

## **An Empowerment Evaluation Model for Sexual Assault Programs: Empirical Evidence of Effectiveness**

**Rebecca Campbell,<sup>1,5</sup> Heather Dorey,<sup>2</sup> Monika Naegeli,<sup>2</sup> Lori K. Grubstein,<sup>2</sup> Kelly K. Bennett,<sup>2</sup> Freya Bonter,<sup>2</sup> Patricia K. Smith,<sup>3</sup> Jessica Grzywacz,<sup>3</sup> Patsy K. Baker,<sup>4</sup> and William S. Davidson, II<sup>1</sup>**

---

Using an empowerment evaluation approach, the Sexual Assault and Rape Prevention (SARP) Evaluation Project brought together university researchers, public health evaluators, state funders, and program staff to provide evaluation training and consultation to all state-funded rape prevention and victim services programs in Michigan. In this paper, the specific activities of the SARP project are described, and process and outcome data are presented that address the effectiveness of this empowerment evaluation approach. Based on subjective reports from program staff and objective data obtained from state funders, results suggest that this evaluation model was successful in helping 90% of the prevention programs and 75% of the victim services programs successfully develop and launch program evaluations. One-year follow-up data indicate that 90% of the programs had sustained their evaluation efforts after the formal work of the SARP project had ended.

---

**KEY WORDS:** SARP; empowerment evaluation; participatory evaluation; collaboration; program evaluation; rape; sexual assault.

---

Program evaluator is one of many roles community psychologists may assume in their work with organizations and communities. Although there are many theoretical approaches to conducting program evaluation, participatory methods, such as empowerment evaluation, may be especially useful to community psychologists. Consistent with the values of community psychology, participatory evaluation emphasizes collaboration and community development. In contrast to traditional methods of evaluation whereby evaluators often function independently, participatory methods favor shared power, control, and decision making. Multiple stakeholder groups are brought together to plan and conduct an evaluation. The

voices and perspectives of program staff shape all aspects of the evaluation, from design to utilization.

Empowerment evaluation is a specific participatory evaluation approach whereby the evaluator provides training and consultation to program staff so that they can conduct their own evaluations. The programs—not the evaluator—own and control the evaluation. Empowerment evaluation seeks to build capacity within organizations and promote social change. The parallels between community psychology and empowerment evaluation are striking. For instance, both pose a fundamental challenge to the assumption of objective neutrality in science. As many community psychologists have noted, values permeate all aspects of science (see Campbell & Wasco, 2000; Fawcett, 1991; Kelly, 1970; Rappaport, 1977). Community psychologists and empowerment evaluators try to make their values explicit. In empowerment evaluation, value is placed on empowering organizations through information and skills that can be learned by conducting evaluation. Indeed, Fetterman (1996, 2001a) has drawn empowerment research from

<sup>1</sup>Michigan State University, East Lansing, Michigan.

<sup>2</sup>Michigan Public Health Institute, Okemos, Michigan.

<sup>3</sup>Michigan Department of Community Health, Lansing, Michigan.

<sup>4</sup>Michigan Family Independence Agency, Lansing, Michigan.

<sup>5</sup>To whom correspondence should be addressed at Department of Psychology, Michigan State University, 127C Psychology Building, East Lansing, Michigan 48824-1116; e-mail: rmc@msu.edu.

community psychology in his articulation of this alternative evaluation method (e.g., Rappaport, 1987; Zimmerman, 2000; Zimmerman, Israel, Schultz, & Checkoway, 1992; Zimmerman & Rappaport, 1988). Similarly, community psychologists have maintained that we should be giving psychology away to the public (Chavis, Stucky, & Wandersman, 1983; Kelly, 1970, 1971, 1979; Rappaport, 1977; Wandersman, 2003). Fetterman (1996, 2001a,b, 2002) argued that the techniques of evaluation must be shared with organizations to build their evaluation capacity. Given these shared philosophies and missions, empowerment evaluation may be a useful approach for community psychologists in their evaluation projects.

The empowerment evaluation literature is a developing one as researchers seek to apply these broad principles to their work with local communities. The purpose of this paper is to add to this literature by describing a multi-year evaluation project that was conducted with all state-funded rape prevention programs and rape victim services programs in Michigan. Based upon the theoretical framework of empowerment evaluation, the primary goal of the Sexual Assault and Rape Prevention (SARP) Evaluation Project was to build the programs' capacities for evaluation so they could design, implement, and sustain evaluations that captured their local initiatives. Specifically, this paper will address three topics. First, a brief review of the empowerment evaluation literature is presented with a focus on summarizing key methodological and definitional debates this method has sparked within evaluation science scholarship. Second, we will describe our application of empowerment evaluation to the work of violence against women organizations, highlighting the specific goals and activities of the SARP project. Finally, we will delve into an unresolved challenge in this literature, namely how to measure the "success" or outcomes of empowerment evaluation. To address this issue, we present a multi-method assessment of the effectiveness of our project.

## AN OVERVIEW OF EMPOWERMENT EVALUATION

Mark (2001) noted that "one of the larger recent trends in evaluation theory and practice is an increased focus on stakeholder participation" (p. 462). In participatory methods, the evaluation is organized as a team project with evaluation consultants and representatives from multiple stakeholder groups (e.g.,

agency staff, funders, clients) (Cousins & Earl, 1992, 1995; Cousins & Whitmore, 1998; Fetterman, 1994, 1995, 1996; Fetterman, Kaftarian, & Wandersman, 1996; Garaway, 1995; Greene, 1988; Mark & Shotland, 1985). The participating stakeholders are directly involved in planning, conducting, and analyzing the evaluation (Torres & Preskill, 2001). Because program staff take part in formulating questions and collecting data, some evaluation scientists have questioned the reliability and validity of participatory approaches (see Mark, 2001 for a review of these debates). Proponents of participatory methods note that these evaluations are done in close consultation with trained evaluators who can attune to these quality concerns (Cousins & Earl, 1992; Fetterman, 1995, 1997, 2001a). Moreover, a study by Fine, Thayer, and Coghlan (2000) suggested that stakeholder participation can improve the quality of an evaluation. In a national survey of 178 non-profit organizations, Fine et al. (2000) found that increased stakeholder involvement was "credited with improving evaluation design (by helping to guarantee that relevant questions are asked and appropriate measures are selected), ensuring available resources to implement the evaluation and its recommendations, increasing stakeholders' understanding and appreciation of an agency, and improving stakeholders' understanding of evaluation" (p. iii). Similarly, Patton (1997a) noted that increased participation is related to increased utilization: The more stakeholders participate, the more likely they are to use the evaluation's findings to improve programs.

