

EVALUATION GUIDEBOOK
FOR PROJECTS FUNDED
BY S.T.O.P. FORMULA GRANTS
UNDER THE VIOLENCE AGAINST WOMEN ACT

By:

Martha R. Burt
Adele V. Harrell
Lisa C. Newmark
Laudan Y. Aron
Lisa K. Jacobs
and others

December 1997

Urban Institute
2100 M Street, N.W.
Washington, DC 20037
(202) 833-7200
<http://www.urban.org>

This project was supported by Grant No. 95-WT-NX-0005 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice or of other staff members, officers, trustees, advisory groups, or funders of the Urban Institute.

TABLE OF CONTENTS

TABLE OF CONTENTS	i
WHERE TO GET RESOURCES DISCUSSED IN THIS <i>GUIDEBOOK</i>	vii
PREFACE	ix
CHAPTER 1: GETTING STARTED: THINKING ABOUT EVALUATION	1
What Should Be Evaluated?	2
Working Together: Participatory Evaluation	4
Respect for Women’s Privacy and Safety	5
How to Use this <i>Guidebook</i>	5
CHAPTER 2: DEVELOPING AND USING A LOGIC MODEL	7
Creating a Logic Model	8
EXHIBIT 2.1: Logic Model for Evaluation of Shelter-Based Services For Domestic Violence Victims	9
EXHIBIT 2.2: Logic Model for Your Program	10
Program Evaluability	12
If Not Impact Evaluation, What?	14
Evaluation Options for Programs that Are Already Under Way	14
CHAPTER 3: MAKING EVALUATION WORK FOR YOU	15
Finding an Evaluator	15
Getting What You Want from Evaluation	17
For State Coordinators	18
For Subgrantees	20
CHAPTER 4: USING EVALUATION INFORMATION	23
Uses for Program Promotion	23
Avoiding Misunderstandings of Your Program	23
Using Evaluation Information to Get Funding	25
Using Evaluation Information to Increase Public Appreciation of Your Program	26
Uses for Program Development	26
Uses for Community Development	29
Collecting Data over Time	30
CHAPTER 5: USING THE SAR AND THE SSS	31
What Are the SAR and the SSS?	31
Getting Information for the SAR and the SSS	31

Using the SAR and SSS Information	33
State-Level Monitoring and Planning	33
Monitoring and Modifying the Subgrant	34
Providing Feedback on the Forms	34
How to Handle Special Circumstances	35
Projects with Several Funding Sources	35
Projects that Enhance Services to the Same Victims	35
Projects that Enhance Service Quality but NOT Numbers	36
Other Ways to Get Useful Information	37
The SAR Form	38
The SSS Form	41
CHAPTER 6: CHOOSING AN EVALUATION DESIGN	51
Impact Evaluation Designs	52
Non-Experimental Impact Evaluations	52
Quasi-Experimental Designs	55
Experimental Designs	57
The Impact Evaluation Design “Decision Tree”	59
EXHIBIT 6.1: Decision Tree	60
Process Analysis	61
Key Elements	61
Design Variations	61
Practical Issues	62
Performance Monitoring	62
Key Elements	62
Design Variations	63
Considerations/Limitations	63
Practical Issues	64
Data Collection Strategies	65
Quantitative Strategies	65
Qualitative Strategies	67
The Planner’s Questions	68
Additional Considerations in Planning your Evaluation	68
What Level of Evaluation to Use	68
Designs for Projects Already in Progress	68
Informed Consent, Follow-Up Arrangements, and Confidentiality/Data Security	69
Addendum: Evaluation and Basic Research Methods	71

INTRODUCTION TO THE RESOURCE CHAPTERS	75
Example 1: Counseling Services	77
Example 2: Special Prosecution Unit	77
Example 3: Court Advocacy Program	78
EXHIBIT LM.1: Logic Model for Counseling Services	79
EXHIBIT LM.2: Logic Model for Special Prosecution Unit	80
EXHIBIT LM.3: Logic Model for Court Advocacy Program	81
CHAPTER 7: VICTIM SAFETY AND WELL-BEING: MEASURES OF SHORT-TERM AND LONG-TERM CHANGE	83
What is Short-Term and What Is Long-Term Change?	83
Measures of Short-Term Change	84
Measures of Long-Term Change: Victim Safety and Well-Being	87
Criteria Used to Select Instruments	88
A Note of Caution	89
Measures of Victim Safety	89
Measures of Psychological Well-Being	99
CHAPTER 8: DESCRIBING VICTIM SERVICES AND SUPPORT SYSTEMS	119
Institutional Affiliation, Organizational Structure, and Linkages	119
Venue	119
Structure of Activities	120
Linkages	121
Services, Activities, and Clients	121
More and Better Services and Advocacy	122
More Clients/Users, More Underserved Clients/Users	124
Increased Cultural Competence	125
TABLE 8.1: Institutional Affiliation, Organizational Structure, and Linkages	126
TABLE 8.2: Services, Activities, and Clients	128
CHAPTER 9: EVALUATING CRIMINAL AND CIVIL JUSTICE AGENCY CHANGES .	131
Don't Forget Background and External Factors	133
TABLE 9.1: Impacts on Law Enforcement Agencies	134
TABLE 9.2: Impacts on Prosecution Offices	141
TABLE 9.3: Impacts on Judges and Courts	147
TABLE 9.4: Impacts on Corrections (Probation and Parole)	153
CHAPTER 10: MEASURES OF COMMUNITY COLLABORATION	159
Elements of Community-Level Collaboration	160
Organizing Structure	160
Composition of Committee/Group	161
Levels/Source of Support	162

Communicating Effectively	163
Developing a Shared Vision	163
Establishing Systems of Conflict Resolution and Evaluation	164
Developing Trust and Mutual Respect	164
Engaging in Joint Activities and Co-Location	165
Reporting	165
Funding	165
Community Collaboration: Ultimate Outcomes	165
Creating Permanent Policy and Practice Changes	165
Treating Victims and Perpetrators Consistently	166
Creating More Options and Resources for Victims in the Justice and Human Service Systems	166
Changing Public Knowledge of VAW and Reducing Unmet Need	166
TABLE 10.1: Community Collaboration Goal 1: Establish an Effective, Stable, and Continuing Community Response to VAW	168
TABLE 10.2: Community Collaboration Goal 2: Achieve System Level Outcomes	170
TABLE 10.3: Community Collaboration Goal 3: Achieve Ultimate Outcomes	173

CHAPTER 11: MEASURING CHANGE IN COMMUNITY ATTITUDES,

KNOWLEDGE OF SERVICES, AND LEVEL OF VIOLENCE	175
Knowledge of Service Availability	175
What to Ask About	175
How, When, and Whom to Ask	177
Attitudes toward Violence Against Women	178
Measuring Community Attitudes	178
“Beyond the Usual Suspects”—Community Allies and Partners	188
Measuring Community Levels of Violence Against Women	190
Addendum: Additional Reading	191

CHAPTER 12: MEASURING PERCEPTIONS OF JUSTICE 195

Perceptions of Justice	195
Distributive Justice	195
Procedural Justice	196
Protection of Victim Rights	197
Addendum: Readings on Perceptions of Justice	198

CHAPTER 13: MEASURING THE IMPACT OF TRAINING 201

Logic Model of a Training Project	201
EXHIBIT 13.1: Typical Training Program	202
Selecting Outcomes to Evaluate	203
Avoiding Inappropriate Outcomes	203
Picking the Right Outcomes	204

Levels of Evaluation for Training Programs	205
TABLE 13.1: Options for Evaluating Training Activities	206
Level 1: Reaction Assessments	207
Level 2: Learning Assessments	209
Level 3: Behavior Change Assessments	212
TABLE 13.2: Comparing Responses Before and After Training	213
Level 4: Problem Impact Assessments	213
Assessing External Influences	214
Other Issues	215
Timing of Training Impact Evaluations	215
Methods to Evaluate Training	215
Using Training Evaluation Results to Improve Services	216
Addendum: Readings on Training Evaluations	217
CHAPTER 14: DATA SYSTEM DEVELOPMENT	221
Evaluation Methodology	222
Logic Model for a Data System Project	223
EXHIBIT 14.1: Data System Development Project	224
Documenting the Purposes and Operations of the Data System	225
Measuring the Accuracy, Reliability, and Utility of the Data System	226
Assessing User Perceptions of Accuracy, Reliability, and Utility	226
Assessing Data Quality	227
Measuring Effects of the Data System on Behavior, Performance, and Policy Change	229
Changes in User Behavior, Job Performance, and System Responses	229
Services Affecting Victims, Offenders, and Other Clients	230
Use of System Data by Government Officials and Others in Policy-Making Positions	230
Expanding the Scope or Policy Application of Studies	231
Increased Variety of Organizations Requesting Information and Data	231
New Approaches to Data Access and Integration	232
CHAPTER 15: SPECIAL ISSUES FOR EVALUATING PROJECTS ON INDIAN TRIBAL LANDS	233
The Indian Tribal Grants Program	233
The Administration of Justice in Indian Country	234
Tribal Sovereignty	234
Tribal Jurisdiction	235
Source and Scope of Tribal Jurisdictional Authority	236
Indian Country	237
Criminal Law in Indian Country	237
State Versus Tribal Jurisdiction	238
TABLE 15.1: State-by-State Overview of Public Law 280	239
Law Enforcement Issues	240

Tribal Court Systems	241
Tribal Codes	242
Protocol	243
Conclusion	245

WHERE TO GET RESOURCES DISCUSSED IN THIS *GUIDEBOOK*

Copies of this *Evaluation Guidebook*:

1. Every state STOP coordinator has both a hard copy and a computer disk copy of the *Guidebook*. Call VAWGO for the contact information for the STOP coordinator in your state. VAWGO phone number: (202) 307-6026
2. Every statewide coalition on domestic violence and/or sexual assault has a hard copy of the *Guidebook*.
 - Call the National Resource Center on Domestic Violence (800 537-2238) or the National Coalition Against Domestic Violence (303 839-1852) for the domestic violence coalition in your state.
 - Call the National Alliance of Sexual Assault Coalitions (860 282-9881) for the sexual assault coalition in your state.
3. The STOP TA Project has both a hard copy and a computer disk copy of the *Guidebook*. TA Project phone numbers: (800 256-5883 or 202 265-0967)
4. The *Guidebook* is available in photoimage and in text formats on the Urban Institute's Internet webpage—<http://www.urban.org> and look under "Recent Publications" or "Research by Author" and then "Burt." VAWGO and NIJ websites also link up to this.

NOTE: Hard copy versions of the *Evaluation Guidebook* are complete, and are legal to copy as many times as needed. Internet and disk versions of the *Evaluation Guidebook* are missing a few pages of Chapter 7 because some publishers do not grant permission to reproduce other than in hard copy format.

