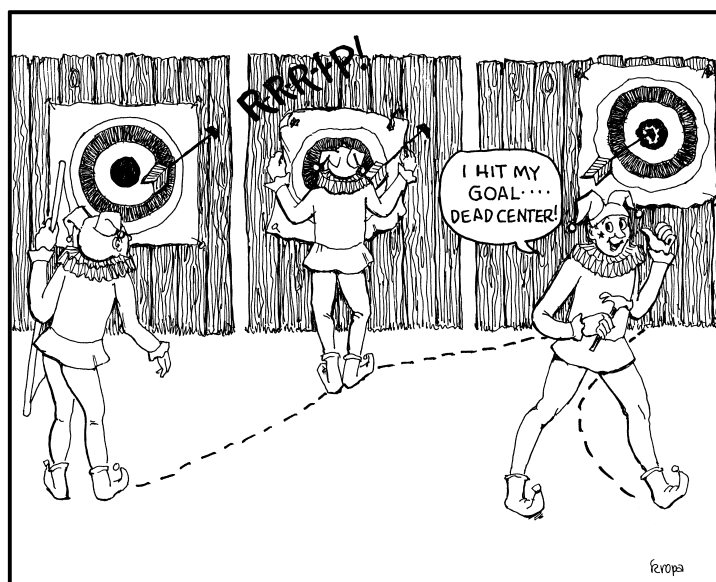
evaluation is always free of the goals of the specific program and *sometimes* free of the goals of the program sponsor. In reality, then, goal-free evaluation is not really goal-free at all, but is simply directed at a different and usually wide decision audience. The typical goal-free evaluator must surely think (especially if he rejects the goals of the sponsoring agency) that his evaluation will extend at least to the level of “national policy formulators.” The question is whether this decision audience is of the highest priority. (P. 11)

Here, then, Alkin raises the question of who the primary intended users are for a goal-free evaluation. In that regard, it should be noted that Scriven’s goal-free proposal assumes both internal and external evaluators. Thus, part of the reason the external evaluators can ignore program staff and local project goals is because the internal evaluator takes care of all that. Thus, goal-free evaluation is only partially goal free. Someone has to stay home and mind the goals while the external evaluators search for any and all effects. As Scriven (1972b) has stated

Planning and production require goals, and formulating them in testable terms is absolutely necessary for the manager as well as the internal evaluator who keeps the manager informed. That has nothing to do with the question of whether the external evaluator needs or should be given any account of the project’s goals. (P. 4)

In later reflections, Scriven (1991b:181) proposed “hybrid forms” in which one part of a comprehensive evaluation includes a goal-free evaluator working parallel to a goals-based evaluator. For our purposes, Scriven’s critique of goals-based evaluation is useful in affirming why evaluators need more than one way of focusing an evaluation.

Evaluation will not be well served by dividing people into opposing camps: progoal versus antigol evaluators. I am reminded of an incident at the University of Wisconsin during the student protests over the Vietnam War. Those opposed to the war were often labeled communists. At one demonstration, both antiwar and



Avoiding Posttraumatic Goal Nonattainment Syndrome

Your program's goals you need a way of knowing.

You're sure you've just about arrived,

But where have you been going?

So, like the guy who fired his rifle at a 10-foot curtain

And drew a ring around the hole to make a bull's eye-certain,

It's best to wait until you're through

And then see where you are:

Deciding goals before you start is riskier by far.

So, if you follow my advice in your evaluation,

You'll start with certainty

And end with self-congratulation.

SOURCE: McIntyre (1976:39).

prowar demonstrators got into a scuffle, so police began making arrests indiscriminately. When one of the prowar demonstrators was apprehended, he began yelling, "You've got the wrong person. I'm *anti-communist!*" To which the police officer replied, "I don't care what kind of communist you are, you're going to jail."

Well, I don't care what kind of evaluator you are, to be effective you need the flexibility to evaluate with or without goals. *The utilization-focused evaluation issue is what information is needed by primary intended uses, not whether goals are clear, specific, and measurable.* Let's consider, then, some other alternatives to traditional goals-based evaluation.

Developmental Evaluation

The only man who behaves sensibly is my tailor; he takes my measurements anew every time he sees me, while all

the rest go on with their old measurements and expect me to fit them.

George Bernard Shaw
(1856–1950)

Developmental Evaluation (DE) is an approach to evaluation in innovative settings where goals are emergent and changing rather than predetermined and fixed. Innovative initiatives are characterized by a state of continuous development and adaptation, and they often unfold within dynamic and unpredictable conditions. DE supports such innovative initiatives by bringing data to bear to inform and guide emergent

choices. I introduced DE in Chapter 4 as one kind of intended use (see Menu 4.1). DE is also listed in Menu 5.1 on *process use* (Chapter 5) as an approach to program and organizational development in which evaluative thinking is infused into and made integral to the development process. In this chapter, we'll look at DE as a major alternative for conceptualizing what evaluation can contribute and how an evaluator can work, an approach informed by insights from complexity science and the particular characteristics of complex, dynamic systems (see Exhibit 8.1).

I originally conceptualized DE as an alternative to formative and summative evaluation (Patton 2005a, 1996, 1994a). The formative-summative distinction was first conceptualized by Scriven (1967) in discussing evaluation of a school curriculum. Summative evaluations were those conducted after completion of the program and for the benefit of some external audience or decision maker to determine whether to continue, expand, or disseminate

EXHIBIT 8.1

Developmental Evaluation Defined

Developmental evaluation supports program and organizational development to guide adaptation to emergent and dynamic realities from a complex systems perspective. Developmental evaluation differs from typical program improvement evaluation (making a program better) in that it involves changing the program model itself as part of innovation and response to changed conditions and understandings. Developmental evaluation doesn't render overall judgments of effectiveness (traditional summative evaluation) because the program never becomes a fixed, static, and stable intervention. Developmental evaluation supports social innovation and adaptive management. Evaluation processes include asking evaluative questions, applying evaluation logic, and gathering real-time data to guide program, product, and/or organizational development. The evaluator is often part of a development team whose members collaborate to conceptualize, design, and test new approaches in a long-term, on-going process of continuous improvement, adaptation, and intentional change. The evaluator's primary function in the team is to infuse team discussions with evaluative questions, data, and thinking to facilitate data-based reflection and decision making in the developmental process.

the program or curriculum (Scriven 1991a, 1991b). Formative evaluations, in contrast, served the purpose of getting ready for summative evaluation by helping work through implementation problems and get the program (or curriculum) sufficiently stabilized to be ready for a summative assessment. Over time, formative evaluation came to refer to any evaluation aimed at improving an intervention or model, but the implication has remained that such improvements are supposed to lead to a stable, fixed model that can be judged as worthy or unworthy of continued funding and dissemination.

But suppose an innovative intervention is being tried out in a highly dynamic environment where those involved are engaged in ongoing trial and error experimentation, figuring out what works, learning lessons, adapting to changed circumstances, working with new participants—and they never expect to arrive at a fix, static, and stable model. They are interested in and committed

to ongoing development. It was precisely this situation that gave rise to DE. I had a 5-year contract with a community leadership program that specified 2½ years of formative evaluation to be followed by 2½ years of summative evaluation. During the formative evaluation, the program made major changes in all aspects of how it operated, from recruitment through program activities and on to follow-up with graduates. At the end of this highly innovative phase of engagement, I pronounced, "From now on, you can't make any more changes in the program because we need it to stay stable so we can conduct the summative evaluation. Only with a fixed intervention, carefully implemented the same for each new group of leaders, can we attribute the measured outcomes to your program intervention in a valid and credible way."

Staff were aghast. They protested, "We don't want to implement a fixed model. In fact, what we've learned is that we need to

keep adapting what we do to the particular needs of new groups. Communities vary. The backgrounds of our participants vary. The economic and political context keeps changing. No. No. No. We can't fix the model. We don't want to do summative evaluation."

Since the purpose of the formative evaluation was to get the program ready for summative evaluation, not doing summative evaluation also meant not doing formative evaluation. It meant doing something else. But what? The answer became *Developmental Evaluation*. DE involved ongoing changes in the program, adapting it to changed circumstances, and altering tactics based on emergent conditions. My two evaluation colleagues and I became part of the design team for the program, which included a sociologist, a couple of psychologists, a communications specialist, some adult educators, a philanthropic funder, and program staff. The design team represented a range of expertise and experiences. Our evaluation role was to bring evaluative thinking and data to bear as the team developed new approaches for new groups, including immigrants, Native Americans, people from distressed rural communities, elected officials, and young people.

