4

Situation Analysis Conducted Jointly With Primary Intended Users

Situations come and situations go.

Situations change and situations endure.

Situations get better and situations get worse.

Situations emerge and situations abate.

All of which begs the question:

What damn situation are we talking about?

Halcolm's Treatise on Situation Analysis

Situation analysis is ongoing. It's not something you do once, congratulate yourself on nailing it, check it off as done, and move on, never to pass that way again. The first three utilization-focused evaluation (U-FE) steps involve situation analysis:

Step 1. Assessing program and organizational readiness for utilization-focused evaluation, essentially analyzing the program situation.

Step 2. Assessing evaluator readiness and competence to undertake a utilization-focused evaluation, essentially analyzing the evaluator's situation.

Step 3. Identifying primary intended users and assessing their readiness to engage in utilization-focused evaluation, essentially analyzing the stakeholder situation.

As the Complex Systems Interconnections Graphic 1-2-3 (p. 86) shows, these steps—and situations—are interdependent. So why is Step 4 situation analysis? I'm glad you asked.

We're working within the constraint of a step-by-step and chapter-by-chapter framework, which is inherently linear, taking one thing at a time. Simultaneously, we're trying to recognize and call attention to the nonlinear, interactive, and iterative nature of utilization-focused evaluation as it unfolds in the real world with actual primary intended users. So Step 4 calls for a focused, intentional, and systematic situation analysis undertaken *with primary intended users* while at the same time reminding you that you've been engaged in situation analysis from the very beginning, and admonishing you to keep analyzing the situation as you move forward.

Step 4 focuses on bringing the primary intended users fully into the situation analysis. Exhibit 4.1 presents some questions for beginning situation analysis aimed at assuring understanding of the program, appreciating stakeholders' interests and potential areas of conflict, understanding the program's prior history and experiences with evaluation, and making explicit the decision context that will affect use, namely: What decisions, if any, will the evaluation findings inform? From these basic situation analysis questions, we can move to a more in-depth understanding of the evaluation use situation by identifying and examining factors that may contribute to use and those that may be barriers. First, let's look more closely at the Herculean challenges of situation analysis.

Context Matters: The Challenge of Situation Analysis in Designing and Conducting Evaluations

Every evaluation situation is unique. A successful evaluation (one that is useful, practical, ethical, accurate, and accountable) emerges from the special characteristics and conditions of a particular situation—a mixture of people, politics, history, context, resources, constraints, values, needs, interests, and chance. The standards and principles of evaluation provide overall direction, a foundation of ethical guidance, and a commitment to professional competence and integrity, but there are no absolute rules an evaluator can follow to know exactly what to do with specific users in a particular situation. As an evaluation unfolds, evaluators and primary intended users must work together to identify the evaluation that best fits their information needs and the program's context and situation. This means *negotiating* the evaluation's intended and desired uses, and adapting the design to financial, political, timing, and methodological constraints and opportunities.

To appreciate how complicated it can be to design an evaluation to fit the program's situation, let's use playing chess as an analogy. Bruce Pandolfini (1998), a world-class chess master, consults with major corporate leaders to teach them the mindset of a chess master so that they can become more skilled at strategic analysis and thinking. He points out that there are some 85 billion ways of playing just the first four moves in a game of chess (that's 85, with 9 zeros—85,000,000,000). Deciding what moves to make requires both strategy

EXHIBIT 4.1

Beginning Situation Analysis

Understand the program

What is the program's history? What situation gave rise to the program?

What are the program's primary goals? To what extent are these goals clear, specific, and measurable?

What are the strategies for attaining these goals?

Who are the intended beneficiaries of the program's intervention? What are their characteristics?

What are staff characteristics?

What's the program's budget?

For existing programs, how has the program changed overtime? What led to those changes?

Identify primary stakeholders and their interests

Where do stakeholders' interests align?

Where do their interests conflict?

What's the political context for the evaluation?

Who will be the primary intended users of the evaluation?

Evaluation history

What prior experiences, if any, has the program had with evaluation?

What are current monitoring and evaluation approaches, if any? How are monitoring and evaluation data currently used, if at all? What factors affect current uses?

What capacities does the program have to engage in evaluation (staff skills, budget for evaluation, information systems, a culture of inquiry, data management, and interpretation capacity)?

Decision and action context

What's the primary intended purpose of the evaluation?

What decisions, if any, is the program facing? What are the time lines for any such decisions?

What uncertainties does the program face? Externally? Internally?

and tactics grounded in an analysis of the situation presented by a particular game and opponent within an overall framework of fundamental chess ideas and concepts, understanding what the different pieces do, how they can be moved, and how they relate to each other. Once the game starts, subsequent moves are contingent on and must be adapted to what one's opponent does and the unfolding situation.

So how does one play the game of evaluation? And to successfully play the game, what is the focus of situation analysis in evaluation? The challenge is to identify factors and forces that will support and enhance evaluation use versus those factors and forces that may constitute barriers and become sources of resistance to use. Playing the evaluation use game successfully means marshalling powerful forces in support of use to overcome resistance to use. Force field analysis helps conceptualize the forces at play in the real-world utilization game.

Force Field Analysis

The idea of force field analysis was developed by eminent social psychologist Kurt Lewin in the 1940s. For any decision, he posited that there are forces that affect movement toward a goal—helping forces versus hindering forces. He applied force field analysis to such issues as behavior change, learning outcomes, group conflict resolution, and organizational morale. We're going to apply it to evaluation use. Exhibit 4.2 shows such a force field analysis.