5. You can buy the *Guidebook* for cost plus shipping from the Urban Institute's clearinghouse. Call or email: (202 857-8687 or pubs@ui.urban.org).

Copies of evaluation instruments referenced in this *Guidebook*:

1. The STOP TA Project has hard copies of the measuring instruments described in Chapters 7 and 11 of this *Guidebook* (only hard copy is available due to copyright requirements). TA Project phone numbers: (800 256-5883 or 202 265-0967)

Further help with your evaluation:

1. The STOP TA Project will soon have the capacity to offer a limited amount of technical assistance related to evaluation. TA Project phone numbers: (800 256-5883 or 202 265-0967). PLEASE DO NOT CALL THE URBAN INSTITUTE; it is not funded to provide technical assistance.

PREFACE

This *Evaluation Guidebook* is intended as a resource for all people interested in learning more about the success of programs that try to aid women victims of violence. It has been written especially for projects funded through STOP formula grants, but has wider application to any program addressing the needs of women victimized by sexual assault, domestic violence, or stalking.

The Violence Against Women Act (VAWA), Title IV of the Violent Crime Control and Law Enforcement Act of 1994 (P.L. 103-322), provides for Law Enforcement and Prosecution Grants to states under Chapter 2 of the Safe Streets Act. The grants have been designated the STOP (Services, Training, Officers, Prosecutors) grants by their federal administrator, the Department of Justice's Violence Against Women Grants Office (VAWGO) in the Office of Justice Programs (OJP). Their purpose is "to assist States, Indian tribal governments, and units of local government to develop and strengthen effective law enforcement and prosecution strategies to combat violent crimes against women, and to develop and strengthen victim services in cases involving violent crimes against women."

A major emphasis in VAWA is also placed on collaboration to create system change, and on reaching underserved populations. System change may result from developing or improving collaborative relationships among justice system and private nonprofit victim service agencies. This emphasis on collaboration is key to the long-term ability of the legislation to bring about needed system change. Historically the relationships among these agencies in many states and communities have been distant or contentious, with little perceived common ground. VAWA is structured to bring these parties together, with the hope that they can craft new approaches which ultimately will reduce violence against women and the trauma it produces. Many groups of women have not participated in any services that were available for women victims of violence. To compensate for these service gaps, VAWA encourages states to use STOP funds to address the needs of previously underserved victim populations, including racial, cultural, ethnic, language, and sexual orientation minorities, as well as rural communities. The Act requires assessment of the extent to which such communities of previously underserved women have been reached through the services supported by STOP grants.

To accomplish VAWA goals, state coordinators in all 50 states, the District of Columbia, and five territories distribute STOP formula grant funds to subgrantees, which carry out specific projects. Subgrantees can be victim service agencies, law enforcement or prosecution agencies, or a wide variety of other agencies. They run projects addressing one or more of the STOP program's seven purpose areas:

- Training for law enforcement and/or prosecution,
- Special units for law enforcement and/or prosecution,

Evaluation Guidebook

- Policy and procedure development for law enforcement and/or prosecution,
- Data and communication systems,
- Victim services,
- Stalking, and
- Needs of Indian tribes.

This *Guidebook* is designed to help subgrantees document their accomplishments, and to offer assistance to state STOP coordinators as they try to support the evaluation activities of their subgrantees. Everyone should read the first four chapters, which introduce the reader to issues in doing evaluations and in working with evaluators:

- Chapter 1 covers reasons to participate in evaluations, and participatory approaches to conducting evaluations.
- Chapter 2 explains logic models, including why you want to have one and how to develop and use one.
- Chapter 3 describes how to get what you want from an evaluation, including how to work with evaluators.
- Chapter 4 discusses way to use evaluation results to improve your program's functioning and performance, promote your program, and avoid being misunderstood.

The only evaluation activity that VAWGO **requires** for projects receiving STOP funds is that they complete a Subgrant Award Report (SAR) at the beginning of their project and for each add-on of funding, and Subgrant Statistical Summary (SSS) forms covering each calendar year during which the project operates. Everything else in this *Guidebook* is optional with respect to federal requirements, although state STOP coordinators may impose additional evaluation requirements of their own. Therefore, if you read nothing else, all state STOP coordinators and STOP subgrantees should read Chapter 5, which explains the requirements of the SAR and SSS.

The remaining chapters of this *Guidebook* form a reference or resource section covering specialized topics. There is no need for everyone to read all of these chapters. Rather, you can go directly to the chapter(s) that provide measures relevant to your particular project. In addition, the Introduction to the Resource Chapters uses logic models to demonstrate how programs of different types might need or want to draw upon the evaluation resources from one or more of the resource chapters:

- ❑ Chapter 6 is a technical discussion of evaluation design, and would be relevant to someone who must actually design an evaluation or who intends to work closely with an evaluator to select an appropriate design.
- ❑ Chapter 7 addresses short- and long-term victim outcomes, and would be relevant to any project that provides direct services to victims.
- ❑ Chapter 8 offers a variety of dimensions that could be used to describe current status and changes in any of the following: one victim service, a “full-service” victim support and advocacy agency, and/or the network of victim services available in a locality.
- ❑ Chapter 9 addresses system changes/outcomes for the criminal and civil justice systems, and would be relevant to projects focused on policies, procedures, protocols, special units, or other types of system change.
- ❑ Chapter 10 describes potential outcomes and measures to assess community-wide system change, and would be relevant to projects attempting community coordination or collaboration activities.
- ❑ Chapter 11 focuses on ways to measure improvements in community attitudes toward violence against women, and would be relevant to any project with a community education component or that wants to assess changed community attitudes.
- ❑ Chapter 12 presents some ways to think about measuring increased perceptions of fairness and justice in the legal and court systems, for those projects interested in seeing whether any changed practices or changed system relationships have made enough difference that victims now see the systems as operating more fairly than they did before.
- ❑ Chapter 13 provides measures for the impacts of training activities, to be used by those projects with a significant training component.
- ❑ Chapter 14 focuses on measures of impact for projects that are developing or installing new data or communication systems.
- ❑ Chapter 15 discusses special issues regarding evaluations of projects on Indian tribal lands.

CHAPTER 1

GETTING STARTED: THINKING ABOUT EVALUATION

This chapter lays the groundwork for thinking about evaluation. Many readers of this *Guidebook* will never have participated in an evaluation before. Even more will never have been in charge of running an evaluation, making the decisions about what to look at, how to do it, and how to describe your results. Some readers may have had negative experiences with evaluations or evaluators, and be nervous about getting involved again. We have tried to write this *Guidebook* in a way that makes evaluation clear and helps all readers gain the benefits of evaluations that are appropriate to their programs.

Rigorous evaluation of projects supported with STOP formula grant funds is vitally important. Evaluation can:

- Document what your project accomplishes,
- Provide evidence of your project's impact and effectiveness in reaching its goals,
- Describe what kinds of participants benefit the most (and least) from project activities,
- Generate information on what strategies work best, how projects should be structured, and how to overcome obstacles, and
- Document project costs and, in some studies, assess the value of benefits.

You can use this information to:

- Determine if your project is accomplishing your objectives, for whom, and how,
- Plan and manage your project by getting feedback to identify areas that are operating according to plan and those that need attention and development,
- Identify unmet needs and gaps in services for those you are trying to reach,
- Publicize your accomplishments, and
- Raise funds for project continuation, expansion or replication.

The kind of information you will get, and what you can do with it, depends on the kind of evaluation you select. You need to start by asking what you hope to learn and how you plan to use the findings. Answer these questions: Who is the evaluation audience? What do they need to know? When do they need to know it?

You can choose from the following types of evaluation:

- ❑ **Impact evaluation** focuses on questions of causality. Did your project have its intended effects? If so, who was helped and what activities or characteristics of the project created the impact? Did your project have any unintended consequences, positive or negative?
- ❑ **Process evaluation** answers questions about how the project operates and documents the procedures and activities undertaken in service delivery. Such evaluations help to identify problems faced in delivering services and strategies for overcoming these problems. They can tell you if your project is doing what you want it to do, in the way you want to do it. They can provide guidance for practitioners and service providers interested in replicating or adapting your strategies in their own projects.
- ❑ **Performance monitoring** provides regular, consistent data on key project activities and accomplishments. Indicators of performance obtained through routine monitoring have several uses. You can use them in a process evaluation to document the activities of component parts of service delivery. You can also use these results in conjunction with project management, to identify areas in which performance expectations are not being attained. Finally, these indicators can be used as part of an impact evaluation to document project accomplishments and help you raise funds for your program.

A comprehensive evaluation will include all of these activities. Sometimes, however, the questions raised, the target audience for findings, or the available resources limit the evaluation focus to one or two of these activities. Any of these evaluations can include estimation of how much the project or project components cost. Impact evaluations can include assessments of how the costs compare to the value of benefits (cost-benefit analysis) or the efficiency with which alternative projects achieve impacts (cost-effectiveness analysis).

WHAT SHOULD BE EVALUATED?

State STOP coordinators may want to develop statewide evaluations that cover many or all of the projects they fund with STOP dollars. Alternatively, state STOP coordinators could decide that the Subgrant Award Reports and Subgrant Statistical Summaries will supply adequate basic information to describe their portfolio of subgrants, and will concentrate special evaluation dollars on special issues. To see longer term program effects on women, they might decide to fund a follow-up study of women served by particular types of agencies. Or, they might decide that they have funded some particularly innovative programs for reaching underserved populations, and

they want to know more about how well known and well respected these programs have become within their target communities.

In addition, regardless of the evaluation activities of state STOP coordinators, STOP subgrantees might want to learn more about the effects of their own particular program efforts. They might devote some of their own resources to this effort, and might have a need to work with an evaluator in the process.

This *Guidebook* is written to help both state STOP coordinators and STOP subgrantees as they develop plans to evaluate STOP projects. It contains information that can be applied statewide, to a subgroup of similar projects throughout the state or one of its regions, or to a single project.

One issue that may arise for anyone trying to use this *Guidebook* is, Am I trying to isolate and evaluate the effects of *only those activities supported by STOP funds*? Or, am I trying to understand whether a particular type of activity (e.g., counseling, court advocacy, special police or prosecution units, providing accompaniment to the hospital for rape victims) produces desired outcomes *regardless of who pays for how much of the activity*?

We strongly advocate taking the latter approach—evaluate the effectiveness of the activity, regardless of who pays for it. Your evaluations of project results for outcomes such as victim well-being, continued fear, children’s school outcomes, or the gradual disappearance of sexual assault victims’ nightmares and flashbacks need to look at the program as a whole. Unless STOP funds are being used to fund *all of an entirely new and separate activity*, it is virtually impossible to sort out what STOP funds accomplished versus what other funds accomplished. Instead, we suggest using your evaluation effort to document *that the type of activity has good effects*. Then you can justify spending any type of money for it, and can say that STOP funds have been used to good purpose when they support this type of activity. Only when your evaluation is focused on simple performance indicators is it reasonable to attribute a portion of the project achievements (e.g., number of victims served, officers trained) to STOP versus other funding sources.