The relationship lasted more than 6 years and involved different evaluation designs each year including participant observation, several different surveys, field observations, telephone interviews, case studies of individuals and communities, cost analyses, theory of change conceptualizations, futuring exercises, and training participants to do their own self-evaluations and community-based evaluations. Each year the program changed in significant ways and new evaluation questions emerged. Program goals and strategies evolved. The evaluation evolved. No final report was ever written. The

Summative-Formative-Developmental Evaluation Metaphors

It is said that formative evaluation occurs when the cook tastes the soup and decides if it needs more ingredients or simmering, while summative evaluation occurs when the guests taste the finished soup (Stake, quoted in Scriven 1991b:169). Developmental evaluation occurs when, before cooking, the chef goes to the market to see what vegetables are freshest, what fish has just arrived, and meanders through the market considering possibilities, thinking about who the guests will be, what they were served last time, what the weather is like, and considers how adventurous to be with this meal.

program continues to evolve—and continues to rely on DE.

Complexity Science and Developmental Evaluation

Complexity science offers insights into the changed role that evaluation can usefully play in highly innovative and dynamic circumstances characterized by uncertainty. Studying how living systems organize, adapt, evolve and transform challenges the largely mechanistic models of most programs—and most evaluations. Complexity science reveals that the real world is not a machine. Complex systems are too dynamic, emergent, and, yes, complex, to be reduced to simple cause-effect predictions and controls. I had the opportunity to become part of a Think Tank on Social Innovation that examined the implications of complexity science for social change and evaluation. What we found is that social innovators are driven not by concrete goals but by possibilities, often ill-defined possibilities expressed as values, hopes, and visions. In the early days of innovation,

280 ■ FOCUSING EVALUATIONS: CHOICES, OPTIONS, AND DECISIONS

when ideas about possibilities are just being formed, the innovative process can actually be damaged by forcing too much concreteness and specificity. That's why brainstorming exercises outlaw criticism, because premature critiques and demand for specificity stifle the imagination. Yet what do evaluators typically bring to these situations? Evaluators are trained to insist that hoped-for changes and visions be specified as clear, specific, and measurable goals. That is typically all the evaluator has to offer, the only conceptual tool in the evaluator's toolkit. That's the moral of the story of Alice's encounter with the Cheshire Cat in Wonderland, the sarcastic observation that if you don't know where you're going, any road will get you there.

In the Social Innovations Think Tank, we called that way of thinking "Getting to Yes" where "Yes" represented clear, specific, and measurable outcomes. As we studied real cases of innovation and social transformation, however, we were struck by the liberating effects of open-ended aspiration, belief in possibility, and visionary commitment. Partly tongue in cheek, we came to understand that social transformation begins not

with a plan for "getting to yes" as much as a commitment to undertake a journey aimed at "getting to maybe." That phrase, *Getting to Maybe*, became the title of our book on how the world is changed (Westley, Zimmerman, and Patton 2006). And what does evaluation have to offer people on their journey to *maybe*? One answer is DE.

Getting to Maybe: The Case of Hope

One of the cases we examined in *Getting to Maybe* described the uncertain journey of the Hope Community in Minneapolis. I'm going to include part of that story here with particular attention to evaluation implications *for this situation* and what DE offers under these kinds of circumstances.

In 1977, three Roman Catholic nuns started St. Joseph's House in the inner city of Minneapolis. Over the years, thousands of women and kids found compassionate shelter, dozens of volunteers came to the inner city, women and children who were and had been homeless built a community around St. Joe's hospitality, and the sisters

Incremental Development over Time

Have you ever been in one of those old houses in a small town in a place like Ontario or Iowa that seem to have grown rather than being built? The house begins as a single-room structure, with, of course windows and a door. Then, as the years pass and prosperity increases, a second, much more gracious room is added, connected to the first by an arched door. That seems a much more welcoming way to enter the house, so this room also adds a door. But this is rather grand so when a kitchen is added in the back, the "real" door is built there. Later, a second wing is added and there too there must be a door. Each room has a door; each room is different—a unique space; all rooms are interconnected into one house.

Social innovation is much like that house, with recognizable rooms, each with its own character and each connected to each other through numerous doors. Whatever room you find yourself in, it is helpful to know that the other rooms exist and that you are likely to pass through them, sometimes repeatedly as you engage the demands of social innovation.

SOURCE: From *Getting to Maybe: How the World Is Changed* (Westley, Zimmerman, and Patton 2006:220).

became leaders in fighting against violence and injustice. But by the early 1990s, their environment had changed. The block surrounding St. Joe's had become the center of a crack cocaine epidemic, drug dealers had claimed the streets, and landlords had abandoned many buildings. St. Joe's guests and families living on the block hid their children inside, police regularly ran through the block with guns drawn, and drug dealers and prostitutes (desperate themselves) broke into abandoned buildings. At the north end of the block where two major Minneapolis streets intersected, once-thriving small businesses (a gas station and grocery store) were abandoned.

When Deanna Foster and Mary Keefe took over the leadership of St. Joe's (now Hope Community, Inc.), they brought a vision of a vital, engaged community and decided to attempt a housing revitalization project. In keeping with their community organizing values, they began by trying to talk with local residents. But they found people afraid to talk, afraid of the drug dealers and perpetrators of violence. They decided to start with some concrete changes. They built a playground at their center and renovated a duplex that shared a driveway with the largest drug house. They put fences around the yard and then added fences around the porch because the drug dealers would run through the porch to get around the fences. The drug house was a triplex filled with little children who were terribly neglected and abused. Those little kids were so desperate for something to do that they would climb onto the garage or climb over the fence, anything to try and get to the playground. They cut a hole in the fence and put the tube through it so that the little kids had their own doorway into the playground. The drug dealers would have to embarrass themselves to crawl through the tube, though some did.

One day in desperation they contacted several donors and raised the money to buy that house. They built a duplex where once there had been a crack house.

When Foster and Keefe tell this story, they don't portray the purchase of the drug house as part of some strategic plan. It was an emergent reaction to what they faced on a daily basis.

Based on their early success in ridding the community of one major drug house and their long-term commitment to that area, the leaders and community came together to shape a new vision and found support for that vision when a door suddenly opened. They garnered unexpected support from a major philanthropic donor in the form of a \$500,000 check. Those funds became the core funding for what is, today, a revitalized neighborhood and a Children's Village.

Foster and Keefe know a lot about traditional planning and evaluation approaches, but those approaches didn't fit them, their fluid and dynamic situation, and how they wanted to engage the community.

We almost had to do it, not backwards, but in alternate order. Normally, when an organization gets half a million dollars they have spent a lot of time in a more linear process thinking through what they are going to do. What is the goal? What is the work plan? What will it cost? Who is the staff? You get the community input, all that stuff, and then have this whopping proposal, right? But it didn't happen that way at all. It was "Here's the vision, here's the money, now, make it happen."

And that very absence of a traditional linear planning process became a source of criticism and complaint.

One of the criticisms we get is that we don't have a linear, goal-directed approach. We don't assume where we are going. We ask: Who's here? What are people experiencing?

282 ■ FOCUSING EVALUATIONS: CHOICES, OPTIONS, AND DECISIONS

What are they believing and hoping? What is their understanding of community? And what is our understanding of all the things we've done? We keep trying things, we keep building understanding and building community around ourselves. We are about uncovering, discovering, and creating. It really unfolds itself. It grows organically. It's just such a natural process.

But it's more complex than that because, at the same time, there's a whole set of strategic thinking that's going on. We also have to ask: Where is the land out there? Where's the money? What are the opportunities? Where are the potential partners? What are the potential pitfalls? How could all this fit together? What would happen if we did this? (Quotations from Westley, Zimmerman, and Patton 2006:173–74; see also Foster and Keefe 2004)

Summative evaluation would not be appropriate for Hope Community because there is no model being created for replication. They are learning and generating principles to inform future action, but that is a far cry from a “best practices” model that can be faithfully replicated in one community after another. The questions Foster and Keefe were asking also differed from familiar formative evaluations, which are focused on establishing programs' strengths and weaknesses, and progress, relative to intended outcomes, as the program unfolds. Instead, Hope's leadership pursued an open-ended approach to data gathering, where the questions and concerns were emergent, and where trial and error experiences were continuously mined for learning.

The Hope Community reality was messy, not orderly; emergent, not controlled; and social innovation was an iterative process of experimentation, learning, and adaptation. The Hope Community leadership was immersed in a complexity perspective. They monitored both the big

picture and the whole picture—national housing, community development, and real estate patterns; interest rates and international finance; government policies, philanthropic funding trends and priorities; research on community revitalization. They had a keen sense of the history of the community. And, at the same time, they were fully enmeshed in the day-to-day reality of work in the community, engaging with residents as well as local government inspectors, city planners, social service agencies working in the community, local businesses, and local funders.

In these kinds of complex situations of rapid change, ongoing adaptation, and shifting priorities, and with these kinds of social innovators who eschew a command-and-control approach to change in favor of engagement and emergence, DE offers a way to infuse systematic evaluative thinking and real-time data into the generative processes of change.