The factors and forces identified along the top are those deemed likely to support and enhance use. Those listed along the bottom are predicted to constrain or undermine evaluation use. The length of each arrow represents the relative strength of the force, with longer arrows representing stronger forces. Broader arrows will have broad effects, while narrower arrows narrower effects on use. Dotted lines are less powerful than solid lines. The estimates of strength, breadth, and power do not have to be accurate. They are relative estimates, allowing rough comparisons of factors. Question marks may be inserted where a factor has been identified but the group is uncertain about its direction or strength. For example, a program may have many staff new to evaluation. Will that be a supporting or constraining factor, or manifest aspects of both? Not sure? Insert a question mark. Then monitor how that factor plays out over time. The force field analysis offers a framework for engaging with primary intended users in thinking strategically about use. At the beginning of the evaluation it is a baseline. When periodically updated, perhaps quarterly or annually, it provides a map of which factors and forces are actually unfolding as likely to affect evaluation use.
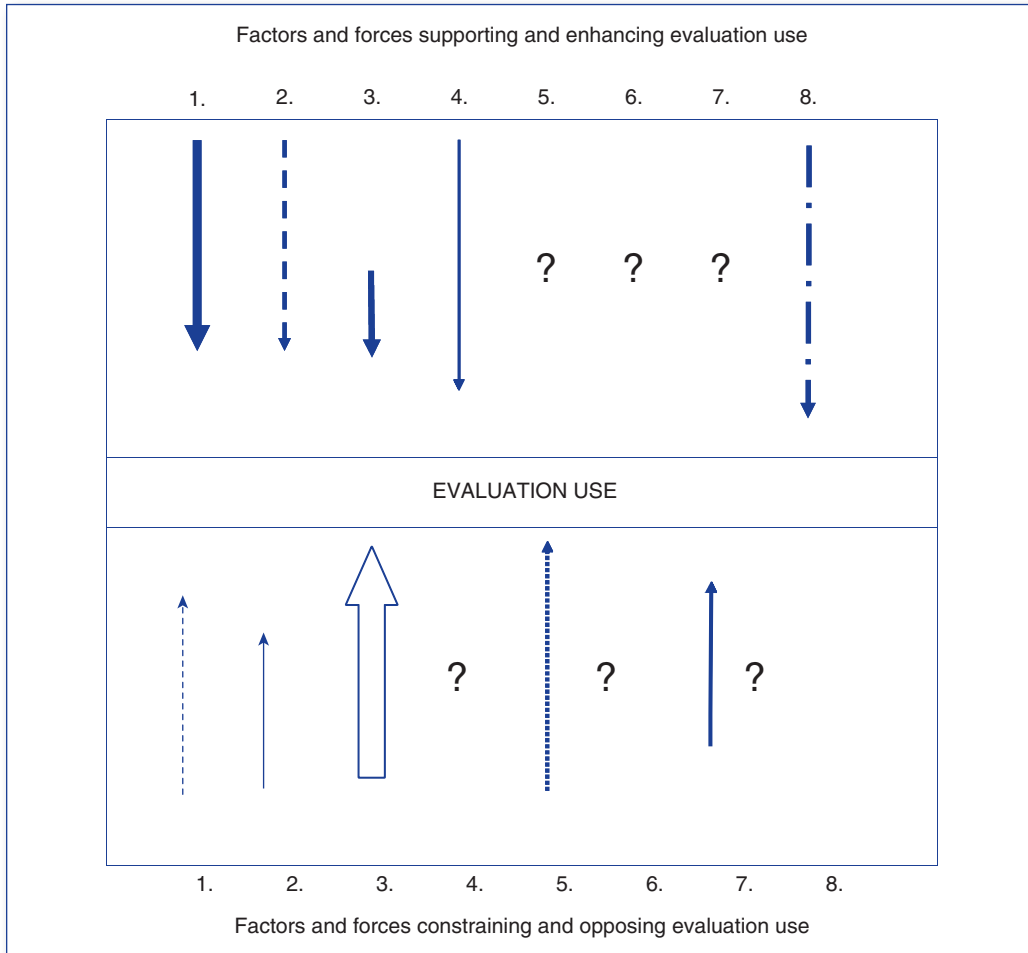
Once the overall force field analysis has been constructed as a baseline situation analysis, Exhibit 4.3 shows a template that can be used to engage primary intended users in considering ways of reinforcing and strengthening factors that will enhance use and counter forces that may undermine use.

Identifying Factors and Forces

With the overall framework of force field analysis in mind, let's turn to the task of identifying and distinguishing positive and negative factors and forces. Then we'll look at an example that applies these factors in a force field situation analysis.

EXHIBIT 4.2

Template for Force Field Analysis:
Factors Supporting and Constraining Evaluation Use

Factors and forces supporting and enhancing evaluation use

1.   2.   3.   4.   5.   6.   7.   8.

?   ?   ?

EVALUATION USE

?   ?   ?

1.   2.   3.   4.   5.   6.   7.   8.

Factors and forces constraining and opposing evaluation use

The length of each arrow represents the relative strength of the force, with longer arrows representing stronger forces. Broader arrows are expected to have broad effects, while narrower arrows will have somewhat less effect on use. Dotted lines are less powerful than solid lines. These are relative estimates, allowing rough comparisons of factors. Question marks may be inserted where a factor has been identified but the group is uncertain about its direction or strength.

EXHIBIT 4.3

Strategic Engagement for Evaluation Use

Factors and forces that support evaluation use:	*What can be done to reinforce and strengthen positive factors and forces?*
1.	1.
2.	2.
3.	3.
4.	4.
5.	5.
Factors and forces that may constrain or undermine evaluation use:	*What can be done to reduce or redirect negative factors and forces?*
1.	1.
2.	2.
3.	3.
4.	4.
5.	5.

Exhibit 4.4 (pp. 93–94) lists just 20 of the situational variables that can affect how an evaluation is designed and conducted, things like number of stakeholders to be dealt with, the evaluation's purpose, staff attitudes toward evaluation, the budget and time line for evaluation, and the program's prior experience with evaluation. These variables are presented in no particular order. Most of them could be broken down into several additional dimensions. If we conceive of just three points (or situations) on each of these dimensions—the two endpoints and a midpoint: for example, low budget, moderate budget, substantial budget—then the possible combinations of these 20 dimensions represent 8,000 unique situational configurations for evaluation.