A different issue you may be facing is that your program is already under way, so you cannot create the perfect evaluation design that assumes you had all your plans in place before the program started. We discuss options for handling this situation in Chapter 2.

A final issue for those thinking about evaluation is, *Is the program ready for impact evaluation*? Not all programs are ready. Does that mean you do nothing? Of course not. Process evaluation is a vital and extremely useful activity for any program to undertake. We discuss this issue in more detail in Chapter 2.

WORKING TOGETHER: PARTICIPATORY EVALUATION

In Chapter 3 we discuss many issues that arise when working with evaluators. However, because we think *attitude* is so important in evaluation, we include here in Chapter 1 a few comments on participatory evaluation.

Some evaluations are imposed from above. Funders usually require programs to report certain data to justify future funding. Funders may also hire an outside evaluator and require funded programs to cooperate with the activities of this evaluator. In many cases, the programs to be evaluated are not asked what *they* think is important about their programs, or what *they* think would be fair and appropriate measures of their own programs' performance.

We think it is important for the entire enterprise of evaluation that those being evaluated have a significant share in the decision-making about what should be evaluated, when, how it should be measured, and how the results should be interpreted. We urge state STOP coordinators and STOP subgrantees to work together to develop appropriate evaluation techniques, measures, and timetables for the types of activities that STOP funds are being used to support.

A participatory approach usually creates converts to evaluation. In contrast, an approach "imposed from above" is more likely than not to frighten those being evaluated, and make them wary of evaluation and less inclined to see how it can help them improve their own program and increase its financial support.

- ❑ **Identify goals.** The mutual design work needs to start with a discussion of what the goals of the various subgrants are. Some of these goals may be short-term, and some may be long-term. They will certainly be different for different types of projects.
- ❑ **Goals should be realistic.** Do not give yourselves the goal of "eliminating violence against women," or "our clients will be violence-free in six months," as the first will probably take a few centuries, and experience indicates that even the second is a long process. On the other hand, there is a lot of room for improvement in how justice agencies treat women victims of violence; significant shifts toward better treatment is a realistic and measurable short-term goal. So are "our clients will understand their options for action," or "our clients will leave with a better support system in place."
- ❑ **Identify data users and what they need.** At the same time, discussion needs to focus on who will use the data collected and the evaluation results, and how to ensure that the data to be collected will be useful to those who must gather it (the program people, usually), as well as to

funderson. This issue is equal in importance to agreement on goals. A good way to ensure that evaluations will fail is to design them so the data are not useful to the program people.

- **Details.** Once goals and uses of the data are agreed upon, state coordinators, subgrantees, and possibly outside evaluators can get down to the details. These include:
 - How the goals should be measured,
 - Who should actually collect the data and how,
 - Who should analyze and interpret the data,
 - Who should review the analysis and interpretation for accuracy and reasonableness, and
 - Who should have access to the data for other analyses that might serve each program differently, and how that access will work.

Including all of the interested parties in designing evaluations helps to avoid the selection of inappropriate measures and promote the selection of measures that both reflect program accomplishments and help guide program practice.

RESPECT FOR WOMEN’S PRIVACY AND SAFETY

Just as program directors and staff want their needs taken into consideration during the design of an evaluation, so too must we think about the needs of the women from whom a good deal of the data for evaluations must come. Everyone who has worked with women victims of violence knows the problems they face with having to repeat their story many times, the concerns they may have about how private information will be used and whether things they say will be held in confidence, and the problems that efforts to contact women may raise for their safety and well-being. No one wants to conduct an evaluation that will make a woman’s life harder or potentially jeopardize her safety. We discuss these issues at greater length in Chapter 6. In addition, Chapter 6 also discusses some approaches you can use to establish permission to conduct follow-up interviews, and to keep track of women between follow-up interviews. We think it is important to note here that including in your evaluation design and data collection planning some women who have experienced violence will give you the opportunity to check your plans against what they consider feasible and safe.

HOW TO USE THIS *GUIDEBOOK*

Most people will not want or need to read this whole *Guidebook*, and there certainly is no need to read it all at one sitting—no one can handle more than 200 pages of evaluation techniques and

measures. Chapters 1 through 4 contain the “basics” that everyone will want to read, and that’s only about 25 pages. After that, you can pick and choose your chapters depending on your evaluation needs. Some will be relevant to you early in your evaluation work, while others may be more useful somewhat later on.

Chapter 5. Since all state STOP coordinators and all STOP subgrantees must use the SAR and SSS to report activities to VAWGO, each of you should read Chapter 5 when your subgrant begins (or now, if you have already started). It will help you plan how you are going to record the data you will need for these reports.

Chapter 6. This chapter gets into the details of evaluation design, how to choose the right level of evaluation and specific research design for your circumstances, and how to protect women’s privacy and safety while conducting evaluations. Everyone doing any type of evaluation should read the sections on choosing the right level of evaluation for your program, and on protecting privacy and safety. The section on choosing a specific research design contains more technical language than most of the rest of this *Guidebook*, and should be used by those who are going to be getting down to the nitty-gritty of design details either on your own or with an evaluator. You may also want to use that section with a standard evaluation text, several of which are included in the chapter’s addendum.

Chapters 7 through 15. Each of these chapters takes a different purpose or goal that could be part of a STOP subgrant, and focuses on the measures that would be appropriate to use in identifying its immediate, short-, and/or long-term impact. Chapters 7 through 12 focus on particular types of *outcomes* you might want to measure (victim, system, community, attitudes). Chapters 13 and 14 address measurement options for two complex types of projects—training and data system development. Chapter 15 addresses special issues involved in doing evaluations on Indian tribal lands.

If you are doing a training project, read Chapter 13 and, if you want ultimate justice system impacts, also read Chapter 9. If you are doing a victim services project or otherwise hoping to make a difference for victims’ lives, read the victim outcomes chapter (Chapter 7) and the victim services chapter (Chapter 8). If your project involves developing and implementing the use of a new data system, read Chapter 14. If you are concerned with changing perceptions of justice within the system as other parts of the system change, read Chapter 12. Use these chapters as best suits your purpose. There is no need to commit them all to memory if they cover something that is not relevant to your project and its goals.

CHAPTER 2

DEVELOPING AND USING A LOGIC MODEL

You need to start your evaluation with a clear understanding of your project's goals and objectives. Next you need to think about the activities your project does, and your beliefs about how those activities will result eventually in reaching your project's goals. Start by drawing a diagram with four columns, A, B, C, and D. Then follow the steps below.

- ❑ **Step 1. Make a list of project goals.** Because it is often difficult to measure and document long-term goals such as reducing violence against women in your community, you want to be sure to start with short-term goals, then include some medium-term and long-term goals also. Short-term goals could include, for example (1) increasing the number of women who receive requested protection orders; (2) increasing police recruit understanding of evidence collection procedures; or (3) getting more women in underserved groups to approach your service.

Many projects have found that the process of discussing and listing goals is useful in clarifying differences in opinions among project staff and participating agencies on goals. Group them into categories and place them in Column D (far right of your diagram). You may want to have two columns: one for immediate program outcomes (D1) and one for longer term goals (D2).

- ❑ **Step 2. Make a list of project services and other activities.** Activities can include project planning steps such as convening meetings or reviewing training curriculum materials, as well as the delivery of services or products such as training police recruits or counseling victims. Group them into categories and place them in Column B.
- ❑ **Step 3. Make a list of all the background characteristics (factors) of the people involved that you think may influence the relationship between activities and goals.** For example, training may have different effects depending on the rank or years of experience of law enforcement officers; the impact of victim services may depend on the person's age, number of children, or health. Place this list at the far left of your diagram in Column A.
- ❑ **Step 4. Make a list of all the events or factors happening during or after your project activities that could influence how or whether your project accomplishes its goals.** These could include, for example, whether a victim has continued contact with an abuser, or whether law

enforcement officers are placed in positions that use the information given during a training program. Place this list in Column C.

CREATING A LOGIC MODEL

A sample logic model is shown in Exhibit 2.1, using an evaluation of a shelter-based program for domestic violence victims as a hypothetical example. Exhibit 2.2 shows a blank diagram that you can use to create a logic model for your own project.

- ❑ **Step 1. Make a list of project goals.** Boxes illustrating different categories of program goals are shown in Column D. These have been divided into immediate outcomes (Column D.1) and longer term outcomes (Column D.2). At this stage, the diagram does not list particular variables or indicators, but lists goals at a conceptual level. Goals of this program are to provide immediate safety for victims, help them get access to legal protection, plan for safety in the event of a new incident of abuse, and help victims with the needs of their children. These outcomes are expected to lead to decreases in future victimization and improvements in the victim's mental health and feelings of security.
- ❑ **Step 2. Make a list of project services and activities.** The services to be provided by the program will include providing safe places, counseling, referrals to needed services, legal advocacy, and services for, or on behalf of, their children. These are shown in boxes in Column B.
- ❑ **Step 3. Make a list of all the background characteristics (factors) of the people involved that you think may influence the relationship between activities and goals.** The problems and resources available to individuals served by the program may influence the outcomes independently of the program and may also influence what services the victims receive. Examples are shown in Column A. Victims who do not speak English will need translators and culturally sensitive counseling. Victims with very severe histories of abuse may have more extensive needs than those who seek shelter after the first incident.
- ❑ **Step 4. Make a list of all the events or factors happening during or after your project activities that could influence how or whether your project accomplishes the objectives.** Column C shows a list of things that may affect the desired outcomes either in combination with the program services or in addition to them. For example, victims with strong ties to friends and family may get additional assistance in finding a safe place or recovering from the trauma of abuse. Some police responses

Exhibit 2.1

Logic Model for Evaluation of Shelter-Based Services for Domestic Violence Victims