One specific model of participatory evaluation is empowerment evaluation. In empowerment evaluation, "evaluators teach people to conduct their own evaluations and thus become more self-sufficient. This approach desensitizes and demystifies evaluation and ideally helps organizations internalize evaluation principles and practices, making evaluation an integral part of program planning" (Fetterman, 1996, p. 9). In other words, evaluators do not function as external evaluators (i.e., conduct an evaluation independently, then report the findings); they teach participants how to develop and conduct their own evaluations. Ownership and control of the evaluation is held by the programs. Fetterman (2001b) outlined five "facets," or defining characteristics, of empowerment evaluation: 1) training—evaluators teach program staff how to conduct their own evaluations; 2) facilitation—evaluators serve as coaches throughout the evaluation process; 3) advocacy—program staff may use their self-evaluations as tools for advocacy and institutional

change; 4) illumination—program staff may develop new insights into roles, structures, and program dynamics through conducting their evaluation; and 5) liberation—over time, organizations may free themselves from pre-existing roles and constraints based on what they learn in their evaluations.

Though empowerment evaluation can provide program staff with useful resources, this method also has its critics. For example, Scriven (1997) argued that empowerment evaluation is not evaluation at all because the evaluator is no longer an objective judge, but rather a coach and advocate. Similarly, Lackey, Moberg, and Balistreri (1997) noted that when this much control is given to program staff (or other stakeholders), it is often not clear how evaluation standards are defined and maintained (see also Stufflebeam, 1994). Even those amenable to the idea of evaluators as trainers and facilitators have critiqued this approach. Patton (1997b) argued that empowerment evaluation lacks conceptual and methodological clarity. Liberation, social change, and empowerment of individuals and institutions are key distinguishing features of empowerment evaluation. Yet, these outcomes are difficult to operationalize and assess, and even if measurable, they may be unrealistic for some smaller-scale or shorter-term projects.

How to measure the “success” of an empowerment evaluation remains a challenge in this literature. Although liberation may be the ultimate goal of this evaluation approach, most practitioners of this method have focused on whether the participating organizations can learn evaluation skills and successfully develop, launch, and utilize evaluations of their programs (see Levin, 1996). Moreover, Wandersman et al. (2004) noted that because building capacity is a fundamental goal of empowerment evaluation, practitioners should assess whether evaluation continues to be a part of an organization’s activities after the evaluation team has ended their formal work with the program. In addition, Wandersman (1999) argued for 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” (p. 96). Taken together, these emerging recommendations suggest that an empowerment evaluation should teach evaluation skills and build capacity. These outcomes should be sustained

after the work of the evaluators has concluded, and program staff should be using their evaluation results in an on-going way to improve program services.

### **A MODEL OF EMPOWERMENT EVALUATION FOR SEXUAL ASSAULT PROGRAMS**

Participatory methods in general, and empowerment evaluation in particular, may be effective approaches for evaluating the work of violence against women organizations. Rape crisis centers and domestic violence shelters often espouse the need for egalitarian models that share power, control, and resources (Koss & Harvey, 1991; Matthews, 1994; Riger, 1999). Participatory models of evaluation, such as empowerment evaluation, are consistent with many of these fundamental tenets of the feminist violence against women social movement (Andrews, 1996; Levin, 1999). Furthermore, it is especially important that program staff actively participate in their evaluations because they possess unique knowledge critical to the success of the evaluation, most notably how to respect and guard the safety and confidentiality of their clients (Sullivan & Cain, 2004). With its focus on teaching program staff how to conduct their own evaluations, empowerment approaches may be useful to rape crisis center and domestic violence organizations, which are being required with increasing frequency to evaluate their programs (see Riger et al., 2002). Without the financial resources to hire external evaluators, violence against women organizations may benefit tremendously from having evaluation expertise “in house.” Though empowerment evaluation requires time and effort (see Levin, 1999 for a discussion on how this is no small concern), program staff can learn skills to conduct evaluations, protect their clients, respond to accountability demands from their funders and communities, and gather useful information for program improvement.

To explore how empowerment evaluation could help sexual assault programs, a multiple stakeholder group in Michigan was formed to provide training and consultation to all state-funded rape prevention and rape victim services programs ( $N = 10$  rape prevention programs:  $n = 4$  rural programs,  $n = 2$  small urban,  $n = 3$  medium urban,  $n = 1$  large urban) ( $N = 24$  victim services programs:  $n = 10$  rural programs,  $n = 6$  small urban,  $n = 8$  medium urban). State funders wanted agencies to evaluate their prevention and victim services programs, but were aware

that such activities can be time consuming and costly. The funders wanted to build the evaluation capacity within each agency so that program staff could conduct their own evaluations. Recognizing the diversity across programs, funders also decided against standardized evaluation protocols, favoring local decision making that addressed local program initiatives. To address these aims, the Sexual Assault and Rape Prevention (SARP) Evaluation Project was created, which brought together university researchers, public health evaluators, state government funders, and rape crisis center staff to develop local evaluations of diverse programs throughout Michigan.

Over the six years of the project, the SARP team engaged in three primary activities. First, to learn more about the rape prevention and victim services programs throughout Michigan, SARP team members visited every program to meet staff and assess their evaluation needs. Second, based on the information gathered in these initial visits, the SARP team developed a series of training manuals for program staff to teach them how to create and execute their own evaluations. Finally, the SARP staff conducted numerous training workshops and informal technical assistance meetings to teach program staff how to evaluate their programs. The SARP team also provided on-going individualized consultation to program staff as they carried out what they had learned in training. Each of these activities is described in detail below.

