

## Empowerment Evaluation: Principles and Action

*Abraham Wandersman, Dana C. Keener,  
Jessica Snell-Johns, Robin Lin Miller,  
Paul Flaspohler, Melanie Livet-Dye, Julia Mendez,  
Thomas Behrens, Barbara Bolson,  
and LaVome Robinson*

Empowerment evaluation (EE) aims to increase the likelihood that programs will achieve results by increasing the capacity of program stakeholders (any individual, group, or organization that has an important interest in how well a program functions) to plan, implement, and evaluate their own programs. This chapter defines EE, proposes a set of principles for the theory and practice of EE, and presents a case example to illustrate EE in action.

### **Getting to Know Empowerment Evaluation**

Although EE is relatively new (it was introduced in 1992), it is gaining acceptance in mainstream evaluation circles. Since the publication of the seminal book on EE (Fetterman, Kaftarian, & Wandersman, 1996), much work has been done in the name of EE. In addition to scholarly contributions and numerous evaluations that labeled themselves as EE, the creation and establishment of the American Evaluation Association's Collaborative, Participatory, and

---

The evaluation of the Night Ministry was made possible by an evaluation team from the University of Illinois at Chicago, which was made up of Robin Miller and several graduate students: Barbara J. Bedney, Terri Bicok, Khari Hunt, Mary Murray, Bernadette Sanchez, and Dana Wardlaw. In addition, we acknowledge leaders and staff from the Middle Tyger Community Center and the Foundation for the Future Initiative—both of Spartanburg, South Carolina—who through their collaboration and commitment to evaluation greatly informed our ability to identify and articulate the empowerment evaluation principles. Finally, we are grateful to Rebecca Campbell, Margaret Davis, David Fetterman, Miles McNall, and Shakeh Kaftarian for their valuable feedback on the chapter.

Empowerment Evaluation topical interest group has provided fertile ground for extended debate and discussion about the EE approach.

### *Definitions of Empowerment Evaluation*

Fetterman (2001) defined EE as “the use of evaluation concepts, techniques, and findings to foster improvement and self-determination” (p. 3). Although this definition of EE has remained consistent since the onset of the approach, the methods and principles of EE have continued to evolve and become more refined over time. Wandersman’s description of EE places an explicit emphasis on results:

The goal of empowerment evaluation is to improve program success. By providing program developers with tools for assessing the planning, implementation, and evaluation of programs, program practitioners have the opportunity to improve planning, implement with quality, evaluate outcomes, and develop a continuous quality improvement system, thereby increasing the probability of achieving results. (1999, p. 96)

### *Purposes of Empowerment Evaluation*

EE expands the purpose, roles, and potential settings of evaluation beyond that of traditional evaluation approaches.<sup>1</sup> However, this does not preclude the need or diminish the importance of more traditional evaluation approaches (Fetterman, 2001). The type of evaluation selected for a given program is best determined by the goals and purposes of the evaluation. Therefore, neither EE nor traditional evaluation is inherently a “better” approach. Instead, each evaluation approach is valuable when applied for purposes that are well-suited to the strengths and intentions of the given approach (Chelimsky, 1997; Patton, 1997).

Chelimsky (1997) described three purposes of evaluation, including: (a) evaluation for development (information collected to strengthen institutions); (b) evaluation for accountability (measurement of results or efficiency); and (c) evaluation for knowledge (acquisition of a more profound understanding in some specific area or field; p. 10). EE is especially appropriate for the purpose of development, because it actively seeks to develop people, programs, and institutions. In addition, EE can also be used effectively for the purpose of accountability, as stated in Wandersman’s definition of EE. The frameworks and methodology (e.g., Prevention Plus III, Getting to Outcomes) created and used by Wandersman and colleagues (e.g., Linney & Wandersman, 1991; Wandersman, Imm, Chinman, & Kaftarian, 2000) are designed not only to improve programs but to measure the results of programs (accountability). Although it

---

<sup>1</sup>Traditional (or independent) evaluation is characterized by greater autonomy of the evaluator, the use of controlled research methods, and is most often used for the purpose of accountability (Rossi, Freeman, & Lipsey, 1999).