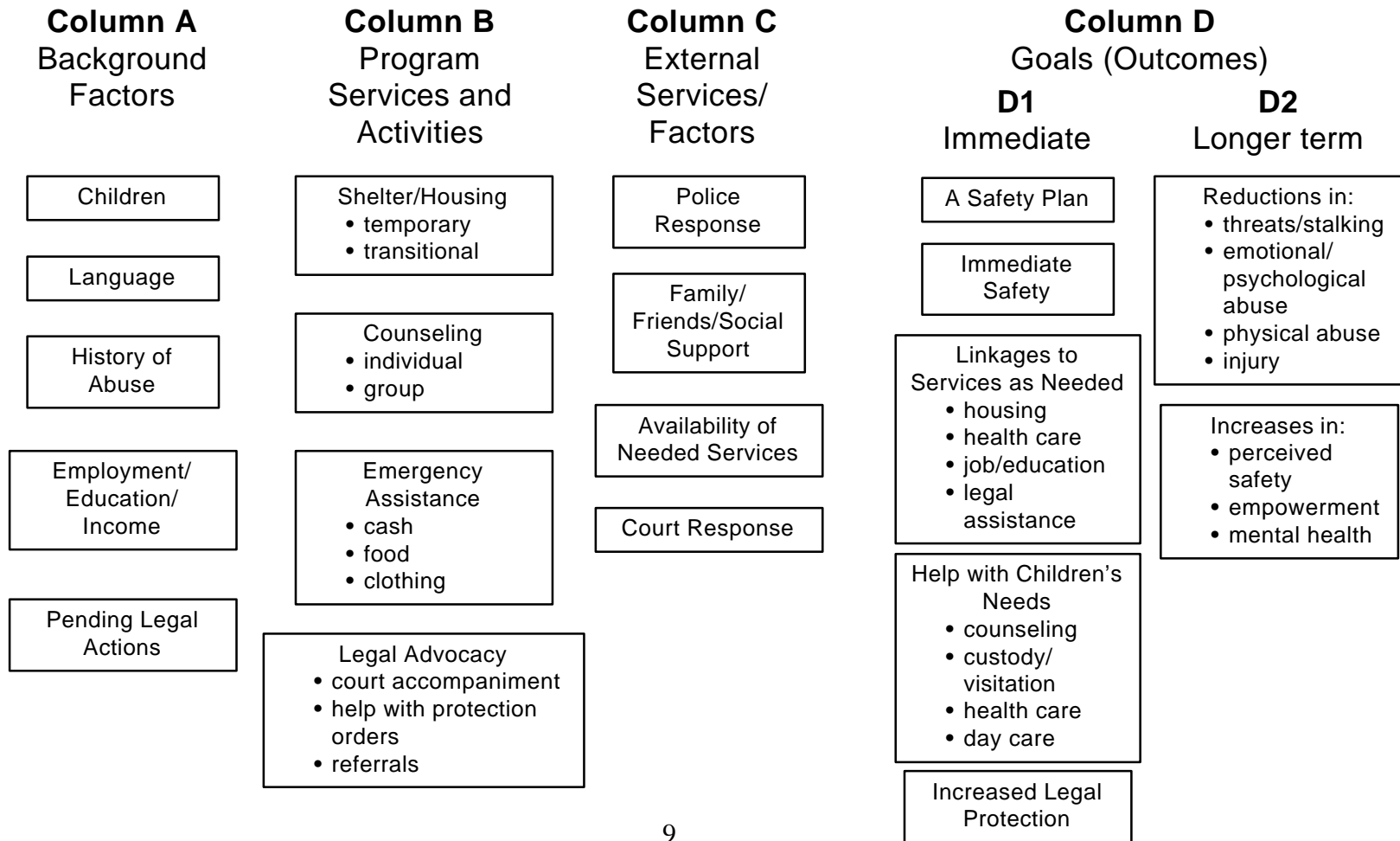


Exhibit 2.2

Logic Model for Your Program

Column A Background Factors	Column B Program Services and Activities	Column C External Services/ Factors	Column D Goals (Outcomes)	
			D1 Immediate	D2 Longer term
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

(arrest of the abuser) may have an independent effect, while others (notifying the shelter that the abuser has been released from jail) may act in combination with shelter services to protect the victim.

The diagrams can be used to plan your evaluation as follows:

- In impact evaluations, the diagram is used to spell out how, and for whom, certain project activities are expected to attain specific goals. This process involves developing a set of hypotheses to be tested in the analysis. Simple hypotheses will usually take the form of “A service from Column B leads to an outcome in Column D.” More complex hypotheses might take the form of “Column B leads to Column D, but happens faster/more if conditions in Column C are right.” Sample hypotheses for the logic model presented in Exhibit 2.1 might be:
 - Participating in group and individual counseling helps women become more empowered.
 - Receipt of legal advocacy increases legal protection and immediate safety.
 - Legal advocacy increases legal protection and immediate safety, but will do it better if the police and court responses are supportive, and will have trouble doing it if these responses are obstructive.
- In process evaluation, the diagram is used to identify expectations about how the project should work—an “ideal type”—which can then be used to assess the deviations in practice, why these deviations have occurred, and how the deviations may affect achievement of goals either positively or negatively. This vision of the program assists project managers and evaluators to identify differences from expectations (including positive and negative unintended consequences), consider possible ways to fine-tune project operations so they look more like the planned approach, or go back to the drawing board to consider changes to the original approach.
- In performance monitoring, the diagram is used to focus on which kinds of indicators (statistics) are appropriate for specific target populations, communities, or time periods. The process of making the list of strategies and outcomes often helps projects identify gaps in planning or differences among staff and partner agencies in assumptions about how the project will operate and the various responsibilities of participants.

The logic model can also be the basis for preparing a list of variables (indicators) needed to measure the concepts identified in each box. To do this, make a separate page for each box in your own logic model. Take the first page and list the variables needed for the first box. Using the example in Exhibit 2.1, the first box is “Children.” Basic variables might be: yes/no; number of children; and, for each child, age, gender, where child lives (with client, where else). Additional variables that might be relevant depending on what types of services your program offers, might be whether child(ren) were also abused, how they were abused, what school they are attending, whether they have attendance problems, etc.

Once you have listed variables, then list next to each variable the source(s) where you will (may) be able to get the information, such as court records, program records, interview with the victim, the program staff, etc. Whenever possible, we recommend getting information on the same variable from multiple sources, to verify that the information is correct (e.g., number of children) or to have several “reads” on the same variable if it is a hard one to measure. For example, you could measure reductions in victimization with information from both the victim and police records—each would tell you something different about the outcome.

Repeat these two steps (listing variables and listing data sources) for each box in your own logic model (Exhibit 2.2).

Once you know what variables you want to measure and where the information will come from, you can begin to develop data collection plans. These can take a while. You need to find out what records are maintained by agencies, decide what information is needed to construct variables, get an understanding of the limits and meaning of the records, arrange permission to get the data if you don’t already have it all in your files, and develop a protocol specifying exactly what data elements will be collected, by whom, and how. You need to decide when and how program participants can be contacted to provide information, what questions should be asked, and who will collect the information. Similarly, you will want to make lists of persons on the program staff and working with other agencies who need to be interviewed and what questions are appropriate for each. You may want to add to or change the forms your own agency uses to be sure that you are collecting the necessary information for each client.

PROGRAM EVALUABILITY

Not every program is ready for an impact evaluation (even if it could still profit from a performance evaluation and/or from performance monitoring). In addition, every program does not need the same level of evaluation. Since there are never enough evaluation resources to go around, it is important to decide whether impact evaluation is justified and feasible for your project, whether it is likely to provide useful information, and what level of evaluation is needed

(monitoring, process evaluation, or impact evaluation). Evaluation planners should ask themselves:

- Does the logic of your diagram stand up to careful scrutiny? Is your proposed link between project activities and objectives reasonable in view of the resources available and findings of earlier research? If project goals are unrealistic or the activities not well grounded in theory and/or prior evidence, then evaluation is not a good investment.
- Is the project operating at its full scope, and is it stable (not just beginning, or not just having changed or being about to change some major aspect of program operations)?
- What kinds of data will be needed, from how many people? What data are likely to be available already from existing data? Evaluations should be designed to maximize the use of available data, as long as these are valid indicators of important concepts and are reliable. Available data may, for example, include government statistics, individual and summary agency records and statistics, and information collected by researchers for other studies. If there are crucial data needs not met with existing data, resources must be available to collect any new data required.
- Are adequate resources and assets available—money, time, expertise, community and government support? Are there any factors that limit or constrain access to these resources?
- Can the evaluation be achieved in a time frame that will permit the findings to be useful in making project and policy decisions?
- To what extent does evaluation information already exist somewhere on the same or a closely related project? The answer to this question can have important implications for action. Any successful previous attempts may yield promising models for replication. Lessons learned from previous unsuccessful attempts may inform the current effort. If sufficient evidence already exists from previous efforts, the value of a new evaluation may be marginal.
- To what extent are the findings from an evaluation likely to be generalizable to other communities, and therefore useful in assessing whether the project should be expanded to other settings or areas? Are there unique characteristics of the projects to be evaluated that might not apply to most other projects? Project characteristics that are critical to success but are not replicable in other settings reduce the value of any evaluation findings.

Negative or unsatisfactory answers to enough of these questions about a particular project suggest that the project is not ready for impact evaluation, and that major impact evaluation resources would be better spent on projects for which these questions can be answered more positively.

IF NOT IMPACT EVALUATION, WHAT?

Even if your program is not ready for impact evaluation, it could still benefit from a process evaluation, and you could still collect useful information through performance monitoring. The mere exercise of answering the evaluability questions just posed often helps a program identify gaps in its practice, mismatches between its practice and its goals, and ways the program could be improved. This is one outcome of a process evaluation; other benefits of a process evaluation include achieving a greater understanding of work flow, successes and issues in service delivery, successes and issues in outreach/client finding and in follow-up, and so on. In addition, performance monitoring will provide you with basic descriptive information about who you are serving, what they are getting from your program, and how they feel about the services they have received. These options are described in more detail in Chapter 6.

EVALUATION OPTIONS FOR PROGRAMS THAT ARE ALREADY UNDER WAY

With the exception of a few carefully planned research demonstrations, most evaluations are done with programs whose operations were already under way before the evaluation started. These circumstances obviously preclude starting “at the beginning.” However, you still have a number of options open to you, including identifying and collecting data from a meaningful comparison group, and monitoring the performance of your agency. Chapter 6 discusses these options at greater length.

CHAPTER 3

MAKING EVALUATION WORK FOR YOU

In Chapter 1 we talked about the many reasons why you would want to participate in an evaluation. Assuming you are sold on the potential usefulness of evaluation for your program, how do you make sure you get a good one?

Sometimes the biggest challenge to getting useful evaluation results is finding an evaluator who understands your program and with whom you feel comfortable working. The main goal is to get what you want. It's not always easy.

This chapter offers some “down home” advice about how to get what you want from an evaluator and an evaluation. It is based on many years of experience working with programs like those being funded with STOP grants, and learning to appreciate the issues these programs face when they try to participate in evaluation.

This chapter is intended to aid state STOP coordinators as they work with evaluators charged with conducting statewide evaluations; subgrantees working with evaluators they have hired to assess their own subgrant activities; and subgrantees whose state coordinators have hired a statewide evaluator and asked the subgrantees to cooperate with a statewide evaluation.