### **Assessment of Needs**

Before evaluation training and consultation began, the SARP team traveled to all state-funded rape prevention and victim services agencies in Michigan to meet on-site with agency staff to learn about their programs. There were two primary goals for these visits. First, the SARP team members needed to explain the purpose of this project. The state funders were requiring all rape prevention grantees to evaluate their programs, and were recommending that all victim services programs also conduct evaluations.<sup>6</sup> We told program staff that the funders had formed and funded the SARP team to help them with their evaluations by providing on-going training and consultation. These services would be provided to all state-funded programs free of charge. We explained the empowerment evaluation approach and emphasized that the control of the evaluations would remain within each

local agency. Simply put, we were there to help them create and conduct their own evaluations—we were not there to evaluate them. In many instances, programs were already engaged in some evaluation activities because multiple funding agencies require evaluation. Thus, an invitation to be part of a project that would help each agency carry out their evaluations was positively received. However, in some instances there was confusion in these initial meetings as to the role and approach of the SARP team. Many program staff assumed we had been hired as external evaluators and were there to evaluate their programs. It required multiple conversations over the first year of the project to assure program staff that we were not external evaluators, but evaluation consultants.

The second goal of these visits was to conduct in-depth interviews with staff to learn about their programs and organizations. Two SARP team members conducted interviews with the each agency's executive director, the sexual assault program coordinator(s), and other members of the agency (e.g., grant coordinators, volunteer coordinators). The interview covered eight topics: 1) organizational history and background; 2) current structure, function, and philosophy of the organization; 3) description of their prevention and/victim services program; 4) information about the kinds of clients served by their programs; 5) information about funding sources and amounts; 6) community relations; 7) existing efforts to evaluate their programs; and 8) their needs from the SARP team. The interviews took approximately three hours to complete, and were tape-recorded with permission from the program staff. The SARP team transcribed these interviews and created summaries for each site that highlighted key programmatic issues and evaluation needs. The organizational histories provided the SARP team with contextual details about the programs and their communities that identified local strengths and concerns. These profiles were updated throughout the six years of the project as circumstances changed within each agency.

### **Development of Evaluation Training Materials**

Informed by these site visit interviews, the SARP team, with input from program staff and funding agency representatives, developed an evaluation training curriculum. Our goal was to create training materials that would: 1) be user-friendly, avoid excessive use of jargon, and not presuppose knowledge of technical as well as non-technical aspects of evaluation; 2) be flexible and allow agencies to use the

<sup>6</sup>The prevention programs also received \$5,000 from their state funders to support the mandated evaluation.

evaluation framework for multiple programs and purposes; 3) be “stand alone” documents that could guide staff through the evaluation steps even if they were not able to attend training; and 4) provide real-life examples from sexual assault prevention and victim services programs. The process of planning and conducting an evaluation was broken down into eight discrete phases, beginning with how to develop and refine program goals and objectives through how to utilize evaluation findings to improve programs (see Table I for a description of each phase). Throughout all eight phases, we emphasized the importance of conducting both process evaluation (i.e., assessing how their program was being implemented and how it was perceived by their clients) and outcome evaluation (i.e., assessing the impact of their program on their clients). We recommended that programs begin with process evaluation and add outcome evaluation as resources permitted.

To teach program staff these eight phases of evaluation, four training manuals were created: introduction manuals and data analysis/utilization manuals for rape prevention programs and victim services programs (SARP Project, 1998a, 1998; 1999a, 1999b). In addition, we created separate resource guides for rape prevention and victim services programs that gathered additional information about evalua-

tion (e.g., methods for qualitative evaluation, inexpensive software for data analysis, web-based evaluation resources) (SARP Project, 2000a, 2000b). The training materials were reviewed by outside experts in the field, including sexual assault prevention and victim services program staff throughout Michigan, and their suggestions for revision were addressed before the materials were used in workshops. The training materials did not tell programs what to evaluate; it emphasized *how* to evaluate. Similarly, whether program staff decided to use their own evaluation measures, samples provided by the SARP team, or instruments created by other researchers or practitioners was not our dictate. Our aim was to support staff through the entire decision making process of conducting an evaluation.

### Evaluation Training and Technical Assistance

To provide training on the eight phases of evaluation, the SARP team offered a series of four training workshops, staged over five years. The first training covered Phases 1–4 (Clarify Program Goals & Objectives through Develop a Data Collection Plan); the second Phases 5–6 (Manage the Data & Analyze the Data); the third Phase 7 (Report Evaluation

**Table I.** The Phases of Evaluation Taught in the SARP Evaluation Model

Phase	Description
<i>Phase 1:</i> Clarify program goals & objectives	Identify major goals of the prevention/victim services program Identify major goals for evaluation of program Identify process goals and outcome goals for the evaluation Identify objectives for all process and outcome goals
<i>Phase 2:</i> Select an evaluation design	Identify appropriate evaluation design for process goals Identify appropriate evaluation design for outcome goals
<i>Phase 3:</i> Select & modify measurement instruments	Review instruments/measures already in use by program Review other instruments/measures (provided by SARP team) Make modification to existing instruments/create new instruments
<i>Phase 4:</i> Develop a data collection plan	Identify key personnel in program for conducting the evaluation Develop an implementation plan & tracking log for the evaluation
<i>Phase 5:</i> Manage the data	Software training (Excel) Create codebook and coding procedures Data entry and data cleaning procedure
<i>Phase 6:</i> Analyze & interpret the data	Descriptive and inferential statistics Data interpretation and presentation of results
<i>Phase 7:</i> Report evaluation findings	Identify positive/negative and intended/unintended evaluation findings Develop summaries of program, evaluation, and findings Writing reports for funders Writing reports for other groups
<i>Phase 8:</i> Utilize evaluation findings	Identify areas for improvement in the evaluation Identify areas for improvement in the program Disseminating findings

Findings); and the fourth Phase 8 (Utilize Evaluation Findings). Each workshop was offered separately for prevention programs and for victim services programs and was conducted regionally throughout Michigan. Additional review workshops were held to train new staff or provide refresher instruction for existing staff. Over the course of the project, the SARP team conducted 18 evaluation trainings. The training workshops were practice-oriented and allowed participants to apply the knowledge and skills to their program evaluation activities. Program staff were divided into small groups and worked with a member of the SARP team to learn each phase and apply it to their work.