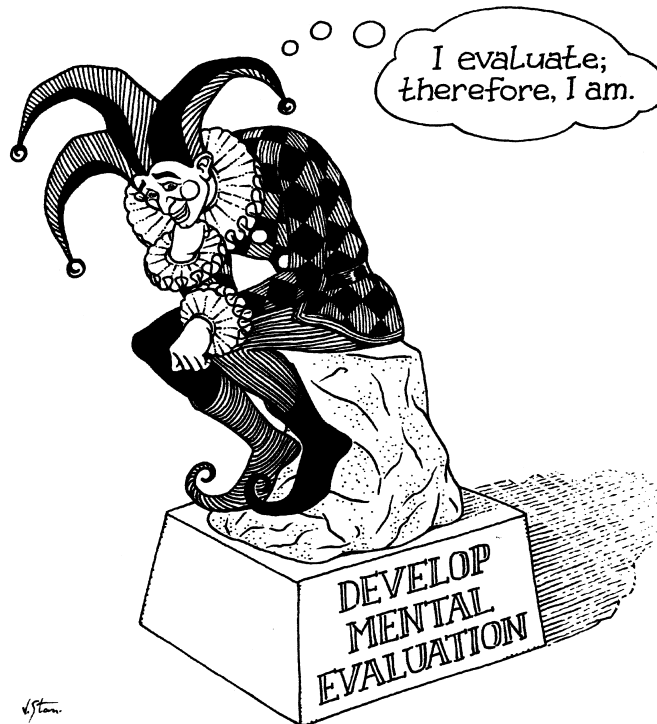
Understanding Evaluation Niches

Utilization-focused evaluation aims to adapt evaluation to the needs of particular information users and decision makers within the specific set of circumstances they face. As the Hope Community case illustrates, one category of primary intended user consists of social innovators who make up what they're doing as they go along. They are engaged in what management guru Tom Peters (1996) advocated in his book *Liberation Management* as

Ready. Fire. Aim.

Instead of Ready. Aim. Aim. Aim. . . .

This runs counter to the conventional wisdom that extensive planning (aiming) should precede action. But planning only



works where you have control and know what the critical factors are. Under conditions of high innovation, uncertainty rules and control freaks perish. Indeed, one of the advantages of *Ready, Fire, Aim* is, paradoxically, its high and rapid failure rate, facilitating fast learning and speedy moving on (Shirky 2007). Traditionally, goal-based evaluation operates under conditions of *Ready. Aim. Fire*. Then, the evaluator determines whether the target was hit. But what is the role for evaluation when the innovator's mode of operating is *Ready. Fire. Aim*. The developmental evaluator still figures out what was hit (if anything), but the analysis is not a comparison of what was hit to a preconceived target. In providing feedback about what the innovator has "hit" (what immediate outcomes are emerging), the developmental evaluator engages the innovator in the following kind of dialogue: What's your

reaction to what you've hit so far? And what you've missed? What does this "hit" tell you? How does what you've done so far align with your values and vision? What does this "hit" (or "miss") suggest about what to do next?

Let me give a concrete example. Some years ago I was invited to join a design team to bring an evaluative perspective to their vision of creating an innovative program aimed at helping chronically unemployed and disadvantaged men of color get living wage, sustainable jobs in established companies with good benefits. The team had been working on a comprehensive plan for a pilot effort. I had just completed a cluster evaluation of 34 projects aimed at supporting families in poverty improve their lives. All those projects had been carefully planned by savvy advocates and experienced community development professionals, usually with participation of people in

284 ■ FOCUSING EVALUATIONS: CHOICES, OPTIONS, AND DECISIONS

The *Ontario Science Centre* in Toronto asked the question: What if Canada could become a world leader in innovation? From this question, an idea was born. What if a science centre could reconceptualize what it meant to be a visitor? What if visitors were participants whose experience would engage visitors directly in scientific experimentation and the gathering of data to take on problems with real-world applications? What if, as in actual engineering and science, participants could guide their research activities without certainty of the results, and have leeway to innovate in their approaches?

These questions led to the development of the Agents of Change initiative. A bold and creative experiment aimed at fostering the development of visitor's thinking about innovation, risk, collaboration, and creativity. What would it take to do this? How would the science centre need to think differently?

Developmental evaluation supported this process as the Agents of Change initiative moved forward with an ambitious time frame. Planning, acting, and adapting were simultaneous as new elements were designed they were immediately tested on the floor and, with rapid observation and feedback, modifications were made daily. At the same time, the developmental evaluation helped in shaping the goal of fostering innovation in visitors.

SOURCE: Gamble (2007).

the communities they were serving. They had written detailed proposals as required by their philanthropic funder. *And not one of those 34 projects had unfolded as planned.* All experienced serious implementation difficulties and most had to substantially revise what they had planned to do, significantly adapting their model to the realities of the people in poverty with whom they were working. The funder supported, indeed, encouraged, those changes despite the delays and mishaps involved. Based on those findings I told design team I had just joined, "It won't make much difference what you plan, it won't be right, so just start doing it and make corrections as you go." What really mattered, I suggested, was experience on a small-scale with constant reengineering.

That innovative program, called *Twin Cities Rise!*, has now been operating for over a decade. Along the way, virtually every aspect of what they do has changed substantially. Their recruitment and orientation processes have developed—not just improved, but fundamentally changed. (They began with an in-depth selection

process trying to select for success. When that didn't work, they went to a more open recruitment process followed by a probation period before full, contractual enrollment in the program.) Their target populations have changed (more immigrants and women), partly in response to changed welfare to work legislation. The political and economic environment has changed. Their objectives and immediate outcomes have changed as they came to better understand what prospective employers wanted. Their "model" today looks nothing like the original "model" they had in mind, though *the fundamental principles and values that led to the program have not changed at all.* It is a values-driven program characterized by constant adaptation. What they disseminate are principles, values, and lessons, not a fixed model of specific practices (Rothschild forthcoming).

In our Social Innovations Think Tank meetings, we examined cases like the Hope Community and *Twin Cities Rise!* to learn about how change had occurred in those cases and to consider evaluation

implications. Inevitably, questions arose: How solid are the data on results and impacts? Can the causal chain between intervention and outcomes be substantiated or even traced? What things didn't work along the way, and how did those involved learn from failure as well as success? How

do complexity science concepts illuminate these change processes? We also looked at what the profession of evaluation had to offer in light of complexity theory. That led to a further refinement of DE and the contrasts between traditional evaluation and DE listed in Exhibit 8.2.

EXHIBIT 8.2

Evaluation Niche Contrasts

<i>Traditional Summative Evaluations</i>	<i>Developmental Evaluations</i>
Measure success against predetermined goals	Develop new measures and monitoring mechanisms as goals emerge and evolve
Based on linear cause-effect modeling	Based on complex systems thinking, nonlinear, emergent dynamics, and interdependent interconnections
Render judgments of success or failure	Provide rapid feedback, generate learnings, support direction, or affirm changes in direction
Position the evaluator outside to assure independence and objectivity	Position the evaluator as a design team member integrated into developmental decision making
Aim to produce generalizable findings so that effective practices can be applied to elsewhere	Aim to produce context-specific understandings that inform ongoing innovation; innovative principles are generalizable.
Accountability focused on and directed to external authorities and funders	Accountability centered on the innovators' deep sense of fundamental values and commitments
Accountability aimed at control and locating source of failures	Learning to respond to lack of control and staying in touch with what's unfolding and thereby responding strategically
Evaluation often a compliance function delegated down in the organization.	Evaluation a leadership function: reality-testing, results-focused, learning-oriented leadership
Evaluator determines the design based on the evaluator's perspective about what is important. The evaluator controls the evaluation	Evaluator collaborates with those engaged in the change effort to design an evaluation process that matches the innovation philosophically and organizationally
Evaluation focuses on bottom line success or failure	Evaluation focuses on learning and adaptation

The Role of the Developmental Evaluator

In Hemmingway's *The Sun Also Rises*, the main character is asked how he went bankrupt. He replies in what has become a famous literary line, "Gradually, then suddenly." Developmental evaluators help monitor what is happening gradually to help anticipate sudden changes.

In DE, the evaluator is incorporated into the program or organizational development decision-making process because those involved value the logic and conceptual rigor of evaluation thought and engagement with data. Moreover, experienced evaluators have accumulated knowledge about patterns of effective programming that can inform options and facilitate discussion of the possible implications of actions as they are considered. My role as developmental evaluator has been to ask evaluative questions of the innovators and hold their feet to the fire of reality testing. Evaluation data are collected and interpreted as part of the feedback process, to be sure, but quite above and beyond the use of findings, these development-oriented decision makers want to have their ideas examined in the glaring light of evaluation logic. *Honing ideas on the whetstone of evaluative thinking is an example of process use.*

Keep in mind we are talking about working with social innovators here: action-oriented, change-obsessed, push-the-envelope, do-it-now people. As I noted in Chapter 4 when I first introduced DE on the menu of possible uses, many such innovators eschew clear, specific, and measurable goals up-front because clarity, specificity, and measurability are limiting. They've identified an issue or problem about which they are passionate and they want to explore potential solutions or interventions.