Nor are these static situations. The program you thought was new at the first session turns out to have been created out of and to be a continuation of another program; only the name has been changed to protect the guilty. You thought you were dealing with only

EXHIBIT 4.4

Examples of Situational Factors in Evaluation That Can Affect Users' Participation and Use

One primary decision maker	1. Number of stakeholders to be dealt with	Large number
Formative purpose (improvement)	2. Purpose of the evaluation	Summative purpose (funding decision)
New program	3. History of the program	Long history
Enthusiasm	4. Staff attitude toward evaluation	Resistance
Knows virtually nothing	5. Staff knowledge about evaluation	Highly knowledgeable
Cooperative	6. Program interaction patterns (administration-staff, staff-staff, staff-client)	Conflict laden
First time ever	7. Program's prior evaluation experience	Seemingly endless experience
High	8. Staff and participants education levels	Low
Homogeneous groups	9. Staff and/or participants' characteristics (pick any 10 you want)	Heterogeneous groups
One site	10. Program location	Multiple sites
No money to speak of	11. Resources available for evaluation	Substantial funding
One funding source	12. Number of sources of program funding	Multiple funding sources

(Continued)

Simple and singular	13. *Nature of the program treatment*	Complex and multidimensional
Highly standardized and routine	14. *Standardization of treatment*	Highly individualized and nonroutine
Horizontal, little hierarchy, little stratification	15. *Program organizational decision-making structure*	Hierarchical, long chain of command, stratified
Well articulated, specifically defined	16. *Clarity about evaluation purpose and function*	Ambiguous, broadly defined
Operating information system	17. *Existing data on program*	No existing data
External	18. *Evaluator(s)' relationship to the program*	Internal
Voluntary, self-initiated	19. *Impetus for the evaluation*	Required, forced on program
Long time line, open	20. *Time available for the evaluation*	Short time line, fixed deadline

one primary decision maker at the outset, and suddenly you have stakeholders coming out your ears, or vice versa. With some programs, I've felt like I've been through all 8,000 situations in the first month of design negotiations.

Now, just in case 8,000 situations to analyze, be sensitive to, and design evaluations for doesn't seem challenging enough, add two more points to each dimension—a point between each endpoint and the midpoint. Now, combinations of the five points on all 20 dimensions yield 3,200,000 potentially different situations. Perhaps such complexity helps explain why the slogan that won the hearts of evaluators in attendance at the pioneering 1978 Evaluation Network conference in Aspen, Colorado, was the lament:

Evaluators do IT under difficult circumstances.

The point of this analysis is to raise a fundamental question: How can evaluators pre-pare themselves to deal with a lot of different people and a huge variety of situations? The research on decision making says we can't systematically consider every possible variable, or even 50 variables, or even 20 variables. What we need is a framework for making sense of situations, for telling us what factors deserve priority based on research and desired results. Such a framework, rather than providing narrow, specific prescrip-tions, should offer questions to force us to think about and analyze the situation. This is essential and reminds us of the title of this book: *Essentials of Utilization-Focused Evaluation*. The *utilization-focused* advice for managing situational complexity in evalu-ation is to *stay focused on use*. For every issue that surfaces in evaluation negotiations, for every design decision, for every budget allocation, and for every choice among alter-natives, keep asking: How will what we do affect use in this situation? Here are some actual situations I've faced.

- The program has a new director. The previous director, much admired and trusted, retired after 15 years of service. The new director was hired after a national search and brought in from the outside although there were two internal candidates for the position. The evaluation process is scheduled to begin after the new director has been on the job for only a month. How might this affect the evaluation and its use? What are the elements of the situation that might support evaluation use? What are the potential pitfalls and barriers to use? How can these be acknowl-edged and managed?
- In another situation, the evaluation will focus on a major new initiative considered very innovative and risky because it involves a creative approach to poverty reduction that hasn't been tried in this community before. The program implementing the new initiative has a long history of solid performance delivering basic community services to people in poverty. The program's stellar reputation flows from effectively implementing government and United Way services in accordance with established procedures and rules. However, the agency lead-ership and staff are not known for being innovative or creative. They have been asked to take on this innovative community engagement program because of their reputation for meeting expectations. Their past approaches to evaluation have been modest and centered on deliver-ing outputs (low-cost services) but with little attention to outcomes (no follow-up on the effects of services). The new initiative includes lots of rhetoric about being outcomes-focused. How might this situation affect the evaluation and its use? What are the elements of the situation that might support evaluation use? What are the potential pitfalls and barriers to use? How can these be acknowledged and managed? What evaluation capacity-building will need to occur?
- A few months after the devastating earthquake in Haiti, the international agencies involved in relief work were asked to cooperate in a collaborative evaluation of lessons learned that might improve future disaster relief efforts. Cooperation and collaboration are important so that the evaluation data collection does not become redundant and intrusive on continuing relief efforts as it surely would if every agency conducted its own evaluation and had to seek data from the

same few, already-overworked Haitian sources. But such a collaborative and comprehensive evaluation has not been undertaken before. The cooperating relief agencies traditionally compete for resources and media attention. Time is short and resources for the evaluation will be limited, but the stakes are high for the agencies and the people of Haiti, as well as future victims of natural disaster. How might this context affect the evaluation and its use? What are the elements of the situation that might support evaluation use? What are the potential pitfalls and barriers to use?

These are just three brief examples of the millions of situations evaluators face. They are meant to give a hint of the challenges of situation awareness. Exhibit 4.5 provides an overview of some common situations that require special evaluator skills.

Internal and External Evaluators

Internal and external evaluators face quite different situations. Internal evaluators are employed by the program, project, or organization being evaluated. External evaluators work as independent contractors. This raises a fundamental situational issue regarding the location of the evaluator inside or outside the entity being evaluated, what is sometimes called the "in-house" versus "outhouse" issue. Let's take a closer look at internal versus external evaluation situations.