**Table 8.1.** Ten Principles of Empowerment Evaluation

Core values	Principle 1. EE aims to influence the quality of programs. Principle 2. In EE, the power and the responsibility for the evaluation lies with program stakeholders. Principle 3. EE adheres to the evaluation standards.
Creating a culture that is ready and interested in improvement	Principle 4. Empowerment evaluators demystify evaluation. Principle 5. Empowerment evaluators emphasize collaboration with program stakeholders. Principle 6. Empowerment evaluators increase stakeholders' capacity to conduct evaluation and to effectively use results. Principle 7. Empowerment evaluators use results in the spirit of continuous quality improvement.
EE is a cyclical and developmental process	Principle 8. EE is helpful at any stage of program development. Principle 9. EE influences program planning. Principle 10. EE institutionalizes self-evaluation among program staff.

is not a current emphasis of EE, in the future we hope to systematically explore the potential usefulness of EE for the third purpose of research knowledge.

### Guiding Principles of Empowerment Evaluation

EE is characterized by principles that represent EE's stance on evaluation ideology and practice. Table 8.1 represents one way to organize the guiding principles of EE. The first category consists of core values that are central to the philosophy and practice of EE. The second category consists of principles that relate to creating a culture that is ready and interested in improvement. The third category includes principles that illuminate how EE is a cyclical and developmental process. It is important to note that: (a) the principles represent ideals; (b) the individual principles are not exclusively associated with EE; and (c) it is the set of principles, taken as a whole, that distinguishes EE from other approaches.

#### *Core Values of Empowerment Evaluation*

When selecting any evaluation approach, it is important to ensure that the core values of the selected approach match the needs of the program and its stakeholders. The core values of EE are that it aims to influence the quality

of programs, that the power and the responsibility for the evaluation lies with the program stakeholders, and that it adheres to the evaluation standards.

**PRINCIPLE 1: EMPOWERMENT EVALUATION AIMS TO INFLUENCE THE QUALITY OF PROGRAMS.** EE values program success. Accordingly, the EE approach seeks to increase the likelihood that programs will achieve their desired outcomes. This is in contrast to traditional evaluation, which values neutrality and objectivity and wants to examine programs in their "natural state." A traditional evaluator tries to avoid influencing a program's outcomes, and an empowerment evaluator strives to positively influence a program's degree of success. Some argue that because empowerment evaluators are not neutral, evaluation findings are more likely to be inaccurate because of the misrepresentation of data, biased research questions, or misinterpretation of results. It is important to recognize that the integrity of an evaluation can be compromised in any setting (e.g., the 2002 ENRON scandal) or evaluation approach. Practitioners working with empowerment evaluators may actually be less likely to misrepresent data than practitioners who feel threatened by evaluation. Empowerment evaluators propose that because they strive for program improvement, they may actually be *more* critical than traditional evaluators (Fetterman, 2001).

**PRINCIPLE 2: IN EMPOWERMENT EVALUATION, THE POWER AND RESPONSIBILITY FOR EVALUATION LIES WITH THE PROGRAM STAKEHOLDERS.** Typically in traditional evaluation, decisions regarding the purpose, design, and use of evaluation results are made by the evaluator and the funder. Alternatively, empowerment evaluators work to blur boundaries that traditionally separate funders, practitioners, and evaluators (Yost & Wandersman, 1998) by giving each group a voice in the decision-making process. Although empowerment evaluators share ideas and provide expert guidance, the stakeholders ultimately make critical decisions about the evaluation, conduct the evaluation, and put the evaluation findings to use. In participatory evaluation designs, decision making is shared by both evaluators and practitioners (Cousins & Whitmore, 1998), whereas in EE, the practitioners are explicitly the ones with the decision-making power.

**PRINCIPLE 3: EMPOWERMENT EVALUATION ADHERES TO THE EVALUATION STANDARDS.** Although the philosophical underpinnings of EE are quite distinct from traditional evaluation approaches, the principles of EE are fully consistent with the standards of evaluation set forth by the Joint Committee on Standards for Educational Evaluation (1994; Fetterman, 2001). The standards serve to provide a common language and set of values for the field and ensure quality across all evaluation philosophies and methodologies. The utility standards are designed to ensure that evaluation serves the information needs of the intended users. The feasibility standards are designed to ensure that evaluation is realistic, prudent, diplomatic, and frugal. The propriety standards are designed to ensure that evaluation is conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by the results. The accuracy standards are designed to ensure that evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated.

*Creating a Culture That Is Ready and Interested in Improvement*

One of the central aims of EE is to create a culture that is ready and interested in improvement. In EE, teaching stakeholders to value opportunities for program improvement is as central to building capacity and achieving results as teaching stakeholders' specific evaluation skills (e.g., logic modeling, survey administration, data entry). Fetterman stresses the importance of "creating a dynamic community of learners—a community where people are willing to share both successes and failures, to be honest, and to be self-critical" (2001, p. 6). We believe that this type of evaluation culture is established when evaluators demystify the process of evaluation, work collaboratively with stakeholders, build stakeholders' capacity, and emphasize that results will be used in the spirit of continuous quality improvement.