Each of these potential users of this *Guidebook* has different needs. We would like to help the process of making these needs compatible rather than competitive or combative. This chapter assumes that readers accept the goals articulated in Chapter 1 for creating evaluations based on the participation of those being evaluated in designing the evaluation. When the evaluator or purchaser of the evaluation is the state STOP coordinator, it is all the more important that the coordinator and subgrantees arrive at a mutual agreement about what should be evaluated, how, and who has responsibility for what, before embarking on a full-fledged evaluation.

FINDING AN EVALUATOR

It may not be so easy to find an evaluator in the first place, let alone find one you will enjoy working with. You have several options: ask around; look at published reports and other published information, including using the Internet; or start at a university.

The best way to find an evaluator you will be happy with is to ask others for recommendations. People you can ask include (1) program directors in other programs like yours; (2) state or local government human services or justice system statistical agency planning agency staff, who sometimes contract out evaluations; (3) planning agency and legislative staff, who often have to use evaluation reports done by others; and (4) your state STOP coordinator. The advantage of seeking personal recommendations is that you will be able to ask your referral sources what they

liked about the people they are recommending, what they did not like, and what advice they can give you about how to work well with the evaluator. In addition, you won't have to know the name of an evaluation organization in order to find out about it.

In the absence of direct access to people who have had good evaluation experiences, you still have a number of options, but you will have to do more sorting out on your own. You can search in a library, or on the Internet, using words like "evaluation," "policy research," "research," and the like. This may turn up names of organizations, but is more likely to turn up published reports of evaluations. You can examine these reports to discover who conducted the research and, if you like their style, take it from there.

Organizations that conduct evaluations include nonprofit organizations, for-profit companies, institutes or centers at universities, and some private individuals. Some but not all of the organizations have "research," "policy research," "policy analysis," "social research," and other such words in their titles.

Universities offer another possible source of evaluation expertise, through faculty who teach research or evaluation methods in the fields of public policy, public health, criminology, sociology, community psychology, and other fields. These faculty may themselves be interested in helping design and/or conduct an evaluation, or they may be able to refer you to others who have the necessary experience.

One potential source of inexpensive or free help with evaluations is university students. If you can design your own evaluation but need help with data collection, you may be able to hook up with students who need to conduct field work interviewing for a class project. You may even be able to get a student placed with your agency for a semester or two as an intern, with the expectation that the student will focus on evaluation activities such as follow-up of previous clients or maintaining a database of services rendered by your agency. One way to establish a connection to students who may be interested in your project is to cultivate a relationship with one or more interested professors of the types described in the previous paragraph.

If you are in any of the search modes other than direct recommendations from people you know and trust, you can try to improve your chances of finding someone good by asking for references. Get from the company, organization, or consultant the names and contact information for three or four clients with whom they have worked recently (preferably within the past year or two). Then check with these people to see how they felt about their experience with the evaluator.

Regardless of the route you take, once you identify an evaluation organization, be sure that you also identify the specific people in that organization with whom programs like yours have had successful experiences. Then make sure that you work with those people, and accept no substitutes. In addition, be sure that the people you want to work with have enough time

available to work with you, and that their schedules can be adjusted to accommodate your getting the work done when you need it, not when it suits the schedule of the evaluator.

GETTING WHAT YOU WANT FROM EVALUATION

The more explicit you can be about what you want from an evaluation, the more likely you are to get it. One thing is absolutely certain—if you don't know what you want, you have only yourself to blame if you go ahead with an evaluation and are unhappy with its results.

A good evaluator can and should help you articulate what you want. If someone you are thinking about hiring as an evaluator walks into your office for the first time with an evaluation plan already on paper or in their head, murmur politely if you must, but get rid of them. They should be walking in with nothing but questions. They and you should spend hours coming to a mutual understanding of what the program is trying to accomplish (goals), exactly how it is trying to get there (its activities, services, or other behaviors), for whom, and all the steps along the way. You having in your own mind (or even better, on paper) a clear logic model or “theory” of how your program's activities accomplish your program's goals is important (see Chapter 2 for how to do this). If you already have one thought out, that is terrific. But a good evaluator will help you create one if you don't have one, or will help you refine the one you have. Do not go forward with an evaluation without one, and make sure it is one that both you and the evaluator understand.

You should be happy at the end of this meeting. You should feel you have been understood, that the evaluator “gets it” with respect to your goals, your problems, your concerns. Think of it as a marriage counseling session, in which A is not allowed to respond to what B has just said until he/she can repeat what B said to B's satisfaction (i.e., B feels understood). If your meeting has gone really well you might even feel inspired, because some of the things you have felt but not previously articulated about your program will now be explicit. Be sure to include all of your key staff in this meeting, especially those who have been around for a long time and who you feel have a good idea of how the program works and why. You should all feel at the end of this meeting that you have some new ideas or new ways of looking at your program.

Use the following do's and don'ts:

Don't be intimidated by evaluators. (Or, don't act that way, even if you are.) **Do** ask questions, and keep asking questions until you are satisfied. If you need to, write out the questions before the meeting. If you get to the end of a first meeting and have more questions than answers, insist on a second meeting. Feel satisfied with the answers to your questions, or ask some more questions. Think about how you feel when your mother comes back from the doctor's. You and she have 14 questions and she doesn't have answers to any of them, because she never asked.

But the doctor probably feels pretty good about the encounter (after all, he/she understands what he/she said, and the money is in the bank).

Don't be overwhelmed or diverted by discussions of methodologies (the "hows"). **Do** make the conversation concentrate on the "whats" until you are sure that the evaluation will focus on the proper program goals and your objectives for the evaluation. Then worry about the best way to do the study.

Don't be intimidated by jargon, about methodologies or anything else. **Do** ask what they mean when they say something you don't understand. Remember, you are proposing to pay these people money to help you. If they cannot explain what they mean in words you can understand, kick them out. They won't get better as time goes on.

If you are not happy, get divorced, or don't get married in the first place. Find another evaluator.

FOR STATE COORDINATORS

Everything does not require the same level of evaluation. Think of the data in the State Statistical Summary (SSS; see Chapter 5) as being the basic evaluation information for your whole state. The minimal information that Congress requires is included on the SSS forms. You **DO** have to ensure that all subgrantees can and do complete the SSS forms. However, evaluative information for some (many) projects can be left at this level. Use special evaluation funding (e.g., special subgrants devoted to evaluation) to focus on specific types of activities, or specific outcomes, that you feel deserve more attention and will not be documented well unless you put more money into the effort.

Remember that you get what you ask for. If you write a Request for Proposals (RFP) or sign a contract that is extremely general and unfocused, you will have to live with whatever anyone wants to send you in response. This is especially true if responses to your RFPs and acceptance of contract deliverables are handled through standard state contracting procedures.

With respect to RFPs:

- Write specific RFPs, or specific sections of RFPs, for specific projects you want to fund. Give as much detail as possible. Specify (1) elements of understanding of the problem/issue, and also of the specific program(s), you expect the bidder to present in the proposal; (2) tasks to be accomplished; (3) timetables; (4) types and levels of staffing expected; (5) areas of expertise or competence you expect the successful bidder to make available to the project; (6) evidence that significant levels of collaboration

will be practiced as part of the project; (7) past experience of the bidder in doing similar projects, including references, with contact information, for you to check and see whether people the bidder has worked with in the past have been satisfied with their work; and (8) evaluation criteria (by which you will judge the proposals) and their relative weights/importance in your assessment of a proposal.

- ❑ The more explicitly you write the project requirements and the evaluation criteria for an acceptable proposal into the RFP, the more grounds you have for rejecting proposals you think will not give you what you want, especially when they are the low bidders.
- ❑ If you think a particular type of agency, or even a specific agency by name, is the only type of agency that can do a good job with a particular type of project, limit eligibility to that type of agency, or issue sole source RFPs (but don't award the contract/grant until you get a good proposal).
- ❑ There's nothing wrong with cultivating a relationship with one or more agencies or programs that you think could do a good job on a particular project for which you fear that you will not get any competent proposals. Then you can either direct the RFP exclusively to those agencies, or, if you must do a general eligibility RFP, at least you can be sure you will have a couple of good proposals.
- ❑ If you know of and have negotiated with a good potential bidder and don't know of any others, and you have looked, you have justification for doing the RFP as a sole source.

With respect to grant/contract provisions:

- ❑ The more explicitly you write the project requirements into any grant or contract you award, the more control you have over the appropriateness and quality of the final project you will get.
- ❑ You can build into the terms of a grant or contract a lot of review, which gives you a chance to reject or require major rewrites before approving things for payment. If you are really worried about a potential grantee or contractor, put into their paperwork several review steps (outline, review, redo outline, write draft of product, review draft, rewrite, re-review, rewrite, etc; draft questionnaire, review, redraft, re-review, redraft, pretest, etc., etc., etc.).
- ❑ You can also use contract language to tie payments to final acceptance of a product, rather than to time spent or other outlays. Reject products if they

do not meet your standards, and refuse to approve payment if the final product is not satisfactory (at least one state STOP coordinator's office has already done this for what it considered a "will do more harm than good" police training manual.)

FOR SUBGRANTEES

You will not have to make decisions about how many projects to evaluate, but you may have to make decisions about which activities of your own program you want to evaluate. The issues to consider in making this decision are the same as those outlined above for state coordinators.

You may be in the position of seeking to hire an evaluator for your own program and, ultimately, of signing a contract or consulting agreement with one. Here too, the comments above for state coordinators are pertinent for you.

The unique position you are in, relative to state STOP coordinators, is that they may be designing an evaluation without your involvement, but nevertheless expecting you to cooperate with its demands. We hope this will not happen, and have written the discussion about participatory research in Chapter 1 to make clear the advantages of involving those being evaluated in designing any evaluation that affects them.

However, if the STOP coordinator in your state appears poised to require you to report data that you do not think will accurately represent your program, it is time to try as many of the following ideas or options as necessary:

- ❑ It is possible that other subgrantees in your state feel they are in the same position. Call around, see what other subgrantees are thinking, and do some organizing to generate some alternative suggestions for more appropriate evaluation. Then submit these to your state coordinator (and the statewide evaluator, if one has already been chosen).
- ❑ If your state has an active Planning Group or Committee and you know people on it or can get to them reasonably easily, call the people with interests closest to yours. Try to enlist their help to get you and other subgrantees involved in the evaluation process before it becomes cast in concrete, or to modify it even if it has become part of contract language.
- ❑ Try going through your state associations and bringing them into the negotiations (e.g., state domestic violence and/or sexual assault coalitions, state prosecutors, or law enforcement associations). Often they will be on the STOP Planning Group for your state.