External evaluators come from universities, consulting firms, and research organizations or work as independent consultants. The defining characteristic of external evaluators is that they have no long-term, ongoing position within the program or organization being evaluated. They are therefore not subordinated to someone in the organization and not directly dependent on the organization for their job and career. External evaluators are valuable precisely because they are outside the organization. It is typically assumed that their external status permits them to be more independent, objective, and credible than internal evaluators. Internal evaluations are suspect because, it is presumed, they can be manipulated more easily by administrators to justify decisions or pressured to present positive findings for public relations purposes. Of course, external evaluators who want future evaluation contracts are also subject to pressure to produce positive findings. In addition, external evaluators are also typically more costly, less knowledgeable about the nuances and particulars of the local situation, and less able to follow through to facilitate the implementation of recommendations. When external evaluators complete their contract, they may take with them a great deal of knowledge and insight that is lost to the program. That knowledge stays "in-house" with internal evaluators. External evaluators have also been known to cause difficulties in a program through insensitivity to organizational relationships and norms, one of the reasons the work of external evaluators is sometimes called "outhouse" work.

EXHIBIT 4.5

Examples of Situations That Pose Special Challenges to Evaluation Use and the Evaluator's Role

Situation	Challenge	Special Evaluator Skills Needed
1. Highly controversial issue	Facilitating different points of view	Conflict resolution skills
2. Highly visible program	Dealing with publicity about the program; reporting findings in a media-circus atmosphere	Public presentation skills Graphic skills Media-handling skills
3. Highly volatile program environment	Rapid change in context, issues, and focus	Tolerance for ambiguity Being a "quick study" Rapid responsiveness Flexibility
4. Cross-cultural or international	Including different perspectives. Being aware of cultural blinders and biases	Cross-culture sensitivity Skills in understanding and incorporating different perspectives
5. Team effort	Managing people	Identifying and using individual skills of team members; team-building skills
6. Evaluation attacked	Preserving credibility	Calm; staying focused on evidence and conclusions
7. Corrupt program	Resolving ethical issues/ upholding standards	Integrity Clear ethical sense Honesty

Most federal, state, local, and international agencies have internal evaluation units to fulfill accountability mandates. It has become clear that internal evaluators can produce evaluations of high quality and high impact while still performing useful service to administrators if they work diligently to establish an image of an independent but active voice in the organizational structure and take a pragmatic approach to helping solve

management problems. Ongoing performance monitoring has become a major activity of internal evaluation systems, and internal evaluation units now support both accountability and learning.

Over the years, I have had extensive contact with internal evaluators through training and consulting, working closely with several of them to design internal monitoring and evaluation systems. I interviewed 10 internal evaluators who I knew used a utilization-focused approach. Their comments about how they have applied utilization-focused principles offer insights into the world of the internal evaluator and illuminate research findings about effective approaches to internal evaluation.

Themes From Internal Evaluators

1. **Actively involving stakeholders within the organization can be difficult** because evaluation is often perceived by both superiors and subordinates as the job of the evaluator. The internal evaluator is typically expected to *do* evaluations, not facilitate an evaluation process involving others. Internal evaluators who have had success involving others have had to work hard at finding special incentives to attract participation in the evaluation process. One internal evaluator commented,

> My director told me he doesn't want to spend time thinking about evaluations. That's why he hired me. He wants me to "anticipate his information needs." I've had to find ways to talk with him about his interests and information needs without explicitly telling him he's helping me focus the evaluation. I guess you could say I kind of involve him without his really knowing he's involved.

2. **Internal evaluators are often asked by superiors for public relations information rather than evaluation.** The internal evaluator may be told, "I want a report for the legislature proving our program is effective." It takes clear conviction, subtle diplomacy, and an astute understanding of how to help superiors appreciate evaluation to keep internal evaluation responsibilities from degenerating into public relations. One mechanism used by several internal evaluators to increase support for real evaluation rather than public relations is establishing an evaluation advisory committee, including influential people from outside the organization, to provide independent checks on the integrity of internal evaluations.

3. **Internal evaluators get asked to do lots of little data-gathering and report-writing tasks** that are quite time consuming but too minor to be considered meaningful evaluation. For example, if someone in the agency wants a quick review of what other states are doing about some problem, the internal evaluator is an easy target for the task. Such assignments can become so pervasive that it's difficult to have time for longer-term, more meaningful evaluation efforts.

4. **Internal evaluators are often excluded from major decisions** or so far removed from critical information networks that they don't know about new initiatives or developments in time to build in an evaluation perspective up front. One internal evaluator explained,

> We have separate people doing planning and evaluation. I'm not included in the planning process and usually don't even see the plan until it's approved. Then they expect me to add on an evaluation. It's a real bitch to take a plan done without any thought of evaluation and add an evaluation without changing the plan. They think evaluation is something you do at the end rather than think about from the start. It's damn hard to break through these perceptions. Besides, I don't want to do the planners' job, and they don't want to do my job, but we've got to find better ways of making the whole thing work together. That's my frustration. . . . It takes me constantly bugging them, and sometimes they think I'm encroaching on their turf. Some days I think, "Who needs the hassle?" even though I know it's not as useful just to tack on the evaluation at the end.

5. **Getting evaluation used takes a lot of follow-through.** One internal evaluator explained that her job was defined as data gathering and report writing without consideration of following up to see if report recommendations were adopted (Step 16 in this book). That's not part of her job description, and it takes time and some authority. She mused:

> How do I get managers to use a report if my job is just to write the report? They're above me. I don't have the authority to ask them in 6 months what they've done. I wrote a follow-up memo once reminding managers about recommendations in an evaluation and some of them didn't like it at all, although a couple of the good ones said they were glad I reminded them.