**PRINCIPLE 4: EMPOWERMENT EVALUATORS DEMYSTIFY EVALUATION.** Before specific steps can be taken to increase stakeholders' capacity and to influence programming, explicit attention must be given to stakeholders' concerns about evaluation. The Northwest Regional Educational Laboratory (1999) highlights common fears about evaluation, including general anxiety among program staff, uncertainty about how to conduct evaluation, and misuse and misunderstanding of evaluation findings, especially by program opponents. Program stakeholders often have good reason to be fearful of evaluation when they lack a role in the evaluation process or when their survival depends on the outcome (e.g., Bicknell & Telfair, 1999). Such fears lead practitioners to feel uneasy about revealing problems and challenges they inevitably face in implementation, which potentially compromises the accuracy and utility of the evaluation results.

The effective empowerment evaluator seeks to overcome these fears and concerns by demystifying evaluation. This is accomplished by using a structured framework to explain the logic and process involved in evaluation. The three steps proposed by Fetterman (2001, pp. 23–33) and the 10 accountability questions proposed by Wandersman et al. (2000) are examples of question-based frameworks that guide stakeholders to make critical decisions about essential elements of effective programs (including evaluation). By translating evaluation methods into specific questions that can be considered by stakeholders, the complex process of evaluation is made concrete. Empowerment evaluators believe that practitioners are more trustful of evaluation results—whether they are positive or negative—when they share ownership of the information and have both the ability and responsibility to use the information to improve their programs.

**PRINCIPLE 5: EMPOWERMENT EVALUATORS EMPHASIZE COLLABORATION WITH PROGRAM STAKEHOLDERS.** One way empowerment evaluators reduce fears and facilitate stakeholders' interest in evaluation is by explaining that EE is not something done by someone, to someone else. Rather, EE is a process in which evaluators work side by side with program stakeholders to implement and evaluate a program in a way that meets the stakeholders' needs. The nature of this collaboration is such that the boundaries traditionally separating funders,

practitioners, and evaluators are intentionally blurred (Yost & Wandersman, 1998). Empowerment evaluators and stakeholders are seen as sharing a common purpose, and all stakeholders are seen as contributing something unique and valuable to that common purpose.

**PRINCIPLE 6: EMPOWERMENT EVALUATORS BUILD STAKEHOLDERS' CAPACITY TO CONDUCT EVALUATION AND TO EFFECTIVELY USE RESULTS.** EE is an ongoing process of building stakeholders' capacity. EE creates a culture that is ready and interested in improvement by tailoring support to the current capacities of stakeholders. Stakeholders come to trust that if they need assistance, their empowerment evaluators are available to guide them. At the same time, empowerment evaluators are eager for stakeholders to operate independently.

Empowerment evaluators teach stakeholders to use techniques that lead to the design and implementation of more effective programs. It is assumed that developing practitioners' abilities to use evaluation to inform decision making will later translate into the use of evaluation to benefit all program activities. Patton (1997) summarized capacity-building as "individual changes in thinking and behavior, and program or organizational changes in procedures and culture, that occur as a result of the learning that occurs during the evaluation process" (p. 90). Snell-Johns and Keener (2000) defined evaluation capacity as the ability to understand and perform skills related to assessing the implementation and effectiveness of a given program and the ability to make changes to this program based on the information gained. Empowerment evaluators believe that when stakeholders know the steps involved in conducting a high-quality evaluation, they are in a better position to understand and use evaluation results, which makes them ready and interested in program improvement.

**PRINCIPLE 7: EMPOWERMENT EVALUATORS USE EVALUATION RESULTS IN THE SPIRIT OF CONTINUOUS QUALITY IMPROVEMENT.** Empowerment evaluators encourage stakeholders to value both positive and negative results; positive results are celebrated and negative results are seen as crucial for additional program development. Negative results are not to be feared, especially in circumstances where the funder also values EE. Because EE emphasizes the use of process and outcome data for program improvement, rather than simply for auditing purposes, practitioners are better able to trust that negative results will not be used arbitrarily or to punish them. By emphasizing that evaluation results will be used for continuous quality improvement, evaluators are able to help establish a dynamic community of learners. Stakeholders learn the value of being honest with one another and with evaluators, and they come to respect and admire others' willingness to be self-critical, perhaps as much as they may have initially valued the discovery of positive results.