They realize that where they end up will be different from what they imagined in the beginning. Where innovative programming is involved, they expect different participants will want different outcomes (an individualized approach as a matter of principle) and that participants themselves should play a major role in setting individualized goals for themselves. This process often includes elements of participatory evaluation, for example, engaging staff and participants in setting personal goals and monitoring goal attainment, but those goals aren't fixed—they're milestones for assessing progress, subject to change as learning occurs. It is in this respect that the primary purpose is program and organizational *development*. As the evaluation unfolds, program designers observe where they end up and make adjustments based on dialogue about what's possible and what's desirable, though the criteria for what's "desirable" may be quite situational and always subject to change.

Social Innovators and Developmental Evaluation

Developmental evaluation is especially appropriate for social innovators who don't value traditional characteristics of summative excellence, such as standardization of inputs, consistency of treatment, uniformity of outcomes, and clarity of causal linkages. They assume a world of multiple causes, diversity of outcomes, inconsistency of interventions, interactive effects at every level—and they find such a world exciting and desirable. They never expect to conduct a summative evaluation because they don't expect the program—or world—to hold still long enough for summative review. They expect to be forever developing and changing—and they want an evaluation approach that supports development and change.

Development-focused relationships can go on for years and, in many cases, never involve formal, written reports. Here are some examples from my own evaluation consulting practice.

Three Examples of Developmental Evaluation

1. *Supporting Diversity in Schools.* A group of foundations agreed to support multicultural education in the Saint Paul Public Schools for 10 or more years. Community members identified the problem as low levels of success for children of color on virtually every indicator they examined, e.g., attendance, test scores, and graduation. The “solution” called for a high degree of community engagement, especially by people of color, in partnering with schools. The nature of the partnering and interim outcomes were to emerge from the process. Indeed, it would have been “disempowering” to local communities to predetermine the desired strategies and outcomes prior to their involvement. Moreover, different communities of color—African Americans, Native Americans, Hispanics, and Southeast Asians—could be expected to have varying needs, set differing goals, and work with the schools in different ways. All these things had to be *developed*.

The evaluation-documented developments, provided feedback at various levels from local communities to the overall district, and facilitated the process of community people and school people coming together to develop evaluative criteria and outcome claims. Both the program design and evaluation changed at least annually, sometimes more often. In the design process, lines between participation, programming, and evaluation were ignored as everyone worked together to develop the

program. The evaluation reports took the form of multiple voices presenting multiple perspectives. These voices and perspectives were facilitated and organized by the evaluation team, but the evaluator’s voice was simply one among many. No summative evaluation was planned or deemed appropriate though a great deal of effort went into publicly communicating the developmental processes and outcomes (see Exhibit 8.3).

2. *Children’s and Families Community Initiative.* A local foundation made a 20-year commitment to work with two inner city neighborhoods to support a healthier environment for children and families. The communities are poor and populated by people of diverse ethnic and racial backgrounds. The heart of the commitment was to provide funds for people in the community to set their own goals and fund projects they deemed worthwhile. A community-based steering committee became, in effect, a decision-making group for small community grants. Grant-making criteria, desired outcomes, and evaluation criteria all had to be developed by the local community. The purpose of the developmental process was to support internal, community-based accountability (as opposed to external judgment by the affluent and distant board of the sponsoring foundation). My role, then, was facilitating sessions with local community leaders to support their developing their own evaluation process and sense of shared accountability. The evaluation process had to be highly flexible and responsive. Aspects of participatory and empowerment evaluation also were incorporated. Taking a 20-year developmental perspective, where the locus of accountability is community based rather than funder based, changed all the usual parameters of evaluation.

EXHIBIT 8.3

Reflective Practice and Developmental Evaluation

For several years, I facilitated a monthly reflective practice process with innovative staff in a suburban adult and community education program. No specific problems or goals were addressed. Instead, they were committed to ongoing program development and organizational change. That meant going wherever their inquiries took them.

They met monthly to share their action research observations for the last month. Their observations focused on whatever issue the group had chosen the previous month. The reflective practice process involved

1. Identifying an issue, interest, or concern
2. Agreeing to try something
3. Agreeing to observe some things about what was tried
4. Reporting back to the group their individual observations with detailed descriptions
5. Identifying patterns of experience or themes across the separate reports (facilitated by the developmental evaluator)
6. Deciding what to try next, i.e., determining the action implications of the findings, and
7. Repeating the process with the new commitment to action

Over several years, this process supported major curricular and organizational change. Evaluation was ongoing and feedback was immediate. The process combined staff and organizational development with evaluative thinking and facilitated reflection. My role, as facilitator, was to keep them focused on data-based observations and help them interpret and apply findings. There were no formal reports and no formative or summative judgments in the usual evaluation sense. Instead, they engaged in an ongoing developmental process of incremental change, informed by data and judgment, which led to significant cumulative evolution of the entire program. They became a learning organization.

3. *Wilderness Education for College Administrators.* In evaluating a wilderness education program, my evaluation partner and I engaged in participant observation and provided daily feedback to program staff about issues that surfaced in our interviews and observations. Over the course of a year involving three 10-day wilderness experiences, staff used our feedback to shape the program, not just in the formative sense of improvement, but in a

developmental way, actually conceptualizing and designing the program as it unfolded. The two leaders of the program expected different participants to take away different things and didn't have pre-set goals or outcomes in mind. Indeed, they wanted to find out what diverse outcomes emerged for those involved and which experiences seemed to support various outcomes for participants. We became part of the decision-making team that

conceptualized the program. Our evaluative questions, quite apart from the data we gathered and fed back, helped shape the program.

An example will illustrate our developmental role. Early in the first trip in the Gila Wilderness of New Mexico, we focused staff attention on our observation that participants were struggling with the transition from city to wilderness. After considerable discussion and input from participants, staff decided to have evening discussions on this issue. Out of those discussions a group exercise evolved in which, each morning and evening, everyone threw their arms about, shook their legs, and tossed their heads in a symbolic act of casting off the toxins that had surfaced from hidden places deep inside. The fresh air, beauty, quiet, fellowship, periods of solitude, and physical activity combined to “squeeze out the urban poisons.” Participants left the wilderness feeling cleaner and purer than they had felt in years. They called that being “detoxified.” Like the drunk who is finally sober, they took their leave from the wilderness committed to staying clear of the toxins.

No one, however, was prepared for the speed of *retoxification*. Follow-up interviews revealed that participants were struggling with reentry. As evaluators, we worked with staff to decide how to support participants in dealing with reentry problems. When participants came back together 3 months later, in the Kofa Mountains of Arizona, they came with the knowledge that detox faded quickly and enduring “purification” couldn’t be expected. Then, the wilderness again salved them with its cleansing power. Most left the second trip more determined than ever to resist retoxification on reentering their urban environments, but the higher expectations only made the subsequent falls more distressing. Many came to the third trip skeptical and resistant.

It didn’t matter. The San Juan River in Utah didn’t care whether participants embraced or resisted it. After 10 days rowing and floating, participants, staff, *and* evaluators abandoned talking about “detox” as an absolute state. We came to understand it as a matter of degree and a process: an ongoing struggle to monitor the “poisons” around us, observe carefully their effects on our minds and bodies, and have the good sense to get to the wilderness when being poisoned started to feel normal. This understanding became part of the program model developed jointly by participants, staff, and evaluators. As evaluators, we led the discussions and pushed for conceptual clarity beyond what staff and participants would likely have been able to do without an evaluation perspective.

Cautions About Developmental Evaluation

It will be clear to the reader, I trust, that my evaluation role in each of the programs just reviewed involved a degree of engagement that went beyond the independent data collection and assessment that have traditionally defined evaluation functions. Lines between evaluation and development became blurred as we worked together collaboratively in teams. I have found these relationships to be substantially different from the more traditional evaluations I conducted earlier in my practice. My role has become more *developmental*.

But once again, a note of caution about language. The term *development* carries negative connotations in some settings. Miller (1981), in *The Book of Jargon*, defines development as “a vague term used to euphemize large periods of time in which nothing happens” (p. 208). Evaluators are well-advised to be attentive to what specific words mean in a particular context to

290 ■ FOCUSING EVALUATIONS: CHOICES, OPTIONS, AND DECISIONS

specific intended users—and to choose their terms accordingly.