Another internal evaluator told me he had learned how to follow up informally. He has 7 years' experience as an internal human services evaluator. He said,

> At first I just wrote a report and figured my job was done. Now, I tell them when we review the initial report that I'll check back in a few months to see how things are going. I find I have to keep pushing, keep reminding, or they get busy and just file the report. We're gradually getting some understanding that our job should include some follow-up. Mostly it's on a few things that we decide are really important. You can't do it all.

Internal Role Definitions

The themes from internal evaluators indicate the importance of carefully defining the job to include attention to use. When and if the internal evaluation job is defined primarily as writing a report and filling out routine reporting forms, the ability of the evaluator to influence use is quite limited. When and if the internal evaluator is organizationally separated

from managers and planners, it is difficult to establish collaborative arrangements that facilitate use. Thus, a utilization-focused approach to internal evaluation will often require a redefinition of the position to include responsibility for working with intended users to develop strategies for acting on findings.

One increasingly important role for internal evaluators is as a resource for infusing evaluative thinking into and throughout the entire organization. This means that rather than only or primarily conducting evaluations, the internal evaluator becomes a trainer, a resource to other units, a facilitator of meetings where evaluative thinking is needed, and an evaluator of the organization's progress in learning and applying those learnings to its work. In this role the internal evaluator works to build evaluation into the organizational culture.

Internal-External Evaluation Combinations

It is important to note that internal *and* external approaches are not mutually exclusive. Actually, there are a good many possible combinations of internal and external evaluations that may be more desirable and more cost-effective than either a purely internal *or* purely external evaluation. Exhibit 4.6 describes some of the points along the external-internal continuum. These constitute varying situation-specific arrangements for locating the evaluation function.

EXHIBIT 4.6

A Continuum of Internal/External Evaluation Relationships

1. *Entirely External.* None of the evaluations of the organization's programs or projects are being completed by internal staff members. No other evaluation activities (e.g., developing program logic models, or evaluation plans) are performed by staff. When evaluation occurs in the organization, it is in response to funders' demands for accountability and is conducted by external evaluation consultants.

2. *Minimal Ad Hoc Internal Evaluation.* Program staff have conducted the evaluation of only a minority of the organization's programs and projects. Those evaluations were ad hoc; that is, they occurred in response to requests from individual managers or funders. Usually, the focus and questions for these evaluations were set by external stakeholders (e.g., providing required performance indicators to funders' accreditation demands).

3. *Occasional Internal Evaluation.* When staff perform evaluations of programs or projects, they usually focus on questions about outputs and processes (e.g., What services were delivered to which clients?). These internal evaluations are conducted by managers or staff who are temporarily given evaluation responsibilities. Core evaluation activities such as having staff create logic models only occur rarely.

4. *Part-Time Internal Evaluator.* The organization has assigned one staff member to perform evaluation tasks on a part-time basis. This person gets their assignments from the Executive Director (or a similar senior manager). Internal evaluations often focus on whether or not the program or project is doing what both the organization and its funders want it to do (e.g., Is the program meeting the goals stated in the program proposal?).

5. *Full-Time Internal Evaluator.* Evaluations performed by internal staff are fairly common in the organization with at least one staff member assigned to evaluation duties on an ongoing basis. Program managers participate in identifying priority evaluation questions and planning evaluations. These internal evaluations often include questions about program outcomes (e.g., How effective was the program? Did clients benefit? To what extent did the program produce its intended outcomes?).

6. *Routine Internal Evaluation.* Evaluation occurs on a regular basis. Several internal staff members have evaluation skills and plan/manage internal evaluations on a regular basis. The organization has policies that require that certain evaluation tasks must occur throughout the organization (e.g., all programs must have a logic model, all programs must collect data on client satisfaction). Results from internal evaluations are routinely reported to managers and staff. These evaluation results are used to inform decisions about the development of the program that was evaluated. Internal evaluations often focus on issues of program costs. Program managers decide which evaluation questions will get asked. The organization has an Evaluation Coordinator or Manager and several staff who have evaluation responsibilities.

7. *Fully Integrated and Highly Valued Internal Evaluation.* Evaluation of all programs and projects is an organizational requirement. An Evaluation Manager leads an internal Evaluation Team. Evaluation staff provide evaluation training and coaching to managers and staff, including how to use findings in their work. Findings are used to improve both individual programs and the entire organization's structures and processes in an ongoing way. Results from internal evaluations are shared with the Board, with partners, and with key stakeholders. Summaries of evaluation findings appear in the newsletter and Annual Report. Evaluation is viewed as central to organizational effectiveness and is an integral part of the organization's culture.

SOURCE: Adapted from Shea and Love (2007).

Accreditation processes are a good example of an internal-external combination. The internal group collects the data and arranges them so that the external group can come in, inspect the data collected by the internal group, sometimes collect additional information on their own, and pass judgment on the program.

There are many ways in which an evaluation can be set up so that some external group of respected professionals and evaluators guarantees the validity and fairness of the evaluation process while the people internal to the program actually collect and/or analyze the evaluation data. The cost savings of such an approach can be substantial while still allowing the evaluation to have basic credibility and legitimacy through the blessing of the external review committee.

I worked for several years with one of the leading chemical dependency treatment centers in the country, the Hazelden Foundation of Minnesota. The foundation has established a rigorous evaluation process that involves data collection at the point of entry into the program and then follow-up questionnaires 6 months, 12 months, and 24 months after leaving the program. Hazelden's own research and evaluation department collects all of the data. My responsibility as an external evaluator was to monitor that data collection periodically to make sure that the established procedures were being followed correctly. I then worked with the program decision makers to identify the kind of data analysis that was desirable. They performed the data analysis with their own computer resources. They sent the data to me, and I wrote the annual evaluation report. They participated in analyzing, interpreting, and making judgments about the data, but for purposes of legitimacy and credibility, the actual writing of the final report was done by me.