Offering these evaluation trainings was important, but not sufficient to create progress in launching and sustaining program evaluations. Our aim was to keep staff engaged in the evaluation process so that what they learned in training was not lost. Therefore, a couple of months after each of these 18 training workshops, the SARP team held regional technical assistance meetings. For these informal meetings, program staff were encouraged to bring their work in progress, and as often as possible, agencies were paired one-on-one with SARP team members to receive individualized assistance with their specific local needs.

### **THE EFFECTIVENESS OF SARP EMPOWERMENT EVALUATION APPROACH**

Because there are no published reports on using empowerment evaluation with sexual assault programs, we wanted to examine whether this evaluation approach was effective with the participating programs. Consistent with emerging recommendations for measuring the "success" of empowerment evaluation projects (see Levin, 1996; Wandersman et al., 2004; Wandersman, 1999), we used a multi-method assessment that included: 1) ratings of participants' satisfaction with the SARP training and consultation; 2) objective and subjective reports of whether the programs had launched their own evaluations; and 3) follow-up interviews assessing whether the programs were: a) still conducting evaluations after the work of the SARP team formally ended; and b) if so, were they using the results obtained in their evaluations to improve program services.

#### **Programs' Satisfaction with the SARP Model**

We routinely assessed participants' satisfaction with the assistance they received from the SARP

team. On a four-point scale (1 = "not at all satisfied;" 4 = "very satisfied"), we asked every participant in every evaluation training or technical assistance meeting to rate three key dimensions of the SARP model: 1) How useful do you think this workshop will be to you in conducting your evaluation? (prevention sites:  $M = 3.84, SD = .37$ ; service sites:  $M = 3.72, SD = .53$ ); 2) How useful do you think the training manuals/training materials will be to you in conducting your evaluation? (prevention sites:  $M = 3.87, SD = .36$ ; services sites:  $M = 3.79, SD = .45$ ); and 3) What is your overall satisfaction with SARP training format (tailored instruction and small group activities and consultation)? (prevention sites:  $M = 3.75, SD = .43$ ; service sites:  $M = 3.82, SD = .45$ ). Staff ratings were consistently very high: between 98 and 100% of the participants in the workshops were very satisfied or satisfied with all three dimensions of the SARP model.

#### **Tracking Programs' Progress in Developing Evaluation Skills**

These satisfaction data speak to how the SARP project was perceived by program staff, but it is also important to ascertain whether this approach was successful in helping agencies conduct their own evaluations. To address this issue, we obtained the evaluation reports that the 10 rape prevention programs filed with their funders for three consecutive years: 1997 (the first complete year SARP began working with these agencies) through 1999 (the year we completed training on the first seven phases).<sup>7</sup> These reports contained complete information about the evaluation activities of the rape prevention programs, and a content analysis was performed to determine if and how the programs had made progress in developing their evaluation skills. However, the funder of the victim services agencies did not require programs to file evaluation reports, so other data sources were used to track the progress of the victim service programs.

The content analysis of rape prevention programs' evaluation reports was conducted by three members of the SARP team and one independent research assistant who had no prior experience working with the SARP team or with these 10 prevention

<sup>7</sup>It was not possible to offer Phase 8 (Utilize Evaluation Findings) training until the later years of the project (2000 and 2001). However, we covered this materially informally during the 1999 technical assistance workshops.

**Table II.** Coding the Rape Prevention Programs' Funders Reports for Evidence of Effectiveness of the SARP Evaluation Model

SARP team's goals	Communities							
	Rural (n = 4)		Small urban (n = 2)		Medium urban (n = 3)		Large urban (n = 1)	
	1997	1999	1997	1999	1997	1999	1997	1999
<i>Phases 1 &amp; 2: Clarify goals &amp; objectives; select an evaluation design</i>								
Developed an evaluation plan	1	4	2	2	0	3	0	0
Conducted both process and outcome evaluation	2	4	2	2	1	3	1	1
Change in sophistication of evaluation	0	4	0	2	0	3	0	0
<i>Phase 3: Select &amp; modify instruments</i>								
Made changes in instruments	0	4	0	2	0	3	0	0
<i>Phase 4: Develop a data collection plan</i>								
Developed an evaluation implementation log	1	4	0	2	0	3	0	0
<i>Phase 5: Manage the data</i>								
Created an internal data management system	1	3	0	2	0	1 <sup>a</sup>	0	0 <sup>b</sup>
<i>Phase 6: Analyze &amp; interpret the data</i>								
Conducted own data analysis	1	3	0	2	0	1 <sup>a</sup>	0	0 <sup>b</sup>
Change in sophistication of analysis	0	3	0	2	0	1 <sup>a</sup>	0	0 <sup>b</sup>
<i>Phase 7: Report evaluation findings</i>								
Reported findings to funder	0	3	0	2	0	3	0	1
Change in sophistication of reporting	0	3	0	2	0	1	0	0
<i>Phase 8: Utilize findings</i>								
Changed curriculum based on findings	0	2	0	2	0	2	0	0
Used findings to access new program settings	0	2	0	2	0	2	0	1
Used findings to secure new funding sources	0	1	0	0	0	1	0	0

<sup>a</sup>Two medium-sized urban sites hired an external consultant for data analysis.

<sup>b</sup>The one large urban site hired an external consultant for the entire evaluation.

programs. Each agency's reports from 1997 through 1999 were reviewed for evidence that they had been utilizing the information provided by the SARP team. These reports were coded on 13 specific goals related to the eight phases of evaluation (see Table II for a complete list of the criteria). A codebook was created that defined each of these 13 criteria, and a binary coding system was selected (yes, there was evidence of that criteria in the report; no, there was no evidence of that criteria in the report). Two people coded all reports (one SARP team member and the independent non-SARP coder), and inter-rater agreement was quite high (94%) ( $\kappa = .87$ ).