One reaction I've had from colleagues is that the examples I've shared above aren't "evaluations" at all but rather organizational development efforts. I won't quarrel with that. There are sound arguments for defining evaluation narrowly to distinguish genuinely evaluative efforts from other kinds of organizational mucking around. But, in each of the examples I've shared, and there are many others, *my participation, identity, and role were considered evaluative by those with whom I was engaged (and by whom I was paid)*. There was no pretense of external independence. My role varied from being evaluation facilitator to full team member. In no case was my role primarily *external* reporting and accountability. When reporting to funders, my developmental role and its implications were made clear.

DE certainly involves a role beyond being solely an evaluator, but I include it among the things we evaluators can do

because program and organizational development are legitimate uses of evaluation processes. What we lose in conceptual clarity and purity with regard to a narrow definition of evaluation (independently judging merit or worth), we gain in appreciation for evaluation expertise. When Scriven (1995) cautions against crossing the line from rendering judgments to offering advice, I think he underestimates the valuable role evaluators can play in design and program improvement based on cumulative knowledge. Part of my value to a design team is that I bring a reservoir of knowledge (based on many years of practice and having read a great many evaluation reports) about what kinds of things tend to work and where to anticipate problems. Young and novice evaluators may be well-advised to stick fairly close to the data. However, experienced evaluators have typically accumulated a great deal of knowledge and wisdom about what works and what doesn't work. More generally, as

Is Developmental Evaluation Really Evaluation?

Developmental evaluation is fundamentally evaluative because it focuses on

Identifying/negotiating/deciding *evaluative criteria*—knowing that our understanding of these will change over time and new criteria may emerge as things change (e.g., new challenges, opportunities, imperatives, such as global warming) and that different people and groups will have different views on what these should be.

Identifying/negotiating/deciding *standards of performance*—knowing that our understanding of these will change over time and the standards may also need to change as things change and that different people and groups will have different views on what these should be.

Identifying/negotiating/deciding what would constitute *credible evidence of performance* and getting it—knowing that our understanding of the best way to do this within ubiquitous limitations will change, and that different people and groups will have different views on what these should be.

Identifying/negotiating/deciding *methods for synthesis of performance information*—knowing that our understanding of the best way to do this will change and the weightings may need to change as things change and that different people and groups will have different views on what these should be.

Patricia J. Rogers (2007b), Professor in Public Sector Evaluation, Founder of CIRCLE (Collaborative Institute for Research, Consulting and Learning in Evaluation), Royal Melbourne Institute of Technology, Australia.

a profession, we know a lot about patterns of effectiveness—and will know more over time. For example, we know that new initiatives will experience implementation problems, that original program (and evaluation) designs will need to be adapted to real, on-the-ground realities, and that the excitement of new, innovative efforts creates halo effects that cannot be sustained over time. That knowledge makes us valuable partners in the design process. Crossing that line, however, can reduce independence of judgment. The costs and benefits of such a role change must be openly acknowledged and carefully assessed with primary intended users and evaluation funders.

Balancing Critical and Creative Thinking

Evaluation requires critical thinking. Development involves creative thinking. These two types of thinking are often seen as mutually exclusive. *Developmental Evaluation* is about holding them in balance. What developmental evaluation does is bring the rigor of evaluation (evidence-based, reality-testing questioning) together with organizational development coaching (change oriented, relational, visionary).

A Menu Approach to Focusing Evaluations

DE offers an approach appropriate for a particular set of contingencies, one where social innovators (the primary intended users) want to use both evaluation processes and findings to support *development* (the primary intended use of the evaluation). Exhibit 8.4 provides a checklist of the contingency variables and situational factors for which DE is especially appropriate. Historically,

lacking the DE option, evaluators have tried to force such situations into formative or summative boxes, often, I believe, constraining or even doing damage to the very process of innovation they were meant to inform.

Evaluation is not benign. Like any powerful tool that is misused, the wrong evaluation approach can do harm despite the intention to do good. Likewise, it would be inappropriate to impose DE on a situation where primary intended users want a rigorous answer to the summative question of whether a specific, well-defined model should be disseminated; under those conditions the evaluation judges whether evidence supports the proposition that the intervention can reliably and consistently produce desired and prescribed outcomes. The challenge, then, is to match the evaluation to the situation, which in utilization-focused evaluation, is determined by the information needs and intended evaluation uses of primary intended users.

There are a variety of ways of focusing evaluations. The transdiscipline of evaluation has become a many-splendored thing, rich with options, alternatives, models, and approaches (e.g., Stufflebeam and Shinkfield 2007). Menu 8.1 at the end of this chapter offers an extensive list of alternative ways of focusing an evaluation. I'll elaborate on only a few of these here. I'm highlighting here alternatives to the traditional goal-based approach to evaluation. These options engage intended users in other ways, always with the purpose of providing useful findings to inform actions, decisions, and understandings.

Focusing on Future Decisions. An evaluation can be focused on information needed to inform future decisions. Proponents and opponents of school busing for

EXHIBIT 8.4

Developmental Evaluation (DE) Checklist Ten Situational Contingencies Indicating DE Would Be Appropriate

1. Situation is characterized by systems complexity: multiple interacting variables and factors interacting dynamically, interdependently, and unpredictably.
2. Working with innovators who are guided by strong values and vision, want to tackle a problem or issue, but aren't yet sure what needs to be done or what specific outcomes they are aiming at, in part because the situation is so complex. (These innovators are the primary intended users of the evaluation.)
3. Innovators want to *develop a solution* through experimentation, trial and error, and seeing what responses they get to what they try. *Development* is the intended use for the evaluation process and findings.
4. Solutions are expected to emerge from engagement and action (not advance planning).
5. The innovation is being tried in a highly dynamic environment, subject to rapid and unpredictable changes and demands.
6. Uncertainty abounds. There is little agreement among people about what should be done. There is little knowledge about the real nature of the problem or what potential interventions will yield.
7. There is a high likelihood of unanticipated *and* unanticipatable consequences.
8. Innovators are open to, indeed, *want to* use evaluative questioning and data to inform their understandings about what is happening and guide their next steps on the indeterminate journey. They can deal with critical questioning and data-based feedback about their creative impulses and "let's try it out and see what happens" approach to change.
9. Developmental evaluator has high tolerance for ambiguity, can react and adapt quickly, and communicate effectively with hyperactive, short-attention-span, action-oriented innovators. (Not all innovators are like that, but it's helpful to be prepared. Developmental evaluation is not an academic exercise. The evaluator will often be engaged in the trenches where the action is happening while it's happening.)
10. Those funding the innovation and evaluation understand the unique niche, constraints, and deliverables of DE.

desegregation may never agree on educational goals, but they may well agree on what information is needed to inform future debate, for example, data about who is bused, at what distances, from what neighborhoods, and with what effects.

Focusing on Critical Issues or Concerns. When the Minnesota Legislature first initiated

Early Childhood Family Education programs, some legislators were concerned about what advice was being given to parents. The evaluation focused on this issue, and the evaluators became the eyes and ears for the Legislature and general public at a time of conflict about "family values" and anxiety about values indoctrination. The evaluation, based on descriptions of what actually occurred and

data on parent reactions, helped put this issue to rest. Over time new issues arose. For example, universal access became a matter of contentious debate. Should the program be targeted to low-income parents or continue to be available to all parents, regardless of income? What are the effects on parents of a program that integrates people of different socioeconomic backgrounds? An evaluation was commissioned to inform that policy debate and examine programming implications (Mueller 1996). These Early Childhood and Parent Education program evaluations, done for the State Legislature, were issue based more than goals based, although attention to differential parent outcomes was subsumed within the issues.

The “Responsive Approach” to Evaluation. Stake (1975) advocates incorporating into an evaluation the various points of view of constituency groups under the assumption that “each of the groups associated with a program understands and experiences it differently and has a valid perspective” (Stecher and Davis 1987:56–57). The focus, then, is on informing each group of the perspective of other groups and providing data on each group’s goals.

Focusing on Questions. In Chapter 2, I described focusing an evaluation in Canada in which primary intended users generated questions that they wanted answered—without regard to methods, measurement, design, resources, precision, goals—just 10 basic questions, real questions that they considered important. After working individually and in small groups, we pulled back together and generated a single list of 10 basic evaluation questions—answers to which, they agreed, could make a real difference to the operations of the school

division. 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. Generating a list of real and meaningful evaluation questions played a critical part in getting things started. Exhibit 2.4 in Chapter 2 offers criteria for good utilization-focused questions.

It is worth noting that formulating an appropriate and meaningful question involves considerable skill and insight. In her novel, *The Left Hand of Darkness*, science fiction author Ursula K. Le Guin (1969) reminds us that questions and answers are precious resources, not to be squandered or treated casually. She shows us that how one poses a question frames the answer one gets—and its utility. In the novel, the character Herbor makes an arduous journey to fortune-tellers who convene rarely and, when they do, permit the asking of only a single question. His mate is obsessed with death, so Herbor asks them how long his mate will live. Herbor returns home to tell his mate the answer that Herbor will die before his mate. His mate is enraged:

You fool! You had a question of the Foretellers, and did not ask them when I am to die, what day, month, year, how many days are left to me—you asked how long? Oh you fool, you staring fool, longer than you, yes, longer than you!