When orchestrating an internal-external combination, one danger to watch for is that the external group may impose unmanageable and overwhelming data collection procedures on the internal people. I saw this happen in an internal-external model with a group of school districts in Canada. The external committee set as the standard doing "comprehensive" data collection at the local school level, including data on learning outcomes, staff morale, facilities, curriculum, the school lunch program, the library, parent reactions, the perceptions of local businesspeople, analysis of the school bus system, and so on. After listening to all of the things the external committee thought should be done, the internal folks dubbed it the Internal-External-*Eternal* model of evaluation.

The point is that a variety of internal-external combinations are possible to combine the lower costs of internal data collection with the higher credibility of external review. In working out the details of internal-external combinations, care will need to be taken to achieve an appropriate and mutually rewarding balance based on a collaborative commitment to the standards of utility, feasibility, propriety, and accuracy.

Evaluation as a Leadership Function

Most writings about internal evaluation assume a separate unit or specialized position with responsibility to conduct evaluations. An important new direction in evaluation is to treat evaluation as a leadership function of all managers and program directors in the organization, including, especially, the executive director. The person responsible for internal evaluation then plays a facilitative, resource, and training function in support of managers and leaders rather than spending time actually conducting evaluations. The best example of this approach I've worked with and observed up close was the position of

associate administrator for Performance Measurement and Evaluation in a county government. The county had no internal evaluation office. Rather, this senior position, as part of the County Executive team, had responsibility to infuse evaluation thinking and systems throughout the county, in every department and program. Every manager in the county received training in how to build outcomes evaluation into ongoing program processes and use data for decision making and budgeting. What made this approach to internal evaluation work, in my judgment, was three-fold:

1. Results-oriented evaluation was defined as a leadership function of every county manager, not just a technical reporting function delegated to data nerds.

2. The overall responsibility for evaluation resided at the highest level of the organization, in the executive team, with direct access to the County Board of Commissioners backed up by public commitments to use evaluation for decision making and budgeting.

3. Because of the prior two commitments, a person of great competence and dedication was selected to fill the position of associate administrator for Performance Measurement and Evaluation, after a national search.

These patterns of effectiveness stand out because so often internal evaluation is delegated to the lowest level in an organization and treated as a clerical or technical function. Indeed, being given an evaluation assignment is often a form of punishment, or a way of giving deadwood staff something meaningless to occupy themselves with. It is clear that, for internal evaluators to be useful and credible, they must have high status in the organization and real power to make evaluation meaningful. And even when this occurs, as in the case just reviewed, it can be difficult to sustain. After less than 2 years, when the county executive changed, the associate administrator for Performance Measurement and Evaluation was lost in a reorganization and the system reverted to treating evaluation as a separate support unit and function.

Elevating the status of evaluation to that of a leadership function may require leadership development. Indeed, as an example of reaching primary intended users, I participated in developing a leadership development workshop that focused on evaluative thinking and practice. We didn't promote it as an evaluation workshop because leaders would not come to such a workshop; they would send lower-level technical staff. To reach the leadership level of organizations with the message and promise of evaluation use, we had to promote the effort as leadership development and embed the evaluation training in that framework. Exhibit 4.7 presents the four functions of results-oriented, reality-testing, learning-focused leadership we used for the leadership training workshop. In this framework, evaluation becomes *an executive leadership responsibility* focused on decision-oriented use rather than a data-collection task focused on routine internal reporting.

EXHIBIT 4.7

Four Functions of Results-Oriented, Reality-Testing, Learning-Focused Leadership

- Create and nurture a results-oriented, reality-testing, learning-focused culture.
- Lead in deciding what outcomes to commit to and hold yourselves accountable for.
- Make measurement of outcomes thoughtful, meaningful, and credible.
- Use the results—and model for others serious use of results.

Evaluation Team Composition Analysis

Evaluations of much size or scope move us from the individual evaluator to a team of evaluators. Team situations raise questions about team composition. This typically involves calculating what combination of skills, knowledge, and experience are needed to conduct the evaluation. The primary intended users can be involved in establishing the diversity criteria for team composition that they believe will lend credibility to the team and ensure that the team has both the skills and diverse perspectives needed to do high-quality work. Here's an example of a diverse, multi-dimensional team of five people that could be assembled to evaluate an environmental initiative:

- An experienced lead evaluator able to coordinate the team and work with diverse intended users;
- A team member with expertise in environmental science;
- One or more culturally competent evaluators who could bring to the team the perspectives of any specific cultural group involved in the initiative: for example, a Native American if the initiative includes American Indian reservations, a Maori evaluator in New Zealand;
- A younger evaluator, newer to the work, who can bring fresh ideas and new technology applications to the team while being mentored professionally in evaluation by more experienced team members; and
- Gender balance to assure that the evaluation addresses and is sensitive to gender issues.

Most larger scale evaluations involve teams: for example, external site visit teams that visit projects for evaluation data collection purposes. Team composition includes calculations about the appropriate mix of expertise, experience, background, and perspective. In a 10-country evaluation of an agricultural extension initiative in the Caribbean, we assembled an evaluation task force of primary intended users that included the international funders, the chief agricultural extension officers of the participating countries, and representatives of the American universities and the University of the West Indies that

were providing technical assistance. The five-member evaluation team reflected this stakeholder diversity. In each country, a farmer and local extension agent joined the team for the work in that country.

Thus, part of situation analysis is determining the appropriate team composition that will bring evaluation expertise, subject matter expertise, cultural competence, methodological expertise, and political credibility to the team.

Balancing Task and Relationship Demands

One of the most fundamental and oft-replicated findings from early research on group effectiveness is that high-performing groups attend to both task completion and relationship-building. In an evaluation context, *the task focus* concerns the primary intended uses of the evaluation and how those uses will be achieved. *The relationship focus* concerns how the evaluator works with and relates to primary intended users to enhance the likelihood of use. Embedded in the utilization-focused evaluation goal of intended use by intended users is attention to both tasks and relationships. Situation analysis and responsiveness involve ongoing assessment of the balance between task completion and facilitating good working relationships. While internal and external evaluators may face different dynamics in this regard, what they share is a need to analyze the situation they face to determine what kind of relationship and process for conducting the evaluation will support task completion and lead to use. It is worth reiterating these challenges in the context of situation analysis.