### *Empowerment Evaluation Is a Cyclical and Developmental Process*

If one considers evaluation and evaluation capacities as existing along a logical developmental continuum, it becomes clear why it is not only possible but

necessary for empowerment evaluators to take on different roles and incorporate various techniques at different points in the life cycle of a program (Fetterman & Eiler, 2001). A key task in implementing EE involves assessing the evaluation capacity of program personnel and their readiness for change. This awareness sets the stage for an empowerment evaluator to choose particular methods to move the program from its current status to the next step along the developmental continuum. Progress along this continuum occurs each time a stakeholder takes greater control over the evaluation design, process, or use of evaluation results.

**PRINCIPLE 8: EMPOWERMENT EVALUATION IS HELPFUL AT ANY STAGE OF PROGRAM DEVELOPMENT.** EE has the explicit value of working with people and programs "where they are at" to move them forward. For example, if a program's staff has never conducted an evaluation before, getting the staff to simply identify their program's target population, goals, and desired outcomes is seen as a legitimate step toward better programming and evaluation. Although it is optimal to begin EE during the early stages of program development, EE can benefit mature programs as well. For example, if a program is in the middle of implementation, EE methods and tools can be used to help assess the quality of implementation. If a program has already been completed and an empowerment evaluator is hired, the evaluator can teach the program stakeholders how to understand the results and how to translate these results into plans for improving the program the next time it is offered. Ideally, empowerment evaluators work with programs and organizations for several years so that a cycle of improvement can be created.

**PRINCIPLE 9: EMPOWERMENT EVALUATION INFLUENCES PROGRAM PLANNING.** It is a truism that programs that are planned better work better. Planning is a necessary, although not sufficient, element of an effective program. The most important questions to answer when developing a good program plan are also relevant to program evaluation. For instance, both a program plan and an evaluation process should identify the goals, target populations, and strategies of a given program. The unfortunate reality is that programs are frequently implemented without giving adequate attention to clearly identifying the needs, target populations, and goals. Often programs are chosen just because they seem like a good idea and not because they are addressing a specific, identified need. Thus, under ideal circumstances, EE begins as soon as an agency decides it wants to address a problem in the community. This means that empowerment evaluators can play a role as early as the grant-writing stage (Yost & Wandersman, 1998). EE can examine the logic or theory of a program before it even gets off the ground, saving time and energy that might be misguided or misused. In doing this, the evaluation process serves as a guide not only for evaluating the program but also for program planning and implementation.

**PRINCIPLE 10: EMPOWERMENT EVALUATION INSTITUTIONALIZES SELF-EVALUATION AMONG PROGRAM STAFF.** As the capacities of the program grow and develop, empowerment evaluators shift from the role of teacher to more of a "critical friend" in the evaluation process (Fetterman, 2001). If EE is successful,

the techniques of evaluation become a part of regular program activity and influence the overall quality of programming. For evaluation to have its optimal impact, it needs to be an ongoing process, which allows EE to become a part of the culture and daily life of a program. In other words, the goal is for EE to eventually become institutionalized within the program setting (Fetterman, 2001). By participating in EE, program stakeholders learn to see the intervention from an evaluator's perspective. This can have a greater, and more durable, impact than the results of a particular evaluation. Fetterman (2001) suggested that institutionalization and the development of a dynamic community of learners allows EE to accommodate changes in the program environment caused by shifting populations, goals, knowledge, and external forces impinging on the program.

### Case Example

The principles of EE are inherent in the ideals of EE, but what does EE really look like? The growth in EE's popularity has resulted in its application in a wide variety of settings (e.g., school readiness initiatives, family resource centers) and with programs of differing scales, ranging from well-funded evaluations of multi-million-dollar, statewide initiatives (e.g., Wandersman et al., 2001) to unfunded evaluations of small local community agencies and programs.

In January 1996, Robin Lin Miller began working as an evaluator with the Night Ministry, a unique faith-based organization located in Chicago. Miller's experience is presented as an illustration of EE's application in a real-world setting. Following is a description of the Night Ministry organization, the conditions that led the evaluation team to choose an EE approach (Fetterman, 1994, 1996; Vanderplaat, 1995), and the process of conducting the evaluation. To conclude, the evaluation effort is examined in terms of the extent to which it conforms to the 10 defining principles of EE.