And with that his mate struck him with a great stone and killed him, fulfilling the prophecy and driving the mate into madness. (Pp. 45–46)

Testing Assumptions. The Greek Stoic philosopher Epictetus observed, “It is impossible for a man to learn what he thinks he

294 ■ FOCUSING EVALUATIONS: CHOICES, OPTIONS, AND DECISIONS

already knows.” With a group that has some trust and is willing to dig deeply into tougher issues, the evaluation can draw on organizational development and action research techniques for questioning assumptions (Dick and Dalmau 1999) and surfacing the “undiscussables”—what is sometimes called naming the elephant in the organization (Hammond and Mayfield 2004). Much of evaluation is framed as finding out what is not known or filling the knowledge gap. But deeper problems go beyond what is not known to what is known but not true (false assumptions) or known to be untrue, at least by some, but not openly talked about (undiscussable). In doing a cluster evaluation for a group of antipoverty programs, the undiscussable was that the staff was almost all white while the clients were virtually all African American. The unexamined assumptions were that there weren’t enough “qualified” black staff and that clients didn’t care about the race of staff anyway, so it wasn’t really an issue or factor. In fact, racism was an undiscussable. It wasn’t until the third year of the evaluation, after trust had been built, some appreciation of evaluative thinking had been established, and those involved were ready to dig more deeply into tougher issues that the group moved inquiry into the effects of racial differences to the top of the list of issues for evaluative inquiry.

When I offer a group the option of testing assumptions and opening up undiscussable, I ask them if they’re ready to take on the evaluation challenge of American humorist Mark Twain who famously observed: “It ain’t what you don’t know that gets you into trouble. It’s what you know for sure that just ain’t so.” This is hard core reality-testing. I’m not inclined to start there with a group.

I find it’s better to start a new group with less threatening issues and build capacity for evaluative inquiry before taking on more challenging and threatening questions.

A “Seat-of-the-Pants” Approach. In our follow-up study of how federal health evaluations were used, we came across a case example of using issues and questions to focus an evaluation. The decision makers in that process, for lack of a better term, called how they focused the evaluation a “seat-of-the-pants” approach. I would call it focusing on critical issues. The results influenced major decisions about the national Hill-Burton Hospital Construction Program. This evaluation illustrates some key characteristics of utilization-focused evaluation.

The evaluation was mandated in federal legislation. The director of the national Hill-Burton program established a permanent committee on evaluation to make decisions about how to spend evaluation funds. The committee included representatives from various branches and services in the division: people from the state Hill-Burton agencies, the Comprehensive Health Planning agencies, and the health care industry, and regional Hill-Burton people. The committee met at regular intervals to “kick around” evaluation ideas. Everyone was free to make suggestions. Said the director, “If the committee thought a suggestion was worthwhile, we would usually give the person that suggested it an opportunity to work it up in a little more detail” [DM159:3]. The program officer commented that the final report *looked* systematic and goals based, but

that’s not the kind of thinking we were actually doing at that time . . . We got started by brainstorming: “Well, we can look at the funding formula and evaluate it.” And someone said, “Well, we can also see what state agencies are doing.” See? And it was this kind of seat-of-the-pants approach. That’s the way we got into it. [PO159:4]

The evaluation committee members were carefully selected on the basis of their knowledge of central program issues. While this

was essentially an internal evaluation, the committee also made use of outside experts. The director reported that the committee was the key to the evaluation's use: "I think the makeup of the committee was such that it helped this study command quite a lot of attention from the state agencies and among the federal people concerned" [DM159:18].

Program Director: Well, I think this was heavily focused toward the major aspects of the program that the group was concerned about.

Interviewer: Did the fact that you focused on major aspects of the program make a difference in how the study was used?

Director: It made a difference in the interest with which it was viewed by people. . . . I think if we hadn't done that, if the committee hadn't been told to go ahead and proceed in that order, and given the freedom to do that, the committee itself would have lost interest. The fact that they felt that they were going to be allowed to pretty well free-wheel and probe into the most important things *as they saw them*, I think that had a lot to do with the enthusiasm with which they approached the task. [DM159:22]

The primary intended users began by brainstorming issues ("seat-of-the-pants approach") but eventually framed the evaluation question in the context of major policy concerns that included, but were not limited to, goal attainment. They negotiated back and forth—until they determined and agreed on the most relevant focus for the evaluation.

Changing Focus over Time: Stage Models of Evaluation

Evaluate no program until it is proud.

—Donald Campbell (1983)

Important to focusing an evaluation can be matching the evaluation to the program's stage of development, what Tripodi, Felin, and Epstein (1971) called *differential evaluation*. Evaluation priorities can vary at

Here, then, we have a case example of the first two steps in utilization-focused evaluation: (1) identifying and organizing primary intended users of the evaluation and (2) focusing the evaluation on their interests and what they believe will be useful. And how do you keep a group like this working together?

the *initiation stage* (when resources are being sought), the *contact stage* (when the program is just getting under way), and the full implementation stage.

In a similar vein, Jacobs (1988) has conceptualized a "five-tier" approach: (1) the preimplementation tier focused on needs assessment and design issues; (2) the accountability tier to document basic functioning to funders; (3) the program clarification tier focused on improvement and feedback to staff; (4) the "progress toward objectives" tier, focused on immediate, short-term outcomes and differential effectiveness among clients; and (5) the "program impact" tier, which focuses on overall judgments of effectiveness, knowledge about what works, and model specification for replication.

The logic of these stage models of evaluation is that, not only do the questions evolve as a program develops, but the stakes go up. When a program begins, all kinds of things

296 ■ FOCUSING EVALUATIONS: CHOICES, OPTIONS, AND DECISIONS

can go wrong, and, as we'll see in the next chapter on implementation evaluation, all kinds of things typically do go wrong. It is rare that a program unfolds as planned. Before committing major resources to overall effectiveness evaluation, then, a stage model begins by making sure the groundwork was carefully laid during the needs assessment phase; then basic implementation issues are examined and formative evaluation for improvement becomes the focus; if the early results are promising, then *and only then*, are the stakes raised by conducting rigorous summative evaluation. It was to this kind of staging of evaluation that Donald Campbell (1983), one of the most distinguished social scientists of the twentieth century, was referring when he implored that no program should be evaluated before it is "proud." Only when program staff have reached a point where they and others close to the program believe that they're on to something, "something special that we know works here and we think others ought to borrow," should rigorous summative evaluation be done to assess the program's overall merit and worth (Schorr 1988:269–270).

An example may help clarify why it's so important to take into account a program's stage of development. The Minnesota State Department of Education funded a "human liberation" course in the Minneapolis public schools aimed at enhancing communication skills around issues of sexism and racism. Funding was guaranteed for 3 years, but a renewal application with evaluation findings had to be filed each year. To ensure rigorous evaluation, an external, out-of-state evaluator was hired. When the evaluator arrived on the scene, virtually everything about the program was uncertain: curriculum content, student reaction, staffing, funding, relationship to the school system, and parent support. The evaluator insisted on beginning at what Jacobs (1988) called the fourth of

five tiers: assessing progress toward objectives. He forced staff, who were just beginning course development (so they were at the initiation or preimplementation stage, tier one) to articulate clear, specific, and measurable goals in behavioral terms. The staff had no previous experience writing behavioral objectives, nor was program conceptualization sufficiently advanced to concretize goals, so the evaluator formulated the objectives for the evaluation.

To the evaluator, the program seemed chaotic. How can a program operate if it doesn't know where it's going? How can it be evaluated if there are no operational objectives? His first-year evaluation rendered a negative judgment with special emphasis on what he perceived as the staff's failure to seriously attend to the behavioral objectives he had formulated. The teaching staff reacted by dismissing the evaluation as irrelevant. State education officials were also disappointed because they understood the problems of first-year programs and found the evaluation flawed in failing to help staff deal with those problems. The program staff refused to work with the same evaluator the second year and faced the prospect of a new evaluator with suspicion and hostility.

When a colleague and I became involved the second year, the staff made it clear that they wanted nothing to do with behavioral objectives. The funders and school officials agreed to a DE with staff as primary users. The evaluation focused on the staff's need for information to inform ongoing, adaptive decisions aimed at program development. This meant confidential interviews with students about strengths and weaknesses of the course, observations of classes to describe interracial dynamics and student reactions, and beginning work on measures of racism and sexism. On this latter point, program staff were undecided as to whether

they were really trying to change student attitudes and behaviors or just make students more “aware.” They needed time and feedback to work out satisfactory approaches to the problems of racism and sexism.