The key question was how many agencies exhibited competency on each of these 13 criteria at the beginning of the project (in 1997) and how many did so after two years of working with the SARP project (in 1999). If the empowerment evaluation approach is effective, then these agencies should be developing their evaluation capacity over time. As can be seen in Table II, the prevention programs made substantial progress in their evaluation initiatives over three years. With respect to Phases 1 and 2 (Clarify Program Goals & Objectives and Select an Evaluation Design), some programs (3 of 10) had evalua-

tion plans in place after working with SARP for only one year, but by 1999, 9 of the 10 prevention programs had specific evaluation plans. A McNemar test, a non-parametric test for related samples, was significant ( $p = .03$ ), indicating that the programs had made significant change on this criteria from 1997 to 1999 (Pett, 1997).<sup>8</sup> In 1997, six of 10 programs were conducting both process and outcome evaluations, and by 1999 all agencies were using combined methods. This increase was not statistically significant (McNemar  $p = .06$ ). By 1999, 9 of 10 programs had made substantial improvements in the sophistication of their evaluation designs, such as adding a comparison group to the design and/or supplementing quantitative findings with systematic qualitative data collection (e.g., focus groups) (McNemar  $p = .002$ ). In Phase 3 (Select & Modify Measurement Instruments), our goal was that program staff would be reviewing their evaluation tools on a regular basis and would be making changes to those tools based on what they were

<sup>8</sup>The McNemar tests compared data from 1997 to 1999 (not across all three time points). These tests were based on a binomial distribution because of the small sample size in these analyses (see Pett, 1997).

learning in their evaluations. By 1999, 9 of the 10 programs were making such changes in their evaluation questionnaires (McNemar  $p = .002$ ). For Phase 4 (Develop a Data Collection Plan), our aim was to help programs develop evaluation implementation logs to track their evaluation methods. In 1997, one program had an implementation log, but by 1999, 9 of 10 had such internal tracking mechanisms (McNemar  $p = .004$ ).

In Phase 5 and 6 (Manage the Data and Analyze the Data), our goal was that program staff develop a data management system and conduct their own data analysis. In 1997, one of the agencies had a data management system, but six had developed one by 1999 (McNemar  $p = .03$ ). By 1999, six programs had conducted their own data analyses and all demonstrated increased sophistication in their data analyses (e.g., one year reporting descriptive data and following years reporting tests of significance with graphs of change) (McNemar  $p = .03$  for both criteria). For these phases in the evaluation process, two of the three medium-sized urban sites decided to hire an external consultant to do their data analysis. For consistency, we coded the progress of these two agencies on these phases even though they worked with an external evaluator. The two sites that hired out their data analysis did not demonstrate progress in these phases in their reports from 1997 to 1999, most likely because the analyses were not yet ready from the external evaluator. The one medium-sized urban program that did work with the SARP team for these phases did complete analyses by 1999.

For Phase 7 (Report Evaluation Findings), we examined whether the programs were reporting substantive evaluation findings to their funder and also assessed how the complexity of their reports changed over time. In 1997 none of the agencies had such a report, but 9 of 10 programs did so by 1999 (McNemar  $p = .002$ ). In addition, 6 of the 10 program made substantial gains in the complexity of their reports (e.g., summarizing large amounts of quantitative data in tables, presenting tests of statistical significance, discussing the limitations of the evaluation and outlining revisions for subsequent years, and linking evaluation findings to needed programmatic changes) (McNemar  $p = .02$ ). Finally, although we offered only informal consultation in Phase 8 (Utilize Evaluation Findings) by 1999, it appears that some agencies were making progress in using their findings to improve and/or promote their programs. By 1999, 6 programs had made changes in their intervention curriculum based on their evaluation findings (McNemar

$p = .02$ ). Similarly, seven of the 10 programs had been able to use their evaluation findings to secure access to new settings for their programs (McNemar  $p = .008$ ). Two programs had been able to secure new funding because of their evaluation findings (McNemar  $p = .25$ ).

In the absence of funders' evaluation reports for the victim services programs, we turned to other data sources that could shed light on whether these programs had successfully launched their own evaluations. Throughout the SARP project, team members kept field notes on their contacts with each program, and a content analysis was performed on these data. This coding was performed by two independent research assistants, neither of whom were SARP team members. A simplified coding scheme was needed because these data were not as complete as the prevention sites' evaluation reports to their funders. Three criteria were assessed using a binary coding system: 1) was there evidence that the program had developed an evaluation plan (Phases 1 and 2); 2) was there evidence that an evaluation had been successfully launched and the program was collecting data (Phases 3 and 4); and 3) was there evidence that the program was analyzing data (Phases 5 and 6). For example, if a SARP team member had indicated in her field notes that she had worked with a particular victim services program on developing an evaluation plan, or had answered questions about implementing an evaluation, or had helped staff with data analysis, then these three criteria were coded affirmatively. Inter-rater agreement for the field notes coding was quite high (96%) ( $\kappa = .92$ ). The results revealed that by 1999, 92% of the victim services programs had developed evaluation plans, 75% had launched evaluations (most [83%] were process evaluations only), and 21% were conducting data analysis.

### Tracking Programs' Longer-Term Evaluation Capacity

As a final measure of the effectiveness of this empowerment evaluation approach, follow-up phone interviews were conducted with all 10 prevention programs and 24 victim services programs. These interviews were conducted in late 2002/early 2003, approximately one year after the work of the SARP project had ended. The principal investigator contacted each agency and asked program staff if they were conducting evaluations of their programs, and if so, to describe their current efforts, what resources they

were using to conduct these evaluations, and whether they had made programmatic changes based on their evaluation findings.<sup>9</sup> Of the 10 prevention programs, 8 were still receiving state funds and were still conducting evaluations one year after the SARP team had ended their work (the other two programs had been defunded and folded). All eight programs had completed two full evaluation cycles, and had made changes to their curricula and/or their evaluation protocols based on their evaluation findings. When asked what resources they were using to help them with their evaluations, all eight spontaneously mentioned the SARP training manuals and notes left for them by previous staff members. One interviewee said, "We have these manuals, I think they're called the SARP manuals. Anyway, I never went to any of their trainings. My executive director told me to read them [the manuals] and follow along. They've been really helpful." Of the 24 victim services programs, 20 were still receiving state funding (four had been defunded and were still operational, though they were not conducting evaluations). Of these 20 programs, 18 (90%) were conducting evaluations (15 were process only; 3 were process and outcome). All programs had made changes to their policies and procedures for helping sexual assault victims based on the results of their evaluation findings. All interviewees spontaneously mentioned the SARP project (manuals and training) as a primary resource for their evaluation needs. For example, one staff member said, "I go back to SARP manuals over and over again. They spell it all out." The participants also mentioned received helpful information from other training manuals (e.g., Sullivan & Coats, 2000) and resources from their state coalition of domestic violence and sexual assault agencies.