#### *The Night Ministry: The Organizational Setting*

Founded in 1976, the Night Ministry is a nondenominational church-based organization that provides physical, emotional, and spiritual services to Chicago's nighttime street communities. Male and female prostitutes, homeless adults and youth, chronically mentally ill individuals, disenfranchised sexual minorities, and substance users are among the Night Ministry's congregants. Locations such as bars, restaurants, liquor stores, street corners, adult bookstores, bathhouses, and parks are its parishes. The Night Ministry reaches out to individuals at the margins of society through three programmatic efforts: (a) a street and health outreach program (the Outreach Health Ministry); (b) a 16-bed emergency shelter for youth and their children (the Open Door Youth Emergency Shelter); and (c) a city-wide partnership of organizations providing emergency shelter to youth (the Youth Shelter Network).

The evaluation activities described in this example were focused on the Outreach Health Ministry (OHM). In OHM, ministers have street parishes in



several communities throughout the greater Chicago area. Ministers spend time in local parishes (e.g., bars, strip clubs, street corners) providing companionship, support, counseling, and referrals to parishioners. The program also operates a bus that travels to each parish between 7 p.m. and 2 a.m. offering hospitality, condoms, clothing, toiletries, health care, STD screening, counseling, and companionship. A minister, a nurse, and a cadre of volunteers staff the bus that travels to about three parishes each night on a regular schedule known to parishioners.

The Outreach Health Ministry is based on the idea that promoting personal and spiritual growth will lead to improved quality of life. Quality of life improvements can be psychological, spiritual, physical, or circumstantial (e.g., getting a job). The program seeks to improve quality of life outcomes by developing relationships with nighttime community members that are characterized by respect and dignity and that provide people with a sense that they are valued and supported. The development of these relationships occurs over long periods of time—in some cases many years. OHM seeks to empower its parishioners through a process of personal transformation and validation. Interactions among community members, staff, and volunteers are reflexive in nature. In other words, staff members believe that personal and spiritual growth evolves through a constant process of reflection and reassessment.

According to the program's proponents, the Night Ministry and its Outreach Health Ministry program represent a radical vision of ministry. OHM puts its faith into action in the nonjudgmental way that they believe Christ might have if He were alive today—going to the people, tailoring ministry to the individual and his or her context, and transforming the role of the minister to that of a mutual learner who is equal to parishioners. Although the program has much in common with traditional outreach and other faith-based initiatives, the Night Ministry asserts that the way it has brought these elements together is unique. It is this nontraditional, fresh approach to pastoral care—an approach that departs from business as usual within the church—that defines the organization's identity.

### *The Empowerment Evaluation*

The EE project with the Night Ministry began in response to the ministry's desire to improve the evaluation process of its Outreach Health Ministry as well as the evaluation skills of its staff. The organization was greatly dissatisfied with the type of data they were previously required to collect, because they believed that the data failed to communicate the program's identity and practice in a meaningful way. In March 1996, Miller attended a meeting of board members and staff to identify the organization's evaluation needs and the target audience of the evaluation results. Later, an evaluation team was formed to include Miller, a team of graduate students from the University of Illinois at Chicago (UIC), ministers, nurses, and divinity student interns. By group consensus, parishioners were not included in the evaluation team, although they were integral to the evaluation process.

A notable feature of this evaluation was that it was a volunteer effort (i.e., not funded). This is despite the fact that funding was offered by a foundation

midway through the project. In this case, the evaluation team chose to decline the funding offer so that the project could proceed unfettered by the timelines and accountability requirements often dictated by funding institutions.<sup>2</sup>

When considering the type of evaluation to use for this project, the EE approach emerged as most consistent with the Night Ministry's culture, values, and practices. By matching the evaluation approach to the values and practices of the program, the evaluation team established a means to overcome passive disinterest among staff and, in some cases, active dislike of evaluation. For instance, EE's emphasis on self-reflection and self-evaluation is also inherent in the discipline of pastoral care. All ministers in the program keep personal diaries to record parishioners' stories and their own experiences on the streets, as well as to process challenging emotions provoked by their work. Such creative freedom to identify and solve problems and the autonomy it implies is one organizational pathway to empowerment (Foster-Fishman, Salem, Chibnall, Legler, & Yapchai, 1998).

Egalitarianism and collaboration are two additional values intrinsic to both the Outreach Health Ministry program and the EE approach. These values are consistent with collectivist and strength-based organizational characteristics identified by Maton and Salem (1995) as empowering. For example, OHM uses storytelling as a way for ministers and parishioners to share themselves with one another and to level perceived inequities in relationships. Personal disclosure is encouraged and considered an essential technique to blur the boundary between personal and professional. Parishioners collaborate with the staff by providing volunteer services on the bus. Volunteers counseled, nurses ministered, and ministers served coffee. The staff displayed a strong sense of community, approaching most aspects of the program as a collective.