By the third year, uncertainties about student reaction and school system support had been reduced by the evaluation. Initial findings indicated support for the program. Staff had become more confident and experienced. They decided to focus on instruments to measure student changes. They were ready to deal with program outcomes as long as they were viewed as experimental and flexible.

The results of the third-year evaluation showed that students’ attitudes became more racist and sexist because the course experience inadvertently reinforced students’ prejudices and stereotypes. Because they helped design and administer the tests used, teachers accepted the negative findings. They abandoned the existing curriculum and initiated a whole new approach to dealing with the issues involved. By working back and forth between specific information needs contextual goals and focused evaluation questions, it was possible to conduct an evaluation that was used for continuous development of the program. The key to use was matching the evaluation to the program’s stage of development and the information needs of designated users as those needs changed over time.

Focusing an Evaluation

Focusing an evaluation is an interactive process between evaluators and the primary intended users of the evaluation. It can be a difficult process because deciding what will be evaluated means deciding what will not be evaluated. Programs are so complex and have so many levels,

goals, and functions that there are always more potential study foci than there are resources to examine them. Moreover, as human beings, we have a limited capacity to take in data and juggle complexities. We can deal effectively with only so much at one time. The alternatives have to be narrowed and decisions made about which way to go. That’s why I’ve emphasized the menu metaphor throughout this book. The utilization-focused evaluation facilitator is a chef offering a rich variety of choices, from full seven-course feasts to fast-food preparation (*but never junk*). The stage approach to evaluation involves figuring out whether, in the life of the program, it’s time for breakfast, lunch, a snack, a light dinner, or a full banquet.

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 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. And evaluators have trouble getting intended users to focus evaluation issues.

Focusing an evaluation means dealing with several basic concerns. What is the purpose of the evaluation? How will the information be used? What will we know after the evaluation that we don’t know now? What actions will we be able to take based on evaluation findings? These are not simply rote questions answered once and then put aside. The utilization-focused evaluator keeps these questions front and center throughout the design process. The answers to these and related questions will

298 ■ FOCUSING EVALUATIONS: CHOICES, OPTIONS, AND DECISIONS

determine everything else that happens in the evaluation. As evaluators and primary users interact around these questions, the evaluation takes shape.

The challenge is to find those “vital few” facts among the “trivial many” that are high in payoff and information load (MacKenzie 1972). The 20-80 rule expresses the importance of focusing on the right information. The 20-80 rule states that, in general, 20 percent of the facts account for 80% of what’s worth knowing (Anderson 1980:26).

In working with intended users to understand the importance of focus, I often do a short exercise. It goes like this:

Let me ask you to put your right hand out in front of you with your arm fully extended and the palm of your hand open. Now, focus on the center of the palm of your hand. Really look at your hand in a way that you haven’t looked at it in a long time. Study the lines—some of them long, some short; some of them deep, some shallow; some relatively straight, some nicely curved, and some of them quite jagged and crooked. Be aware of the colors in your hand: reds, yellows, browns, greens, blues, different shades and hues. And notice the textures, hills and valleys, rough places and smooth. Become aware of the feelings in your hand, feelings of warmth or cold, perhaps tingling sensations.

Now, keeping your right hand in front of you, extend your left arm and look at your left palm in the same way, not comparatively, but just focus on the center of your left palm, studying it, seeing it, feeling it. . . . Really allow your attention to become concentrated on the center of your left palm, getting to know your left hand in a new way. (Pause.)

Now, with both arms still outstretched I want you to focus, with the same intensity

that you’ve been using on each hand, I want you to focus on the center of both palms at the same time. (Pause while they try.) Unless you have quite unusual vision, you’re not able to do that. There are some animals *who* can move their eyes independently of each other, but humans do not have that capability. We can look back and forth between the two hands, or we can use peripheral vision and glance at both hands at the same time, but we can’t focus intensely on the center of both palms simultaneously.

Focusing involves a choice. The decision to look at something is also a decision not to look at something. A decision to see something means that something else will not be seen, at least not with the same acuity. Looking at your left hand or looking at your right hand, or looking more generally at both hands, provides you with different information and different experiences.

The same principle applies to evaluation. Because of limited time and limited resources, it is never possible to look at everything in great depth. Decisions have to be made about what’s worth looking at. Choosing to look at one area in depth is also a decision not to look at something else in depth. Utilization-focused evaluation suggests that the criterion for making those choices of focus be the likely utility of the resulting information. Findings that would be of greatest use for program improvement, decision making, and/or development focus the evaluation.

A Cautionary Note and Conclusion

Making use the focus of evaluation enhances the likelihood of, but does not guarantee, actual use. There are no guarantees. All one can really do is increase the likelihood of use. Utilization-focused

evaluation is time-consuming, frequently frustrating, and occasionally exhausting. The process overflows with options, ambiguities, and uncertainties. When things go wrong, as they often do, you may find yourself asking a personal evaluation question: How did I ever get myself into this craziness?

But when things go right; when decision makers care; when the evaluation question is important, focused, and on target; when you begin to see programs changing even in the midst of posing questions—then evaluation can be exhilarating, energizing, and fulfilling. The challenges yield to creativity, perseverance, and commitment as those involved engage in that most splendid of human enterprises—the application of intellect and emotion to the search for answers that will improve human effort and activity. It seems a shame to waste all that intellect and emotion studying the wrong issues. That’s why it’s worth taking the time to carefully focus an evaluation for optimum utility.

Follow-Up Exercises

1. Conduct a goal-free inquiry with two or three participants in a program. Interview them about what has brought them to the program, what they feel they need, and what they think they are getting through program participation. Conduct the interview without reference to the program’s stated and official goals. Analyze the results. After the interviews, compare participants’ needs and reported results with the program’s goals. Comment on similarities and differences. Reflect on your experience doing goal-free evaluation interviews with participants.
2. In 2006, Muhammad Yunus won the Nobel Peace Prize for his innovative

microcredit work, which began in Bangladesh and became the Grameen Bank with 2.5 million borrowers worldwide, most of them women and all of them poor. It’s easy to locate the story on the Internet. Imagine that you were a developmental evaluator working alongside Yunus. Describe how you, as an evaluator, would have supported his innovation with developmental evaluation. What kinds of data could you have provided him? What kinds of decisions would you have helped him with? What role would you have played?

Or pick another social innovation as your case study. Consider the case of Candy Lightner and the founding of Mothers against Drunk Driving; or Bob Geldof and his work on Live Aid concerts; or any example of major social innovation. Put yourself into an example of social innovation as a developmental evaluator. Describe your role and the data you would provide to inform and guide the innovative process.

3. Review Menu 8.1 in this chapter. Select three quite different approaches, types, or areas of focus. Compare and contrast them emphasizing what factors, circumstances, and contingencies would lead you, as an evaluator, to recommend each one because of its particular suitability and utility for an evaluation situation and challenge you describe.

4. Review the section near the end of this chapter on matching evaluation to the stage of a program’s development. Use an example of an actual program that has been in existence for some time. Learn about the stages of that program’s development and match evaluation questions and data to those different stages.

MENU 8.1

Alternative Ways of Focusing Evaluations

Different types of evaluations ask different questions and focus on different purposes. This menu is meant to be illustrative of the many alternatives available. These options by no means exhaust all possibilities. Various options can be and often are used together within the same evaluation, or options can be implemented in sequence over a period of time, for example, doing implementation evaluation before doing outcomes evaluation, or formative evaluation before summative evaluation.

<i>Focus or Type of Evaluation</i>	<i>Defining Question or Approach</i>
Accountability focus	How resources have been appropriately used to accomplish intended results? Key issue: Who is accountable to whom for what? (Rogers 2005a)
Accreditation focus	Does the program meet minimum standards for accreditation or licensing? (Hughes and Kushner 2005)
Appreciative inquiry	What is best about the program? (Preskill 2005a)
Artistic evaluation (evaluation as art)	Emphasize the artistic and creative elements of evaluation design. (Callahan 2005; Donmoyer 2005a; Lincoln 1991; Patton 1981)
Attribution focus (also casual focus)	Determine the relationship between the program (as a treatment) and resulting outcomes: To what extent can the program be said to have caused the documented outcomes?
Beneficiary Assessment	The perspective of intended beneficiaries about what they have experienced, both processes and outcomes (Salmen and Kane 2006).
Capacity-building focus	Doing evaluation in a way that enhances the long-term capacity to engage in evaluation more systematically. (Baizerman, Compton, and Stockdill 2005; McDonald, Rogers, and Kefford 2003)
CIPP Model	Evaluation of an entity's context, inputs, processes, and products. (Stufflebeam 2005)
Cluster evaluation	Synthesizing overarching lessons and/or impacts from a number of projects within a common initiative or framework (Russon 2005).
Collaborative approach	Evaluators and intended users work together on the evaluation.
Comparative focus	How do two or more programs rank on specific indicators, outcomes, or criteria?