- ❑ If your state has set up a committee to design a statewide evaluation, but its members are mostly or entirely researchers, try to get appointed to the committee, or get representatives of programs similar to yours onto the committee, or get one or more evaluators whom you know understand your type of program onto the committee.
- ❑ Find one or more state legislators or other politicians (e.g., someone in the Governor's office, the Governor's spouse, or the Governor) who care about violence against women, and convince them that a good and *reasonable* evaluation will help all women victims of violence. Then get them to talk with whomever necessary in the agency that administers the STOP grants.
- ❑ Call the STOP TA Project (800 256-5883, or 202 265-0697 in the Washington DC metropolitan area) for more ideas.

CHAPTER 4

USING EVALUATION INFORMATION

Evaluation is empowering. By participating in evaluation, you have had a hand in shaping the information through which someone can come to understand your program's purposes and accomplishments. You have also provided yourselves with a powerful tool for improving and expanding your program and its activities in fruitful ways. And you are very likely also to be in a position to influence the further development of supports for women victims of violence in your community. The benefits provided by a good evaluation can make all the effort seem worthwhile.

Once you have gone to the trouble to participate in evaluation activities, how can you make sure that you receive the benefits. The first, and obvious, way to use evaluation information to show others what you have accomplished—that is, you would use the information to promote your program, and to avoid having your program's purposes and achievements misunderstood by various audiences. The second way to use evaluation information is to improve your program. The third way to use it is as a lever for stimulating your community to make changes. We discuss each of these three uses of evaluation results in this chapter. In addition, other chapters examine in greater depth the issues of using data for program improvements (Chapter 6) and using data for community change (Chapter 10).

USES FOR PROGRAM PROMOTION

A good evaluation should help you promote your program. It should make your goals clear, and should produce information to show how you are moving toward achieving them. You can include this information in news stories to make sure your program is properly understood, use the information in grant proposals to get more money for your program, or use it to tell a variety of audiences about what you do, that you do it well, and that they would benefit from using your services if they ever need them.

Avoiding Misunderstandings of Your Program

When you are responsible for designing your own evaluation, you will rarely be in a situation where you feel the data collected do not adequately reflect your program's goals or accomplishments. But what if you are in the situation of having an evaluation imposed on you by a funder or by a higher level unit in your own organization, and you feel that the data requested will not tell the story of your program? What can you do then?

We hope this will not happen with the STOP grants and subgrants. Throughout this *Guidebook* we have stressed the idea of participatory evaluation—of having those being evaluated participate in decisions about evaluation design, methods, and what gets measured. In Chapter 3 we

presented many specific aspects of how this might work, and what you, as part of a program being evaluated, might do if you feel an evaluation is NOT going to do justice to your program. Here we want to say a few words about what you can do if, after all that, you are still living with an evaluation requirement imposed from above that you feel does not capture the essence of your program.

Suppose, for example, that you run a counseling service and the only thing the formal evaluation wants to know about you is who your clients were. They are asking for the usual demographic information (age, race, number of children); but they also are asking for some types of data that you think reflect an underlying suspicion of women who are victims of violence (e.g., chemical dependency of the victim, number of times victimized, number of prior relationships in which victimized [if a battering case], or prior sexual behavior [if a sexual assault case]). Further, they are NOT asking anything about what the women have done or are doing to help themselves. They are NOT asking for any documentation of the array of services you have given to the women. Nor are they asking for any evidence that the women appreciated your service, changed their behavior because of your service, felt better because of your service, understood their rights better because of your service, understood their feelings better because of your service, or any other impact of your service.

What should you do? There is no EASY solution to this problem, but that does not mean there is NO solution. What you would need to do is to collect, on your own, the information you feel DOES properly reflect your program's goals (see Chapter 7 for how to measure these facets of your program). Once you have this information, be sure to present it alongside the formal evaluation every time those formal data get reported. Develop materials that show what your program really does, and how women really feel about it. Presenting this information augments any impressions created by the minimal victim profile data with the broader picture of your clients and their involvement with your services.

There are also several last steps you could try. First, even before you collect this additional data, you could try working with other programs in your position to develop a data collection protocol that you commit to use *in common*. Then there will be strength in your numbers, both methodologically (that is, you will have more data), and politically (that is, you will have more programs working together). Second, once you have your additional data, you should send the results to whomever imposed the oversimplified evaluation. Ask that they be examined and incorporated in any reports that are issued describing evaluation results. Finally, you could try to reopen negotiations with the overall evaluation/evaluator/funder to see if in the future, the evaluation can be expanded to include more appropriate indicators of program activities and accomplishments. This might be particularly appealing since you will have demonstrated that you can provide the relevant data.

Using Evaluation Information to Get Funding

To use evaluation information to support fund-raising efforts, you would (1) identify the primary mission of any funding source you want to approach; (2) select the data from your evaluation that shows that you are addressing this mission and doing so successfully; and (3) using these data, gear your presentation to the funding source's interests and purpose.

Suppose, however, that you think there are some things that have to be accomplished before the funding source's goals can be accomplished. Your program does some of these things, and you want a funding source to support your activities, even though you can't promise that you will reach their goals every time, or quickly. This is where your logic model (see Chapter 2) can stand you in good stead, assuming that you have thought it out well and that your evaluation has documented all the intermediate steps between what you do and some ultimate goals.

Take training as an example of a preliminary step, and "more convictions" as the ultimate goal. You do training for police officers, and cannot in any way guarantee "more convictions." But if you have done your logic model well, you will be able to document a number of steps that should, ultimately, produce more convictions. For instance, you could be able to show (if you collect the right data) that compared to those who did not receive training, police receiving training are more thorough and careful in their collection of evidence, are more likely to keep the evidence secure and proof against challenge, provide better and more complete case documentation, testify more convincingly in court, and do everything with greater speed. Further, victims who interact with them feel they have been treated with courtesy and respect. Suppose also that you can show that more cases prepared by trained police go forward to the prosecutor's office than do cases prepared by untrained police.

Assume for the moment that, all other things equal, better prepared cases are more likely to result in prosecution. You now have a good argument for why your funding source should support your training activities.

Suppose, however, that after several years of police training the court records show that there have not, in fact, been increases in convictions, and your funder wants to end its support for your activities. You can use evaluation techniques to learn where the problem lies. Have prosecutors been bringing more cases using the better prepared evidence, but judges and juries resist conviction? The answer may lie with combinations of prosecutor training in how to make better presentations of evidence, how to use expert witnesses, how to protect witnesses from damaging cross-questioning, etc.; judicial training; and ultimately in community education to affect jurors' attitudes. Or, has there not been an increase in cases because prosecutors are not acting on the work of the newly trained police? Then the answer may lie in prosecutor training, in developing a special prosecution unit, etc. In any event, having a well-thought-out logic model and evaluation data to support it will let you work with your funding sources to help them understand that your

“piece of the pie” is still vital, but that more needs to be done in other sectors for the community to benefit from the good work the police are now doing.

*Using Evaluation Information to Increase
Public Appreciation of Your Program*

Program evaluation data can also be used in all sorts of promotional ways. You can include them in brochures, flyers, and annual reports that you distribute all over town. You can hold an “Open House” to which you invite reporters, representatives of other agencies that have contact with women victims of violence, representatives of agencies and organizations concerned in general with women’s well-being, and so on. Create attractive posters showing your program in action, and include on each poster one or two sentences or statistics that tell what you have accomplished. You can develop good relations with local reporters, if you don’t already have them. When they write stories about your program, be sure the stories include one or two “boxes” or “sidebars” that use evaluation data to show your successes.

USES FOR PROGRAM DEVELOPMENT¹

Perhaps the most important use for evaluation results is program improvement and development. Evaluation data are particularly useful for helping you look at your program and see what is going wonderfully, what is okay but could be improved, and what cries out for fixing (most programs have some of each). Think about evaluation as an ongoing opportunity for reflection, and for comparing your performance against what you hope your program will achieve.

From the beginning, getting ready for evaluation helps you think about your program and how it is organized. You must be able to sit down and describe to an outsider what you are trying to accomplish (your goals), how the activities you perform every day increase the chances that you will reach your goals, and what internal and external factors might make it easier or more difficult for you to reach your goals. You cannot do this well on the spur of the moment. Many programs set up a retreat or off-site meeting to do this, with an evaluator and possibly also with a facilitator. This gives program staff the luxury of sitting back and thinking about how their program works; usually no one has done this for many years, if ever. In doing this exercise, many programs identify strengths of which they are proud. However, they usually also identify weaknesses, areas of misdirected energy, issues of whether current time allocations are the best use of resources for accomplishing their goals, etc. In short, the opportunity for reflection afforded by preparing for an evaluation can stimulate program improvements even before the evaluation formally begins.

¹ See also Chapter 6, sections on process evaluation and performance monitoring, for more discussion of using evaluation data to help you improve your program.

A premise on which a good evaluator operates should be “no surprises.” You don’t want to get to the end of an evaluation, having heard nothing about findings for its duration, and be hit all of a sudden with a thick report. Even if the findings are overwhelmingly good, waiting until the end to learn them gives you very little opportunity to absorb them and figure out what they mean for potential changes in program operations. If there are some negative findings or findings about areas that need improvement, it is a lot more useful to learn about these findings as they emerge so you can discuss them and decide what to do about them. Getting a lot of negatives dumped on you at one time is discouraging and not likely to be productive, in addition to which it does not make people feel good about continuing to do evaluations.

You should set up regular feedback sessions with your evaluator to discuss evolving findings related to program processes and activities, as well as to get the perceptions and feelings of the evaluator as she or he spends time in your program and with your clients. This interaction can help both the program and the evaluator. The program staff can help the evaluator interpret the meaning of emerging findings and offer suggestions for how to gather additional information relevant to developing a full understanding of anything interesting. At the same time, program staff can use the feedback to think about whether the program needs to make some changes in the way it operates, either to enhance good performance or compensate for areas where performance may need to improve.

Another source of feedback on program operations is performance monitoring data. If you have set up data collection on program activities, services, and clients so your data system produces monthly or quarterly reports, you could present them as part of regular staff meetings and invite discussion about what they mean and how your program should respond to them. This type of open and shared discussion can help bring all staff back to an awareness of overall program goals and how daily behavior is or is not contributing to them. In most programs it is all too easy for everyone in a busy program to get overwhelmed with daily coping so they never have this type of discussion.

Regular feedback sessions from the evaluator and from a data system throughout the evaluation are the ideal. However, even if you get the results of evaluation all at once through a report from an outside evaluator, you should still set aside time for the whole staff to review them, absorb their meaning, and make decisions about how the program might want to change in light of the results. A retreat is often a helpful way to accomplish this. The evaluator should be present for part of the time, to help interpret and understand the results, and perhaps might be absent for part of the time while staff discuss what they have heard and what they might want to do about it.