Both EE and the Night Ministry strive to encourage social justice, a feature of empowering organizations (Maton & Salem, 1995). The Night Ministry is founded on the principle that justice belongs to all and has challenged other congregations, local citizens, and government representatives to advocate for the rights of homeless, mentally ill, and other individuals who are socially disenfranchised. For example, the Night Ministry established a policy advocacy coalition of providers to bring the needs and concerns of homeless and runaway youth, who are not wards of the state, to the attention of local and state government officials.

### *Selecting Appropriate Evaluation Strategies*

The ultimate challenge faced by the evaluation team was to develop an evaluation strategy that could capture the quality of interactions with the diversity of nighttime community members, the longitudinal nature of relationship development, and the variety of desired outcomes, while simultaneously helping the staff learn to develop their own evaluation tools, collect data, and use

---

<sup>2</sup>As funding agencies become more knowledgeable of EE and the benefits of this approach, they are more likely to be flexible and supportive of the methods and timelines necessary to build capacity and implement an empowerment evaluation (W. K. Kellogg Foundation, 1999).

evaluation results to inform program changes. To accomplish these aims, the evaluation team first had to develop a rich description of what actually happened in the program. This understanding was necessary before they could help the organization develop useful evaluation tools, determine the most important evaluation questions, and assess the extent to which evaluation was already part of the fabric of organizational life.

Another task was to create a framework by which tools could be developed in collaboration with the staff so that evaluation skills were cultivated within the organization and so that the logic or small theory of the program was evident. The team wanted to capitalize on the reflexivity of staff (i.e., ability to be reflective), both in the process of transmitting skills and in building a system for long-term evaluation of the program.

In accomplishing all tasks, the evaluation team sought to work in ways that did not disrupt or violate the relationships with the parishioners that the organization had worked so long and hard to develop. Many of the parishioners are highly mistrustful of formal institutions, given their association with illegal activities (e.g., prostitution, drugs) and their marginalized position in society. The university-based evaluation team avoided procedures that community members might associate with the formal institutions that they fear and dislike. An organizational ethnographic approach was used to understand the meaning of the program. The ethnographic paradigm was well-suited to the goals of the evaluation team because its naturalistic research strategies enabled researchers to gain experiential knowledge of a phenomenon and close understanding of what objects, activities, events, and relationships mean to people. The ethnographic paradigm is reflexive in ways that paralleled the program's style. For instance, descriptive accounts of places and events simultaneously give rise to those phenomena and are shaped by those phenomena. Participant observation was the primary means of data collection, supplemented by interviews and document reviews. Participant observations were seen as least disruptive to the operation of the program. This method also allowed the evaluation team to understand the various roles within the ministry while portraying the longitudinal nature of relationships within the program. The entire evaluation team functioned in the dual roles of participant and observer from April 1996 to June 1998. Each team member was responsible for volunteering on a particular night of the week with the bus and documenting the interactions among the staff, volunteers, and community members. By systematically covering each night of the week, the evaluation team was able to observe the different shifts and program participants and gain a complete picture of program operations as they unfolded over time. By having a consistent schedule, each team member also had the opportunity to build relationships with community members at the various stops on that night's itinerary.

While on the bus, team members served coffee and cookies, distributed condoms and clothing, counseled parishioners, and worked alongside the staff. Team members also shadowed ministers on walks through their parishes. These activities provided the team with direct access to the program and insight into the challenges of measuring it. It also led staff to perceive the evaluation team members as standing alongside them on a journey toward transformation, much like staff hoped to be perceived by parishioners.

### *Initial Observations*

The evaluation team's initial observations and informal interviews provided important insights about the program's operation and how its philosophy of practice was enacted. Some initial observations had particular bearing on the evaluators' subsequent actions. One was that program decision making was handled democratically. This was evident at staff meetings and during program implementation. Student interns, volunteers, nurses, and ministers were vested with equal responsibility and authority for program operations. All staff shared in creating a sense of community on the streets each evening. Decisions were based on everyone's experiences, rather than only favoring the perspectives of program management. As a result, staff members shared a sense of ownership of the program.