<i>Focus or Type of Evaluation</i>	<i>Defining Question or Approach</i>
Compliance focus	Are rules and regulations being followed?
Connoisseurship approach	Specialists or experts apply their own criteria and judgment, as with a wine or antiques connoisseur. (Donmoyer 2005b)
Context focus	What is the environment within which the program operates politically, socially, economically, culturally, and scientifically? How does this context affect program effectiveness?
Cost-benefit analysis	What is the relationship between program costs and program outcomes (benefits) expressed in dollars? (Levin 2005a)
Cost-effectiveness analysis	What is the relationship between program costs and outcomes where outcomes are <i>not</i> measured in dollars? (Levin 2005b)
Criterion focused	By what criteria (e.g., quality, cost, client satisfaction) should the program be evaluated?
Critical issues focus	Critical issues and concerns of primary intended users focus the evaluation.
Culturally responsive	Focusing on the influences of cultural context and factors on program processes and outcomes. (Hood 2005)
Decisions focus	What information is needed to inform specific future decisions?
Deliberative democratic approach	This approach uses concepts and procedures from democracy to arrive at justifiable conclusions through inclusion, dialogue, and deliberation. (House 2005a; MacDonald and Kushner 2005; House and Howe 2000)
Descriptive focus	What happens in the program? (No “why” question or cause-effect analysis)
Developmental evaluation	The purpose is program or organizational development and rapid response to emergent realities in highly dynamic and complex systems under conditions of uncertainty.
Diversity focus	The evaluation gives voice to different perspectives on and illuminates various experiences with the program. No single conclusion or summary judgment is considered appropriate.
Effectiveness focus	To what extent is the program effective in attaining its goals? How can the program be more effective?
Efficiency focus	Can inputs be reduced and still obtain the same level of output or can greater output be obtained with no increase in inputs?

(Continued)

MENU 8.1 (Continued)

<i>Focus or Type of Evaluation</i>	<i>Defining Question or Approach</i>
Effort focus	What are the inputs into the program in terms of number of personnel, staff/client ratios, and other descriptors of levels of activity and effort in the program?
Empowerment evaluation	The evaluation is conducted in a way that affirms participants' self-determination and political agenda. (Fetterman and Wandersman 2005)
Equity focus	Are participants treated fairly and justly?
Ethnographic focus	What is the program's culture?
Evaluability assessment	Is the program ready for formal evaluation? What is the feasibility of various evaluation approaches and methods?
Extensiveness focus	To what extent is the program able to deal with the total problem? How does the present level of services and impacts compare to the needed level of services and impacts?
External evaluation	The evaluation is conducted by specialists outside the program and independent of it to increase credibility
Feminist evaluation	Evaluations conducted for the explicit purpose of addressing gender issues, highlighting the needs of women, and promoting change through increased social justice. (Seigart 2005; Seigart and Brisolaro 2002)
Formative evaluation	How can the program be improved?
Goals-based focus	To what extent have program goals and intended outcomes been attained?
Goal-free evaluation	To what extent are actual needs of program participants being met (without regard to stated program goals)?
Horizontal evaluation	Evaluation, knowledge sharing, and program development within a horizontal network (Thiele et al. 2007)
Inclusive evaluation	Emphasizes stakeholder inclusiveness, dialogical data collection methods, social justice, cultural pluralism, and transformation. (Mertens 2005)
Impact evaluation	What are the direct and indirect program impacts, over time, not only on participants, but also on larger systems and the community? Impact evaluation often includes a focus on determining the extent to which results can be attributed to the intervention.

<i>Focus or Type of Evaluation</i>	<i>Defining Question or Approach</i>
Implementation focus	To what extent was the program implemented as designed? What issues surfaced during implementation that need attention in the future?
Inputs focus	What resources (money, staff, facilities, technology, etc.) are available and/or necessary?
Internal evaluation	Program employees conduct the evaluation.
Intervention-oriented evaluation	Design the evaluation to support and reinforce the program's desired results.
Learning-oriented evaluation	Focusing the evaluation on practice improvement and organizational learning (Rogers and Williams 2006).
Judgment focus	Make an overall judgment about the program's merit, worth, and/or significance (see also summative evaluation).
Judicial model	Two evaluation teams present opposing views of whether the program was effective, like a legal trial (Datta 2005).
Knowledge focus (or lessons learned)	What can be learned from this program's experiences and results to inform future efforts?
Logical framework	Specify goals, purposes, outputs, and activities, and connecting assumptions: for each, specify indicators and means of verification
Longitudinal focus	What happens to the program and to participants over time?
Metaevaluation	Evaluation of evaluations: Was the evaluation well done? Is it worth using? Did the evaluation meet professional standards and principles? (Scriven 2005b).
Mission focus	To what extent is the program or organization achieving its overall mission? How well do outcomes of departments or programs within an agency support the overall mission?
Monitoring focus	Routine data collected and analyzed routinely on an ongoing basis, often through a management information system.
M & E (Monitoring and evaluation)	M & E: Integrating, monitoring, and evaluation (Jackson 2005; Kusek and Rist 2004).

(Continued)

304 ■ FOCUSING EVALUATIONS: CHOICES, OPTIONS, AND DECISIONS

MENU 8.1 (Continued)

<i>Focus or Type of Evaluation</i>	<i>Defining Question or Approach</i>
Needs assessment	What do clients need and how can those needs be met? (Altschuld and Kumar 2005)
Needs-based evaluation	(See Goal-free evaluation.)
Norm-referenced approach	How does this program population compare with some specific norm or reference group on selected variables?
Outcomes evaluation	To what extent are desired client/participant outcomes being attained? What are the effects of the program on clients or participants?
Participatory evaluation	Intended users, usually including community members, program participants, and/or staff, are directly involved in the evaluation. (Salmen and Kane 2006; King 2005)
Personalizing evaluation	Portrayal of people's lives and work as contexts within which to understand a program. (Kushner 2005, 2000)
Personnel evaluation	How effective are staff members in carrying out their assigned tasks and in accomplishing their assigned or negotiated goals?
Process focus	Evaluating the activities and events that occur as part of implementation: What do participants experience in the program? What are strengths and weaknesses of day-to-day operations? How can these processes be improved?
Product evaluation	What are the costs, benefits, and market for a specific product?
Program theory evaluation	Making explicit and testing the program's theory of change: What is the program's theory of change and to what extent do empirical findings support the theory in practice? (Rogers et al. 2000; Rogers 2000a, 2000b).
Quality assurance	Are minimum and accepted standards of care being routinely and systematically provided to patients and clients? (Williams 2005a).
Questions focus	What do primary intended users want to know that would make a difference to what they do? The evaluation answers questions instead of making judgments. (Russ-Eft 2005)
Realist evaluation (also realistic evaluation)	What are the underlying mechanisms (possible mediators) of program effects? What values inform the application of findings for social betterment? What works for whom in what circumstances and in what respects, and how? The result is a context-mechanism-outcome configuration. (Pawson and Tilley 2005; Mark, Henry and Julnes 2000)

<i>Focus or Type of Evaluation</i>	<i>Defining Question or Approach</i>
Real-world evaluation	How can evaluation be done under budget, time, data, and political constraints? (Bamberger, Rugh, and Mabry 2006)
Reputation focus	How the program is perceived by key knowledgeable and influentials. Ratings of the quality of universities are often based on reputation among peers.
Responsive evaluation	What are the various points of view of different constituency groups and stakeholders? The responsive evaluator works to capture, represent, and interpret these varying perspectives under the assumption each is valid and valuable. (Stake and Abma 2005)
Social and community indicators	What routine social and economic data should be monitored to assess the impacts of this program? What is the connection between program outcomes and larger-scale social indicators, for example, crime rates?
Social justice focus	How effectively does the program address social justice concerns? (House 2005b)
Success case method	Compares highly successful participants with unsuccessful ones to determine primary factors of success. (Brinkerhoff 2005, 2003)
Summative evaluation	Should the program be continued? If so, at what level? What is the overall merit and worth of the programs?
Systems focus	Using systems thinking, concepts, perspectives, and approaches as the framework for evaluation. (Williams and Iman 2006; Williams 2005b)
Theory-driven evaluation	On what theoretical assumptions and model is the program based? What social scientific theory is the program a test of and to what extent does the program confirm the theory? (Rogers, personal communication, e-mail February 24, 2007; Chen 2005a, 2005b)
Theory of change approach	What are the linkages and connections between inputs, activities, immediate outcomes, intermediate outcomes, and ultimate impacts?
Transformative evaluation	Diverse people are included in the evaluation in a way that is genuinely and ethically respectful of their culture, perspectives, political and economic realities, language, and community priorities? (Mertens 2007)
Utilization-focused evaluation	<i>Intended use by intended users:</i> What information is needed and wanted by primary intended users that will actually be used for program improvement and decision making? (Utilization-focused evaluation can include any of the other types above.)