## CONCLUSIONS

The empowerment evaluation approach used by the SARP team appeared to be effective in helping vi-

olence against women agencies throughout Michigan conduct methodologically sound and programmatically useful evaluations. Throughout the six years of the project, 90% of the rape prevention programs made substantial progress in planning and conducting evaluations. Approximately one year after the SARP team had ended their work with the prevention sites, all funded programs were still conducting evaluations. However, the prevention programs were required by their funders to do evaluations and they received financial support to do so. That so many were successful in initiating and sustaining evaluations may not reflect the SARP teams' efforts, but may instead represent compliance with a state mandate. However, other evaluation scholars have noted that just because evaluation is required does not necessarily mean that it is actually conducted or conducted well (Fetterman et al., 1996; Love, 1991; Patton, 1982; Rossi et al., 1999). The empowerment evaluation approach used in this project may have been instrumental in transforming a potentially onerous mandate into a useful activity.

The rape victim services programs in Michigan were not required to evaluate their programs, so their evaluation progress provides a different perspective on the effectiveness of the SARP approach. Because there was no mandate for evaluation, these programs did not file evaluation reports with their funders, making it impossible to collect objective outcome data. However, analysis of the projects' field notes indicated that 75% of the victim services programs had also successfully launched program evaluations. At the one-year follow-up interview, 90% of the funded programs were still conducting evaluations even though they were not required to do so by their primary state funder. It should be noted, however, that some of the victim services agencies initially were not interested in developing their own evaluations and stated they would prefer to use standardized evaluation protocols and tools. They expressed that they did not have the time to devote to such a labor-intensive effort and wanted external evaluators to create materials for them. In these instances, the SARP team emphasized that we would help all agencies work as efficiently as possible and would provide examples, templates, and other resources. We recommended already developed tools that reflected an awareness of the dynamics of sexual assault (e.g., Sullivan & Coats, 2000) and continued to work with these agencies in providing on-going assistance with their implementation needs. For some programs, being supportive of their local needs meant guiding them

<sup>9</sup>It is possible that the participants' answers may reflect social desirability biases as they were talking with the principal investigator of the project. However, due to high staff turnover at these agencies, only eight staff members (of 34) had worked with the SARP project. The remaining 26 staff were new to their organizations (i.e., they had started at their agencies near the end of the SARP project and had not worked with SARP staff, or they had started after the SARP project had ended) and they did not know the principal investigator. There were no systematic differences in the participants' answers as a function of whether they knew the principal investigator. However, it should be noted that these data are still subject to self-reporting biases.

to established resources; for others, it was helping them develop their own designs and tools. In either context, providing regular consultation and assistance was paramount, as it would be extremely difficult for any one evaluation manual to provide guidance on the day-to-day activities of conducting an evaluation.

Empowerment evaluation emphasizes the process of conducting evaluation—training staff and providing consultation to help them conduct their own evaluations. However, as Mark (2001) noted, “by placing more emphasis on stakeholder *processes*, the empowerment and transformative literature may place less emphasis on the evaluation *findings*” (p. 463) (emphases in original) (see also Sechrest, 1997). In other words, an empowerment approach may help programs conduct their evaluations, but what were the results of those evaluations? To date, the emphasis in the empowerment evaluation literature has indeed been on the process of working collaboratively with program staff, but in practice, we found ourselves very involved in the results of their evaluations. Program staff were keenly focused on their findings—what did they mean for their programs, how could they improve services, what would their funders think? These agencies’ evaluation results were remarkably consistent with findings reported in the academic literature on rape prevention (see Lonsway, 1996). Overall, the prevention programs showed positive short-term effects—significant increases in participants’ awareness of sexual violence and modest, yet significant, short-term attitude change. The victim services programs found that their advocacy services were consistently rated as very helpful to their clients, and their counseling programs had significant positive effects in facilitating survivors’ recoveries. These findings are also consistent with other published findings on victim services programs (Howard, Riger, Campbell, & Wasco, 2003; Riger et al., 2002; Wasco et al., 2004).

Wandersman (1999) took this issue further, arguing that evaluation results should be used to improve services. In this project, all program staff interviewed at the one-year follow-up stated that they had made programmatic improvements based on their evaluation findings. For the prevention sites, the most commonly reported programmatic changes included: adding new content to their curricula to address new topics; adding more material on existing topics that were not receiving adequate coverage, and using their evaluation findings to advocate for more time in their school districts to conduct their prevention programs. For the victim services sites, common programmatic

changes included: outreach to communities of Color to increase services to ethnic minority women, advocacy with the criminal justice system for improved responses to victims, and advocating for changes within their organization to increase the number of counseling sessions that could be provided to rape survivors. Program staff consistently mentioned that they had “made peace” with evaluation requirements because they were learning useful information about their programs, which was improving their services.

Although this project appears to have been successful in teaching evaluation skills, as well as building and sustaining these organization’s evaluation capacity, there are several limitations of this project specifically and of the empowerment evaluation approach more generally that merit examination. First, it is essential to note that we cannot definitely attribute the evaluation successes of these programs to the work of the SARP team. Though the available data suggest that the SARP model was certainly helpful and probably instrumental in the development of these organizations, we did not assess how other resources available to these agencies (e.g., staff expertise, other evaluation materials) may have affected their progress in evaluation. Second, the data on the effectiveness of the SARP model are largely self-report—the program staff said that the SARP project was helpful. Coding the funders’ reports for the rape prevention programs provided a more objective view into the success of these programs in developing evaluation skills, but recall that such data were not available for the victim services programs. Although it is not common in the empowerment evaluation literature to examine the effectiveness of such projects (and indeed, this should become more common), it is a strength of this study that such an attempt was made, but future projects must strive for greater methodological rigor when evaluating empowerment approaches.