There were, however, drawbacks of the democratic and nonhierarchical process. First, the value placed on group decision making and full consensus made it difficult to make more than small changes to the program. Second, because everyone was allowed to put forth their views so freely, over time the attitudes and beliefs of some individuals came to characterize those of the group, whether those beliefs were valid or not. Some of the collective myths observed among staff included the belief that the program could not be evaluated because of its spiritual base and that evaluation findings suggesting program failure were a result of a misunderstanding of the program. Such myths served as armor to protect the staff from accountability demands. Through these myths, the staff had found a way to accept an operating program model that departed from the desired program model. Freidman (2001) referred to this phenomenon, in which groups develop defensive routines, as designed blindness. Thus, the democratic culture had serious implications for evaluation, including defining who the evaluation stakeholders were, what was believed to merit evaluative scrutiny, and how evaluation data were to be processed by staff.

Another important observation made by the evaluation team was that many factors beyond the staff's control heavily affected (or interfered with) the staff's ability to implement the program as intended on a day-to-day basis. These factors included the presence and timing of police sweeps; the burdens on the shelter and drug treatment systems, as well as the restrictions imposed on those systems by regulatory agencies; the time of the month in relation to the distribution of government support payments; the impact on traffic and public behavior caused by events such as the Bulls winning an NBA championship; and the weather. Recognition of these factors highlighted the importance of viewing the program as a series of interactions that unfolded over time, rather than as a discrete interaction between a parishioner and the program. Indeed, the reason that the evaluation team earned trust where other evaluators had failed was because staff believed that they would not be judged for implementation barriers that were outside of their control.

Another crucial observation was that staff had difficulty describing the purpose of the program, how they wanted the program to function, and how specific program activities led to particular outcomes. Indeed, staff rarely talked about outcomes. When asked to describe a success, staff would describe the

same case that they had just offered moments before as a failure. This conceptual muddiness was related to the underlying pastoral philosophy of the program. Pastoral aims, such as spiritual transformation and enlightenment, were not obviously connected to secular aims such as using a condom, kicking a habit, or seeking employment. The staff saw an individual's willingness to change his or her behavior as intricately linked to the spiritual self, a perspective that many funding institutions do not necessarily share. In OHM, trusting relationships were seen as both a means and an end, such that obtaining the means but not the ends could be seen as a success and a failure simultaneously. Getting someone to come back and hang out four or five times over the course of a few months was a success, even if they seldom spoke or were never sober on any of those occasions. Staff members were not driven by a rigid set of outcomes that were invariant across individuals. Achievements were grounded in relationship achievements and in the ebb and flow of the various dyadic exchanges in which parishioners and the Night Ministry personnel engaged.

### *Building an Improvement-Oriented Culture*

After 6 months of observation and participation in program activities, the evaluation team developed a simple logic model of the theory underlying the Outreach Health Ministry program. During a group exercise, the model was presented to the OHM staff for group reflection and feedback. After lengthy discussion, the group agreed that the model accurately represented the intended goals and strategies of the program.

Next, the OHM staff identified the practices associated with each of the model's components. Staff worked in small groups to describe the required activities in each area of the model, creating program templates (Scheirer, 1996) for the logic model. For example, the logic model contained an element titled "build quality relationships." Staff identified behaviors that were necessary to develop a relationship with a parishioner, such as using open body language, being nonjudgmental, remembering personal information about people, and engaging in active listening. The model provided the theoretical guide for the evaluation team's evolving participant observations and served as the initial template for decision making regarding how the program was to be evaluated.

The staff used the logic model and program templates as the basis for generating key questions about program implementation. Staff teams worked together over a period of several months to develop evaluation tools for documenting program processes related to the evaluation questions, each of which loosely correspond to the domains of the logic model. Each small team included a nurse, at least one minister, and a student evaluator. Some teams also included ministerial interns. Teams regularly presented their ideas to the other teams for feedback. Each team ultimately produced at least one measure that was pilot-tested by the staff. Staff then worked through a series of revisions to the measures and procedures for collecting the data. These measures included forms to document the characteristics of the parishes in which the work was carried out, services provided by the bus, and interactions with parishioners in each of the parishes.

### *Using Results to Build Capacity in the Spirit of Continuous Improvement*

During this same period, the evaluation team regularly presented information about what they had learned. For example, the team developed a flow chart that described how interactions between parishioners and staff are initiated and the sequence of interactions that typically follow them. Data revealed that parishioners initiated conversations with staff more often than staff initiated conversations with parishioners and that staff avoided some approaches. Staff was curious to understand the reasons for these findings, resulting in new evaluation questions. The presentations also proved to be a vehicle to promote critical evaluative thought among program staff. It was through the presentations of ethnographic data that staff was best able to stand outside themselves and get a new perspective into the program operations, program assumptions, and the inconsistencies between program practice and program design.