Some evaluators focus only on getting the evaluation designed, the data collected and analyzed, and the report written. They are entirely task focused and want no part of relationship-building. Indeed, they wear their independence as a badge of pride, justifying their inattention to relationships and process as fundamental to their credibility. In this approach to evaluation, independence, neutrality, distance, and credibility are the cornerstones of utility.

At the other end of the continuum are evaluation consultants who make regular interaction with clients a priority and give at least as much attention to relationship-building as getting the work done. Such a consultant once told me, "Building a strong relationship with the client is the task." This evaluator viewed trust, in-depth knowledge, shared values, and close connection to the client as the pillars that support utility.

These two evaluators defined the situational challenge differently. For our purposes, these examples raise questions that can only be answered within the context of a particular situation.

- What kind of relationship to specific intended users will enhance use given the purpose of the evaluation?
- How much distance is needed to establish credibility? How much closeness is appropriate to assure relevance and trust?

- How much ongoing interaction with intended users supports mutual understanding and keeping everyone informed? When does regular communication become burdensome and overdone? What constitutes too little communication to maintain an appropriate degree of interest and engagement?
- How does the relationship with intended users change over the course of an evaluation as the tasks change (from design to data collection, to analysis, to reporting, to use of findings)?
- To what extent is it appropriate for an evaluator to have different relationships with different intended users? Some stakeholders are likely to be more interested in both task and process while others are less so. How does an evaluator deal with these variations, by having different relationships with different intended users, without creating conflicts and distrust in the group as a whole?

There can be no standardized, recipe-like answers to these questions. The answers flow from the situation analysis, including stakeholders' concerns and perspectives, which inform the entire evaluation process. Answering these questions means asking them—seriously, thoughtfully, astutely, and pragmatically—and then letting what you come to understand guide your engagement. And, of course, you don't just ask these questions once at the beginning. As the evaluation unfolds, it's important to evaluate how well the tasks are getting done (quality of work being completed) and how the relationships are unfolding. Feedback from intended users along the way provides critical guidance about whether more or less interaction and communication is needed to enhance use.

Risk Assessment and Contingency Thinking

Contingency thinking for situational responsiveness inevitably involves some degree of risk. Designing an evaluation involves an informal cost-benefit analysis in which potential benefits—for example, using results to improve the program—are considered in relationship to costs, which include financial resources, evaluator and staff time, and opportunity costs (what else could have been done with the money spent on evaluation). Introducing the notion of risk into evaluation design and relationship decisions is a way of acknowledging that things seldom turn out exactly the way they are planned. Calculating risk involves asking the following kinds of questions:

1. What can go wrong in this evaluation?

2. What is the likelihood that it would go wrong?

3. What are the consequences and how bad would they be if things go badly?

The intent of such front-end risk assessment, done with primary intended users, is *not* to deepen the illusion that one can anticipate and thereby prevent all difficulties. Rather, it

is to lay the foundation for contingency thinking as a basis for evaluator-user negotiations and revisions as the evaluation unfolds. Risk analysis should push evaluators and intended users to be prepared for contingencies. Contingency thinking and planning acknowledges the reality that every design will run into execution problems. What distinguishes one evaluation from another is not the absence of problems but the preparation for and ability to solve them. Examining what can go wrong should include thoughtful consideration of what can really be accomplished with available resources.

Risk analysis requires evaluators and stakeholders to become explicit about different scenarios and how they might behave in each.

Three Types of Risk

Risk is traditionally defined as the probability of an occurrence multiplied by the severity of the consequences associated with the hazard. In dialogues with intended users, it can be helpful to break the "what can go wrong?" question into three interdependent categories: idea risk, implementation risk, and evidence risk. Idea risk increases when new, untested, and innovative ideas are being tried out. Implementation risk increases when new organizations or inexperienced staff are implementing a program, and where the environment in which implementation is occurring is turbulent (for example, politically unstable or an arena of conflict). Evidence risk increases when data will be hard to collect or access to needed data is problematic. Two of these three types of risk—idea risk and implementation risk—derive from classic program design risk analysis. The premise is that the more risky the intervention (either because of idea risk or implementation risk, or both), the more uncertain may be the evaluation situation and therefore the more risk that could be entailed in conducting the evaluation due to those uncertainties. The third type of risk, evidence risk, is a fundamental evaluation issue. However, this is a risk shared by the program and the evaluation because the harder it is to evaluate a program, the more that program may be at risk of losing funding or other support. Exhibit 4.8 compares the three kinds of risk for programs and evaluations.

The process of risk analysis should reveal instances in which what is at risk is not just wasted money or useless findings but includes the relationship between the evaluator and intended users by failing to converse openly and honestly about actual and potential problems.

What's Worth Knowing

Focusing an evaluation involves figuring what's worth knowing *in this situation*. Situation analysis, conducted jointly by the evaluator and primary intended users, involves strategic contingency thinking and learning to be active-reactive-interactive-adaptive as a foundation for ongoing situational responsiveness.

EXHIBIT 4.8

Risk Assessment

Nature of Risk	Program Risk Assessment	Evaluation Risk Assessment
Idea/design risk	How clear, well tested, and logical is the intervention idea?	How routine is the evaluation design? How accepted and valid are the measurement approaches?
Implementation risk	What are the challenges to implementing the idea?	What are the challenges to implementing the evaluation design?
Evidence risk	How difficult will it be to evaluate the effectiveness of the idea and/or its implementation?	What are the threats to the evaluation's credibility, utility, feasibility, accuracy, and propriety?