In this project, we struggled with two issues that are fundamental to an empowerment evaluation approach. The focus of our work was teaching program staff how to evaluate their work, and in doing so, the hope was that this process would be empowering to the staff, and ultimately to their organizations. Such a focus on the staff begs the question, what about the programs’ clients? Were the clients empowered? It is not necessarily a limitation of empowerment evaluation inasmuch as our use of it in this project that leaves this question hanging. We focused on working with the staff and encouraged them to check in with their clients regarding the appropriateness and feasibility

of different evaluation options, which they did. In this project, program staff indicated that this indirect link between the evaluators and the clients was appropriate in order to protect the safety, privacy, and confidentiality of their clients. However, this example reminds us to consider carefully who should be brought to the table in an empowerment evaluation project because this influences who may be "empowered" by this process. Another struggle for all involved in this project was the time, effort, energy, and resources it required. By far, the time commitment was the most common complaint the program staff had about the SARP project—it took a long time to do this work, which is to be expected in capacity building, but in the day-to-day operations of an organization, an empowerment evaluation approach may sometimes be too labor intensive (Levin, 1999). The effort that goes into conducting evaluation must be balanced against the time staff need to work with their clients. Evaluation can enhance program services, but care must be taken so that it does not detract from agencies' primary mission. As the demand for evaluation continues to grow within the violence against women movement, future work must continue to identify collaborative evaluation practices that respect the needs of victimized populations and the service providers with whom they work.

## ACKNOWLEDGMENTS

The authors thank Courtney Ahrens, Marcia Clark, Gloria Aponte Clarke, Dee Hurlbert, Stacy Meade, Joan Moore, Molly Smeltzer, Stephanie Townsend, and Sharon Wasco for their assistance throughout this project; and Robin Miller, Cris Sullivan, and Eric Walters for their helpful comments on previous drafts of this manuscript. This publication was supported by the Preventive Health and Health Services (PHHS) Block Grant from the Centers for Disease Control and Prevention (CDC). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the CDC

## REFERENCES

- Andrews, A. B. (1996). Realizing participant empowerment in the evaluation of nonprofit women's services organizations: Notes from the front line. In D. M. Fetterman, S. J. Kaftarian, & A. Wandersman (Eds.), *Empowerment evaluation: Knowledge and tools for self assessment and accountability* (pp. 141–160). Thousand Oaks, CA: Sage.
- Campbell, R., & Wasco, S. M. (2000). Feminist approaches to social science: Epistemological and methodological tenets. *American Journal of Community Psychology, 28*, 773–792.
- Chavis, D. M., Stucky, P., & Wandersman, A. (1983). Returning basic research to the community: A relationship between scientist and citizen. *American Psychologist, 38*, 424–434.
- Cousins, J. B., & Earl, L. M. (Eds.) (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning*. London: Falmer.
- Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14*, 397–418.
- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation, 80*, 5–23.
- Fawcett, S. B. (1991). Some values guiding community research and action. *Journal of Applied Behavioral Analysis, 24*, 621–636.
- Fetterman, D. M. (2002). Empowerment evaluation: Building communities of practice and a culture of learning. *American Journal of Community Psychology, 30*, 89–102.
- Fetterman, D. M. (2001a). *Foundations of empowerment evaluation*. Thousand Oaks, CA: Sage.
- Fetterman, D. M. (2001b). Empowerment evaluation and self-determination: A practical approach toward program improvement and capacity building. In N. Schneiderman, M. A. Speers, J. M. Silva, H. Tomes, & J. H. Gentry (Eds.), *Integrating behavioral and social sciences with public health* (pp. 321–350). Washington DC: American Psychological Association.
- Fetterman, D. M. (1997). Empowerment evaluation: A response to Patton and Scriven. *Evaluation Practice, 18*, 253–267.
- Fetterman, D. M. (1996). Empowerment evaluation: An introduction to theory and practice. In D. M. Fetterman, S. J. Kaftarian, & A. Wandersman (Eds.), *Empowerment evaluation: Knowledge and tools for self assessment and accountability* (pp. 3–46). Thousand Oaks, CA: Sage.
- Fetterman, D. M. (1995). In response to Stufflebean (1994). *Evaluation Practice, 16*, 179–199.
- Fetterman, D. M. (1994). Empowerment evaluation. *Evaluation Practice, 15*, 1–15.
- Fetterman, D. M., Kaftarian, S. J., & Wandersman, A. (1996). *Empowerment evaluation: Knowledge and tools for self assessment and accountability*. Thousand Oaks, CA: Sage.
- Fine, A. H., Thayer, C. E., & Coghlan, A. (2000). *Program evaluation practice in the nonprofit sector*. Washington, DC: The Aspen Institute.
- Garaway, G. B. (1995). Participatory evaluation. *Studies in Educational Evaluation, 21*, 85–102.
- Greene, J. G. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review, 12*, 91–116.
- Howard, A., Riger, S., Campbell, R., & Wasco, S. M. (2003). Counseling services for battered women: A comparison of outcomes for physical and sexual abuse survivors. *Journal of Interpersonal Violence, 18*, 717–734.
- Kelly, J. G. (1979). Tain't what you do, it's the way that you do it. *American Journal of Community Psychology, 7*, 244–261.
- Kelly, J. G. (1970). Antidotes for arrogance: Training for community psychology. *American Psychologist, 25*, 524–531.
- Kelly, J. G. (1971). Qualities for the community psychologist. *American Psychologist, 26*, 897–903.
- Koss, M. P., & Harvey, M. R. (1991). *The rape victim: Clinical and community interventions*. Newbury Park, CA: Sage.
- Lackey, J. M., Moberg, D. P., & Balistreri, M. (1997). By whose standards? Reflections on empowerment evaluation and grassroots groups. *Evaluation Practice, 18*, 137–147.
- Levin, H. M. (1996). Empowerment evaluation and accelerated schools. In D. M. Fetterman, S. J. Kaftarian, & A. Wandersman (Eds.), *Empowerment evaluation: Knowledge and tools for self assessment and accountability* (pp. 49–64). Thousand Oaks, CA: Sage.
- Levin, R. (1999). Participatory evaluation: Researchers and service providers as collaborators versus adversaries. *Violence Against Women, 5*, 1213–1227.