EE pushed the organization to confront its own myths. One example of this occurred at a staff meeting in which the evaluation team reviewed the data that staff had collected during street interactions over a two-month period. The data indicated that the staff had more contacts (i.e., superficial/basic interactions) with women and African Americans/Blacks than with men and Caucasians/Whites. However, inconsistent with the number of contacts, data revealed that a higher proportion of substantive conversations (or more in-depth relationships) occurred with men and Caucasian/White persons. In this same meeting, staff also learned that, despite the perception that they talked about AIDS frequently, only one HIV risk-reduction counseling interaction had been recorded during the two-month period. These two findings, from the EE process, provided staff with surprising revelations about their work and reinforced the importance of collecting evaluation data.

Staff generated multiple hypotheses about the surprising findings. One hypothesis was that they had not collected data thoroughly enough. A second was that they had not adequately trained the staff (who were primarily White and male) to address the more in-depth concerns of women and people of color. The hypotheses they developed were testable, suggested clear action, and created opportunities to challenge the program status quo. The data created a forum for women to talk about how the program might include their voices and perspectives and for the ministers to talk about how reporting the content of their interactions with parishioners clashed with their training about the privacy of the pastoral relationship. The ethnographic data suggested clear areas for program improvement. More recent data suggest that the proportion of substantive encounters by gender and race now closely mirrors the proportion of superficial contacts. Contacts with women have increased modestly.

Staff members continue to use the measures that evolved out of the evaluation to inform organizational changes, to make the program more inclusive, and to keep the program from becoming an agent of the status quo. In addition, staff developed—on their own—a way to monitor client outcomes after their formal relationship with the evaluation team had ended, a sign of increased skill. The Night Ministry continues to include the OHM logic model in their grants and uses the model to articulate their vision of ministry to other outside

entities. The Night Ministry honored the evaluation team with a Living the Mission Award in recognition of its success in putting the values of the organization into action. The Night Ministry also asked the evaluation team to remain involved in the organization to assist them in establishing a long-range plan for a training institute, creating a five-year strategic plan, and continuing their empowerment EE efforts.

In an attempt to reflect on the successes and challenges of this particular EE, Revs. Thomas Behrens and Barbara Bolsun of the Night Ministry graded the efforts of the evaluation team based on its adherence to each of the 10 principles of EE. Table 8.2 presents the results of this informal exercise. As the grades indicate, the weakest areas of performance were in increasing capacity (Principle 6) and institutionalizing an evaluation culture (Principle 10), though the grades were still high. Relatively lower grades in these areas can be attributed to nearly 100% turnover rate of OHM staff in 1999 and, simultaneously, to all research at the University of Illinois at Chicago being suspended for 8 months by the U.S. Office for Human Research Protections beginning in August 1999. Overall, however, the evaluation team did well in adhering to the principles (this is especially impressive because the principles were proposed after the project was completed). Because the team took the time to fully understand the program, the staff, and the context in which it operated, the team was perceived by staff to be "clueful" rather than "clueless," placing it in a strong position to demystify evaluation (Principle 4) and collaborate with stakeholders (Principle 5). The evaluation team's thorough understanding of the program, acquired through long-term collaboration, was key to the staff's acceptance of the UIC evaluation team and the team's ability to work effectively with staff. The evaluation team was viewed as part of the program culture. The small-group process was empowering to staff and helped staff articulate a framework for what they do. They used this framework to enrich their own understanding of their work, communicate with others, and develop more rigorous ways of self-examination (Principle 2). Staff acquired the sense that they could describe and measure what they did, so they did not have to rely on the myth that it was not measurable (Principle 4). Staff came to understand the value of data as a tool for planning (Principles 8 and 9) and program improvement (Principles 1 and 7).

### Conclusion

Community psychology advocates collaboration, partnership, sharing, and parity among researchers, evaluators, community members, and organizations. This ideal, however, seems especially challenging to achieve within the realm of program evaluation. This is likely because of the perpetuation of the belief that only the "experts" (i.e., those with formal training) can serve in the role of evaluator. Indeed, program evaluation does involve knowledge and use of research designs, methods, and statistical analyses. However, if experts remain the sole designers of evaluation methods and the primary interpreters of evaluation findings, the potential utility and impact of the evaluation process would be limited. EE shifts the power of evaluation into the hands of stakeholders.