You can't know everything about a situation—and you can't figure it all out at the beginning. Perfect knowledge is not the goal. Situational and contextual sensitivity is. You work with primary intended users to learn the most important things that are likely to affect evaluation use. That's the focus.

Distinguishing informational wheat from chaff requires determining what's important. The challenge of making such a distinction is nicely illustrated by a story about the founder of the Ford Motor Company. A visitor to Ford's factory encountered the famous Henry Ford himself while being a given a tour of the factory. Looking at a car being built, Ford told the visitor authoritatively: "There are exactly 4,719 parts in that model." The visitor was subsequently introduced to the engineer who oversaw production and, having been impressed that the president had a grasp of such details, reported what Henry Ford had said. The engineer shrugged, clearly unimpressed, and said, "I don't know if that's true, but I can't think of a more useless piece of information" (Fadiman & Bernard, 2000, p. 210).

Figuring out what information will be useful and then delivering that information to the people who can use it is the challenge of utilization-focused evaluation. Having developed a deeper understanding of the situation and context, the next step involves determining the priority purpose of the evaluation.

"Juggling four pins at once is easy with practice. The hard part is all the other stuff you have to watch out for."

PRACTICE EXERCISES

1. *Practice conducting a situation analysis.* Identify a program for which you might design an evaluation. Use Exhibits 4.1 and 4.4 to conduct a beginning situation analysis. Use Exhibit 4.5 to identify any special conditions for the evaluation.

2. *Conduct a utilization force field analysis.* Use Exhibits 4.2 and 4.3 to undertake a force field analysis of an evaluation. What factors are likely to be positive forces that support and enhance evaluation use? What factors are likely to create resistance or barriers to use? How might the positive forces be used to minimize the negative forces?
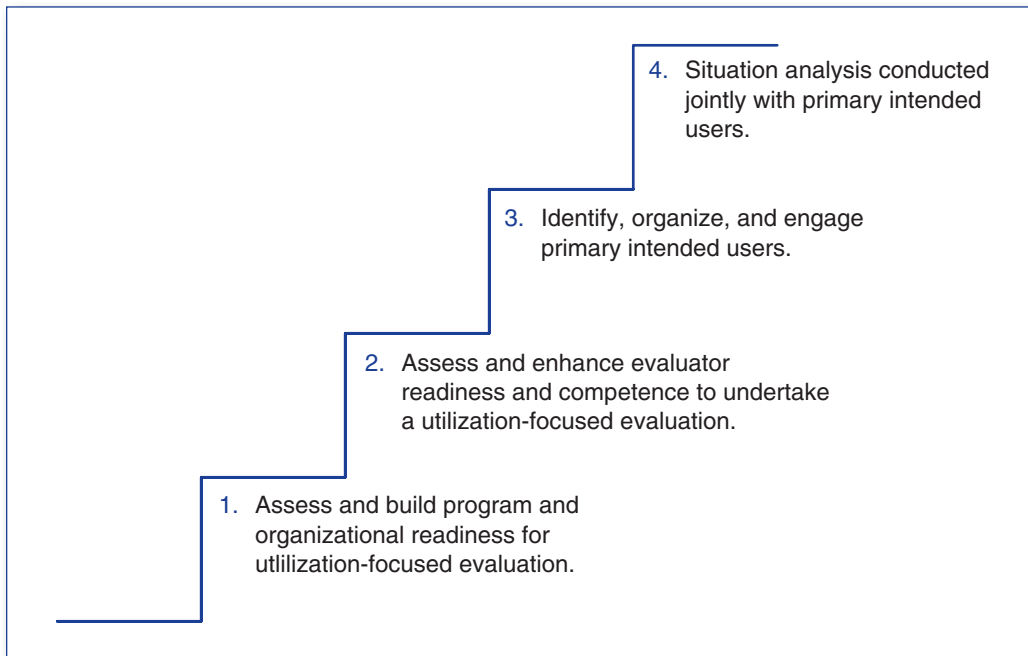
(Continued)

(Continued)

3. *Write a job description for an internal evaluator.* First, describe an organization that you know (or fabricate an organization). Specify the organization's mission, staff size, organizational structure (e.g., different units or program areas), and the challenges the organization is currently facing. Now, write a job description for an internal evaluator in this organization. Include specification of where the internal evaluator will be located, what the evaluation priorities will be, what relationship qualities will be important, and what tasks must get done. Give a rationale for your job description by explaining how the job description is attuned to and appropriate for the organizational situation.

4. *Evaluation risk assessment.* Identify an example of an innovative program that has been evaluated or that has an evaluation design. Use Exhibit 4.8 to discuss evaluation risks in relation to program risks. Elucidate the relationship between program risks and evaluation risks using a concrete program example.

**INTERLUDE, STEPS 1 THROUGH 4.
COMPLEX DYNAMIC SYSTEMS INTERCONNECTIONS:**

**Assessing the Alignment Among the Steps and
Integrating the Situation Analysis From Steps 1 Through 4.**

The steps in the checklist are necessarily linear and sequential.

4. Situation analysis conducted jointly with primary intended users.

3. Identify, organize, and engage primary intended users.

2. Assess and enhance evaluator readiness and competence to undertake a utilization-focused evaluation.

1. Assess and build program and organizational readiness for utlilization-focused evaluation.

Details about what is involved in each step are provided in the summary *U-FE Checklist* in the concluding chapter. (See pages 406–411.)

But to depict utilization-focused evaluation as a complex adaptive system, each new step in the checklist also becomes another element in the complex dynamic system graphic that is emerging as we proceed through the steps of the checklist. This graphic attempts to portray the interdependence of and interactions among the steps.

Complex Dynamic System
Interactions Among U-FE Steps 1 Through 4

Situation analysis

4. SITUATION

ANALYSIS

4. SITUATION

ANALYSIS

1. Assess program readiness, capacity, and commitment

2. Assess evaluator readiness, capacity, and commitment

3. Identify, organize, and engage primary intended users