Several examples may make these potential benefits more concrete. Think of the following:

- ❑ A hotline program sets up a system to collect “location in the city” information from callers. The program matches this information (a type of

performance monitoring) with police data showing areas of the city where many acts of violence against women occur, and pinpoints areas of “mismatch” (many acts of violence, but few calls). The program then establishes an outreach activity to those areas of the city.

- ❑ A program runs a court advocacy project, but does not see much change in women’s ability to obtain protection orders. An evaluator is hired to observe activity in the courtrooms served by the advocacy project. The evaluator observes several things: many women coming into court enter through a door far away from where the advocate is stationed; no information is available to let the women know the advocate is there to help them; the judges will not let the advocate speak to the women in the courtroom itself. The program uses this information to change its publicity tactics (more posters, flyers, and a videotape in several languages showing how to fill out forms); change the sitting position of the advocate; and work with court personnel to refer women to the advocate.

- ❑ Performance monitoring data indicate that a good deal of the time of highly trained counselors is taken up answering phone calls and making referrals for women who are not appropriate to the program (but who do have other needs). Program staff discuss this finding and rearrange the phone-answering tasks so that less highly-trained people answer all phone calls, freeing up counselor time for direct services to counseling clients. This results in more counseling hours becoming available to women in the community. To help the non-professional staff feel comfortable taking all the phone calls, counselors offer training in phone interviewing techniques and also develop a detailed and easily used guide to community resources to help the non-professional staff offer appropriate referrals.

- ❑ A computerized data system puts information about a batterer’s prior civil and criminal record (other DV, other violence, protection orders, plus other family-related issues such as divorce, property settlement and child custody) at a judge’s fingertips. All that is necessary is to enter a name and push a button on the computer. But sentencing practices do not change. Process evaluation reveals that most judges never push the button. The evaluation also learns that some judges do not do it because they do not trust computers; others do not do it because they believe they should treat each case as a separate offense; still others do not do it because they never received training in how to use the system. Extensive education and computer training with both judges (by other judges) and with the court

personnel who do pre-trial and pre-sentence investigations is able to change this picture substantially.

USES FOR COMMUNITY DEVELOPMENT

The example described earlier of the evaluation of police training demonstrates one of the ways that evaluation data can be used for community development. If the results of training are not as expected and further investigation documents one or more “missing links” in your community’s structure for helping women victims of violence, you have the evidence you need for trying to bring the remaining agencies on board.

Even in situations where all relevant agencies are “at the table,” ongoing evaluation results can be used to improve community coordination and perhaps develop new and more appropriate services. Suppose you have a council, task force, or coordinating body in your community, and it is looking for information that will help to prioritize new projects. Feedback from victims, systematically collected through a common questionnaire regardless of which agency or service they use, could be one way to pinpoint what needs doing the most. Polls of staff in member agencies about where the system breaks down, what types of help they need from other agencies so they can do their job better, etc., are another source of useful information for establishing priorities. Cross-training sessions are also a way to begin community development. In these sessions, staff of each agency help staff from other agencies learn the agency’s primary mission, purposes, ways of functioning, issues with other agencies, and needs for cooperation from other agencies. Everyone has a turn, and everyone stands to gain.

One could even think of developing regular forums for sharing agency news, new programs, findings from research that might interest people in other agencies, and so on. Some communities have monthly lunch meetings attended by as many as 50 or 60 people from every agency in town whose work involves women victims of violence. After eating (an important point of attraction for getting them all there), there is a short presentation (not more than 30 minutes, including discussion). Sometimes one agency is “on,” to bring the others up to date on things happening in that agency of relevance to women victims of violence, sometimes to share data, sometimes to propose new activities. Other presentations may be about a particular problem or issue, such as having someone come to explain the implications of a new law, or deciding that everyone needs to be present at the beginning of a discussion of women whose cases fall through the cracks still left in the system. Once or twice a year, these meetings can be used to present evaluation data and discuss their implications.

Collecting Data over Time

Whatever route you take to using data for community development, having the same (or better) data year after year can make an important contribution. When your community begins its activities, data about current functioning can give you a baseline against which to track your progress. Getting feedback every year (or more often if possible) about how you are doing on major community-wide goals can be a source of renewed commitment, as well as a way to reassess where you are going and what might help you get there faster. In order for this exercise to feel good, it will have to include feedback on intermediate goals so there is some hope that there will be accomplishments to report. In Chapter 2 when we talked about logic models, we emphasized the importance of adopting a realistic time frame for achieving your goals, and including many steps along the way so you could track progress and not get discouraged. The same applies to tracking community-wide progress only more so, as it is harder to change a whole community than it is to change one agency.

CHAPTER 5

USING THE SAR AND THE SSS

WHAT ARE THE SAR AND THE SSS?²

The Subgrant Award Report (SAR) is the form you complete when subgrant awards are made, to inform VAWGO how states are allocating their funds and what the subgrant projects plan to accomplish. An SAR is completed for each new or continuation subgrant awarded with each year's funds. Reports on subgrants funded with FY 1995 funds were done in the fall of 1996, and reports on FY 1996 subgrants are due in fall 1997. Beginning with reports for FY 1996 funds, the SAR can be completed and submitted to VAWGO on the computer.

The Subgrant Statistical Summary (SSS) is the follow-up to the SAR, and provides VAWGO with information on any changes in each subgrant's time period, funding, goals, or activities. It also provides feedback on what the project has accomplished during the reporting period. The SSS is a very useful tool for performance monitoring of immediate results. However, it does not give any information about the ultimate impact of a project. More and different data would be needed to determine how a specific training activity changed officers' attitudes, knowledge, and practices; whether a particular policy has been put into practice and with what effects; how a victim service program has helped women improve their lives; and so on.

The SSS is completed once a year for subgrants which were in operation during the previous calendar year. The first round of SSS reports is due in fall 1997, and will provide information on all subgrants active at any time during calendar year 1996. These might be subgrants funded with FY 1995 funds, FY 1996 funds, or both. They may have started before 1996, they may end during 1996, or they may continue into 1997. You will report on all activities from the project's start date through the end of 1996 (or the project's end date, whichever comes first). The reports will be completed on paper this year but will be available electronically beginning next year.

GETTING INFORMATION FOR THE SAR AND THE SSS

Your first question might be, How do I get the information I need to complete these forms? Both forms ask for fairly fundamental information on the wide range of subgrant projects which might be funded, and many project directors should already use record keeping systems which capture this information. However, in some cases you may need to collect additional information. This section describes the basic types of information needed for these forms and where you are likely to find them.

² The final SAR and the provisional SSS forms can be found at the end of this chapter. Instructions for both forms, plus the electronic version of the SAR, are available from VAWGO (if you are a state coordinator) or from your state coordinator (if you are a subgrantee).

The SAR may be completed by either the state administrator or the subgrantee. It asks for information which was very likely provided in the subgrant application materials. This includes fundamental descriptive information such as the amount and type of STOP funding and match; project start and end dates; geographical target area; type of violence against women to be addressed; type of subgrantee agency; and subgrant purpose areas. It also asks for information on other important concerns expressed in the STOP legislation, including efforts to address full faith and credit and underserved populations. If any information requested in the SAR was not provided in the application materials or other existing sources, state administrators and subgrantees should take the opportunity to discuss how the subgrantee will provide the information, and also what information should be included in future subgrant applications.

The new electronic version provided for the FY 1996 subgrants asks for much of the same information as the earlier paper version used for the FY 1995 subgrants, so state administrators are likely to have already established procedures for getting this information. The major difference between the forms over these two years is that where the 1995 form asked for narrative write-in answers, the 1996 form provides answers to be checked off, based on the most common types of answers received in the earlier year. *This should make it easier and quicker for you to complete the new SAR.*

The SSS will most likely be completed by the subgrantee and forwarded for review by the state administrator, unless the state agency has established procedures for subgrantees to report this type of information to the state, and for the state agency to maintain the data. The SSS is divided into nine parts (I through IX). Which parts and how many parts each subgrantee will complete depends on the project's purpose areas.

Part I of the SSS must be completed for all subgrants. It provides updates on some of the fundamental descriptive information which may have changed over the course of the project. Each of the next seven parts (Part II through Part VIII) relates to one of the seven VAWA purpose areas and asks for very basic information for each purpose area the subgrant addressed.

The number of parts you complete depends on how many purpose areas your subgrant addressed. Many subgrantees will only complete one part, as their subgrant addresses only one purpose area. So, if your subgrant was focused entirely on delivering direct victim services, you would complete only Part VI; if it focused exclusively on stalking, you would complete only Part VII.

However, some of you may have to complete more than one part if your subgrant covered two or more purpose areas (e.g., if your subgrant includes training, policy development, and data system development, you would complete Parts II, IV, and V).

Much of the information you will need to complete the SSS should be available from records routinely kept by projects, such as sign-in sheets for training sessions or logs of calls to a hotline.

Other information may be obtained from knowledgeable project staff, such as what other training-related activities were done, how many people staff the special unit, what topics are addressed in the policies you developed, the type of database you have created, and so on.

Part IX is very important. Everyone who serves victims directly should complete Part IX. Part IX asks you to describe the demographic and background characteristics of the victims you served, and is information that VAWGO is required to collect by the VAWA. ***Subgrants that focus on or have components that focus on special units, victim services, data systems, stalking, or serving Indian populations will all need to complete Part IX if they provided any services directly to victims.***

USING THE SAR AND SSS INFORMATION

Information from these sources is invaluable to VAWGO and the Urban Institute as they monitor, evaluate, and report to Congress on how the STOP program is working. ***But what can it do for you?*** If you are a state administrator, it can help you keep track of how funds are being distributed across the state, what types of projects are working particularly well, and useful directions for future funding. If you are a subgrantee, this information can help you monitor the implementation and accomplishments of your project, make improvements as needed, and document achievements to justify future funding.

State-Level Monitoring and Planning

SAR and SSS information can be used to identify statewide patterns in STOP funding; what needs are being addressed; and what needs are remaining unmet. For example, you may be interested in knowing how many subgrants or how much funding is devoted to sexual assault versus domestic violence versus stalking. You can aggregate the information from these forms to get total spending for each type of crime for a single year, or identify spending trends over time as the STOP program continues. You may also want to do the same analysis of which underserved populations are being targeted for services, or which areas of the state have not received funding. Knowing how the funding has been allocated according to whatever factors are most important in your state can be very useful in deciding future funding priorities and strategies for soliciting and receiving the proposals you are most interested in.

You can also gain very valuable information by comparing SARs and SSSs on a subgrant-by-subgrant basis. This comparison will let you answer questions about how subgrant goals or activities change over time, and to what extent the projects are achieving their objectives. Suppose, for example, you compare the SARs and the SSSs and you find that the prosecution training projects are having more difficulty staying on time or reaching potential trainees than the

law enforcement training projects. This may lead you to look for differences between law enforcement and prosecution in their training projects (e.g., what type of agency receives the funds, how the training is structured) or the general training environment (e.g., requirements, coordinating agencies, available curricula). You may then be able to identify what it is about the law enforcement training projects or environment that is more conducive to project success, and how these features can be borrowed or adapted to benefit the prosecution training projects. This may help you select future prosecution training proposals with a higher likelihood of success, and may indicate special projects needed (e.g., you may wish to fund a project to establish a coordinating agency for prosecution training).