- Lonsway, K. A. (1996). Preventing acquaintance rape through education: What do we know? *Psychology of Women Quarterly*, 20, 229–265.
- Love, A. J. (1991). *Internal evaluation: Building organizations from within*. Newbury Park, CA: Sage.
- Mark, M. M. (2001). Evaluation's future: Furor, futile or fertile? *American Journal of Evaluation*, 22, 457–479.
- Mark, M. M., & Shotland, R. L. (1985). Stakeholder-based evaluation and value judgments: The role of perceived power and legitimacy in the selection of stakeholder groups. *Evaluation Review*, 9, 605–626.
- Matthews, N. A. (1994). *Confronting rape: The feminist anti-rape movement and the state*. New York: Routledge.
- Patton, M. Q. (1997a). *Utilization-focus evaluation* (3rd ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (1997b). Toward distinguishing empowerment evaluation and placing it in a larger context. *Evaluation Practice*, 18, 147–164.
- Patton, M. Q. (1982). *Practical evaluation*. Newbury Park, CA: Sage.
- Pett, M. A. (1997). *Nonparametric statistics for health care research: Statistics for small samples and unusual distributions*. Thousand Oaks, CA: Sage.
- Rappaport, J. (1987). Terms of empowerment/exemplars of prevention: Toward a theory for community psychology. *American Journal of Community Psychology*, 15, 121–148.
- Rappaport, J. (1977). *Community psychology: Values, research, and action*. New York: Holt, Rinehart, & Winston.
- Riger, S. (1999). Working together: Challenges in collaborative research on violence against women. *Violence Against Women*, 5, 1099–1117.
- Riger, S., Bennett, L. W., Wasco, S. M., Frohmann, L., Schewe, P. A., Camacho, J., & Campbell, R. (2002). *Evaluating domestic violence and sexual assault services*. Thousand Oaks, CA: Sage.
- Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (1999). *Evaluation: A systematic approach*. Thousand Oaks, CA: Sage.
- Scriven, M. (1997). Empowerment evaluation examined. *Evaluation Practice*, 18, 165–175.
- Sehrest, L. E. (1997). Book review (Empowerment Evaluation). *Environment & Behavior*, 29, 422–426.
- Sexual Assault and Rape Prevention (SARP) Evaluation Project. (2000a). *Resources for evaluating sexual assault prevention programs*. Okemos, MI: Michigan Public Health Institute.
- Sexual Assault and Rape Prevention (SARP) Evaluation Project. (2000b). *Resources for evaluating sexual assault service delivery programs*. Okemos, MI: Michigan Public Health Institute.
- Sexual Assault and Rape Prevention (SARP) Evaluation Project. (1999a). *Evaluation training and practice for sexual assault prevention: Part two (data analysis)*. Okemos, MI: Michigan Public Health Institute.
- Sexual Assault and Rape Prevention (SARP) Evaluation Project. (1999b). *Evaluation training and practice for sexual assault service delivery: Part two (data analysis)*. Okemos, MI: Michigan Public Health Institute.
- Sexual Assault and Rape Prevention (SARP) Evaluation Project. (1998a). *Introduction to evaluation training and practice for sexual assault prevention*. Okemos, MI: Michigan Public Health Institute.
- Sexual Assault and Rape Prevention (SARP) Evaluation Project. (1998b). *Introduction to evaluation training and practice for sexual assault service delivery*. Okemos, MI: Michigan Public Health Institute.
- Stufflebeam, D. L. (1994). Empowerment evaluation, objectivist evaluation, and evaluation standards: Where the future of evaluation should not go and where it needs to go. *Evaluation Practice*, 15, 321–338.
- Sullivan, C. M., & Coats, S. (2000). *Outcome evaluation strategies for sexual assault service programs: A practical guide*. Okemos, MI: Michigan Coalition Against Domestic and Sexual Violence.
- Sullivan, C. M., & Cain, D. (2004). Ethical and safety considerations when obtaining information from or about battered women for research purposes. *Journal of Interpersonal Violence*, 19, 603–618.
- Torres, R. T., & Preskill, H. (2001). Evaluation and organizational learning: Past, present, and future. *American Journal of Evaluation*, 22, 387–395.
- Wandersman, A. (2003). Community science: Bridging the gap between science and practice with community-centered models. *American Journal of Community Psychology*, 31, 227–242.
- Wandersman, A. (1999). Framing the evaluation of health and human service programs in community settings: Assessing progress. *New Directions for Evaluation*, 83, 95–102.
- Wandersman, A., Keener, D. C., Snell-Johns, J., Miller, R. L., Flaspohler, P., Livet-Dye, M., Mendez, J., Behrens, T., Bolson, B., & Robinson, L. (2004). Empowerment evaluation: Principles and action. In L. A. Jason, C. B. Keys, Y. Suarez-Balcazar, R. R. Taylor, & M. I. Davis (Eds.), *Participatory community research: Theories and methods in action* (pp. 139–156). Washington DC: American Psychological Association.
- Wasco, S. M., Campbell, R., Howard, A., Mason, G., Staggs, S., Schewe, P., & Riger, S. (2004). A statewide evaluation of services provided to rape survivors. *Journal of Interpersonal Violence*, 19, 252–263.
- Zimmerman, M. A. (2000). Empowerment theory: Psychological, organizational, and community levels of analysis. In J. Rappaport & E. Seidman (Eds.), *Handbook of Community Psychology* (pp. 2–54). New York: Plenum.
- Zimmerman, M. A., Israel, B. A., Schultz, A., & Checkoway, B. (1992). Further explorations in empowerment theory: An empirical analysis of psychological empowerment. *American Journal of Community Psychology*, 20, 707–727.
- Zimmerman, M. A., & Rappaport, J. (1988). Citizen participation, perceived control, and psychological empowerment. *American Journal of Community Psychology*, 16, 725–750.