Monitoring and Modifying the Subgrant

The information needed for the SAR and SSS can also be very useful to administrators of subgrants. The SAR provides a record of the activities and goals that you and the state STOP administrator agreed your project would address when the award was first made. The SSS you complete every year (for multiyear projects) provides a record of progress and milestones you have achieved. They can be used to improve how your program is implemented or renegotiate activities and goals if any have proven beyond the scope of the project. Of course, you need not assess progress only once a year when the SSS is due; you can monitor the project more often by getting quarterly or even monthly reports from the record keeping system you have established to provide the information requested in the SSS.

The SSS also documents what your project has accomplished—how many victims you have served, how many staff you have trained, or what policies you have developed. This information can be very useful in helping you assess what still needs to be done and how best to do it based on your experience working in the area. Being able to show that you have done successful work in the past and can identify and address unmet needs is very impressive to potential funders.

Providing Feedback on the Forms

Another very important type of information that you can provide is your opinion on the SAR and SSS forms themselves. Let your VAWGO Program Manager know what parts were difficult to use or interpret; what parts were not useful to you or did not really capture what your project is about; and what additional information you would like to provide to give a better understanding of what your project is about, or that would be more helpful in meeting your goals as you monitor, improve, and seek additional support for your statewide strategy or individual project. These forms can be modified from year to year, and your experience with them would provide very valuable insights about how they can be improved to meet everyone's needs for information.

HOW TO HANDLE SPECIAL CIRCUMSTANCES

In some cases it might not be that easy to determine program outputs—how many personnel were trained, victims served, and so on—under STOP funds. This might be because the project’s support came from more than one source, because the project is giving more or different services to the same victims it would have served without STOP funds, or because the project enhances service quality but does not increase the numbers of victims served or services offered. How do you report project activities and accomplishments on the SSS then?

Projects with Several Funding Sources

Suppose, for example, STOP funds supported a project which established a sexual assault special unit in a law enforcement agency and developed a policy on investigation and charging. This project is being supported with equal amounts of STOP, VOCA, and Byrne funds. The unit has five full-time staff who investigated 300 cases of sexual assault through 1996. The question is whether (1) to report on the SSS that the STOP funds were used to develop one-third of a policy, employ one and two-thirds staff, and investigate 100 cases; (2) to report some other allocation (perhaps the STOP funds were in fact used exclusively for policy development); or (3) to report the total outputs of the project without allocating them among funding sources.

You will provide the clearest and most comprehensive picture of what the project has accomplished by reporting in Parts III, IV, and IX the total outputs of the project (e.g., 5 staff, 1 policy, and 300 cases), even though it was supported by three sources of funding. To reflect the funding situation, make sure that your answer to Question 4 in Part I shows ALL the non-STOP funding sources being used to support this activity, including the dollar values assigned for in-kind support. Analyses of projects with multiple sources of funding done by state coordinators or national evaluators will then combine your answers to Parts III, IV, and IX with Part I, Question 4 to assess cost per unit of service, efficiency, or productivity. That is, they will compare the total inputs (STOP, other federal, state, local, and private funds and in-kind cash equivalents) to the total outputs (300 cases, 1 policy, and 5 staff). In some cases this may tell us how much was spent per staff, or per case, or per training session, or per policy, and so on. We may or may not be able to allocate costs to specific funding sources, but since you don’t need to know where the funds came from to know whether a service is efficient or productive, this is a useful approach and provides the most comprehensive information on the projects STOP funds are supporting.

Projects that Enhance Services to the Same Victims

When STOP funds support a project that lets you serve victims you would not otherwise have reached, it is clear that those are the victims you should report on in Part IX. But suppose you

were running a counseling service for sexual assault and domestic violence victims, and (Case A.1) you are using your new STOP funding to add court accompaniment for about 20 percent of your current clients, or (Case A.2) you are now able to offer victims up to 10 free counseling sessions whereas before you had to limit them to 5 free sessions. How do you characterize your services, and what victims do you include in Part IX?

CASE A.1: In Part VI, indicate on Questions 28 and 29 that your STOP subgrant gave new types of service to the same victims you would have served even without the STOP funds. Then, in Part IX, provide the characteristics of **ONLY** the victims who received court accompaniment (in addition to the counseling services you were already offering before STOP funding).

CASE A.2: In Part VI, indicate on Question 28 and 29 that your STOP subgrant gave more of the same types of service to the same victims you would have served even without the STOP funds. Then, in Part IX, provide the characteristics of the victims who received the increased number (6 or more) of free sessions.

Projects that Enhance Service Quality but NOT Numbers

Questions might also arise when STOP funds are supporting goals and activities which are not easily translated into numbers of services or numbers of victims. Suppose (Case B.1) you are using your STOP money to add a nighttime staff person to your shelter to make it more secure and make women feel safer, but you will not be serving any more women, unless (Case B.2) more women are now willing to stay in the shelter because they know it is safer. How do you report on your project and the victims it served?

CASE B.1: In Part VI, indicate on Question 28 and 29 that your STOP subgrant gave enhanced or improved services to the same victims you would have served even without the STOP funds. Then, in Part IX, provide the characteristics of **ALL** the victims who stayed in your shelter once the nighttime staff person was on board.

CASE B.2: In Part VI, indicate on Question 28 and 29 that your STOP subgrant gave enhanced or improved services to the same victims that you would have served even without STOP funding, **AND ALSO** to different victims than you would have served without STOP. Then, in Part IX, provide the characteristics of **ALL** the victims who stayed in your shelter once the nighttime staff person was on board.

However, this leaves you dissatisfied because what you really want to show is the impact of your project—that women feel safer when they stay in your shelter. You want to know whether women feel safer under the new arrangements because if they don't, you want to do something different. This is because *you want them to feel safer*—you are not doing this just to fill out forms

for VAWGO. To answer this question you will need to go beyond the performance monitoring data you provided in the SSS and gather some impact data. You can ask around informally and stop there. Or, you can use one of the outcome measures described in Chapter 7 to assess the level of safety your clients feel and whether it has changed from before to after the new security measures were put in place. In fact, as you will usually have some time between notification of award and actually getting the money, you can use this time to collect data from the women who stay in your shelter without the enhanced nighttime security to make your “before” assessment. Then use the same measures to ask the clients who stay there with the enhanced security how they feel, and you will have your “after” assessment. Compare the two to see whether you have achieved your goal.

OTHER WAYS TO GET USEFUL INFORMATION

The SSS is not the only way you can find out what is going on across your state or how your project is coming along. Your clients and colleagues have many insights and suggestions they would no doubt be delighted to share if you asked them.

State administrators can turn to the statewide planning group or advisory board for advice and assistance. Members who represent law enforcement, prosecution, and victim services agencies and associations should have a good idea how projects are working in their areas and what additional work is still needed. State administrators can also consider convening regional or statewide subgrantee conferences to get their feedback directly and to promote cross-fertilization of ideas from one community to another. A model emphasizing multiagency community teams, similar to the model used in the VAWGO conference in July 1995, might be a useful approach. State administrators may also wish to borrow another page from the federal approach and support a state-level technical assistance project. Not only would this enhance project activities, but monitoring the types and sources of questions and issues may indicate areas for further work.

Subgrantees may also gain valuable insights by consulting other members of their community. You may find that your project is having unintended side effects, positive or negative, on other agencies, and can then work with those personnel to identify appropriate responses to spillover effects. You may also find that other agencies could contribute to or benefit from expansion of your project to include them; either way your project accomplishes more. And don't overlook the people who participate in your project—you can gain many valuable insights that would otherwise have never occurred to you by directly asking prosecutors how the special unit is coming along, asking court personnel what makes your database easy or difficult to use, asking law enforcement officers what they thought of the training you provided, or asking victims whether the project's services were helpful and what else they needed. STOP project staff may be able to hold a small number of informal interviews, but larger or more structured focus groups or surveys may be best undertaken by local or state evaluators involved with the project.



**U.S. DEPARTMENT OF JUSTICE
OFFICE OF JUSTICE PROGRAMS**

SUBGRANT AWARD REPORT

STOP VIOLENCE AGAINST WOMEN FORMULA GRANT PROGRAM

1. Federal Grant Number: _____		2. Amount of Subgrant: \$ _____	
3. A) State Agency: _____ B) Subgrant Number: _____			
4. Subgrant Project Period: Start Date ____/____/____		End Date ____/____/____	
5. Type of Award: <input type="checkbox"/> New <input type="checkbox"/> Continuation		7. Match Required: YES <input type="checkbox"/> NO <input type="checkbox"/>	
6. Subgrant Project Title: _____ Subgrant Program Title: _____ (Only if different from Project Title)		Total Amount of Match: \$ _____ Cash <input type="checkbox"/> \$ _____ In-Kind <input type="checkbox"/> \$ _____	
8. Subgrantee Agency Name: _____ Address: _____ _____ _____ (State) (Zip Code)		9. Subgrantee Agency Administrator: _____ Telephone: _____ Fax Number: _____ E-Mail Address (if available): _____ Subgrant Contact Person: _____ Telephone: _____ Fax Number: _____ E-mail Address (if available): _____	
10. Focus of Subgrant: (If the subgrant has multiple focuses, please indicate the amount of funds used to address each.) <input type="checkbox"/> Domestic Violence \$ _____ <input type="checkbox"/> Stalking \$ _____ <input type="checkbox"/> Sexual Assault \$ _____ <input type="checkbox"/> Other : _____ \$ _____		11. Category of Funding (If the subgrant has multiple categories, please indicate the amount of funds addressing each): <input type="checkbox"/> Law Enforcement \$ _____ <input type="checkbox"/> Prosecution \$ _____ <input type="checkbox"/> Victim Services \$ _____ <input type="checkbox"/> Undesignated \$ _____	
12. Scope of Project: (Check one) <input type="checkbox"/> State-wide <input type="checkbox"/> Regional <input type="checkbox"/> County <input type="checkbox"/> Local (city or town) <input type="checkbox"/> Indian Tribe <input type="checkbox"/> Other: _____			
13. If "state-wide" is checked above, proceed to #14. If "regional," "county," or "local" is checked, please provide the name or a description of the geographic area(s) to be served: _____			
14. Does this subgrant address enforcement of protection orders issued by other jurisdictions within the state? <input type="checkbox"/> Yes <input type="checkbox"/> No		15. Does this subgrant address enforcement of protection orders issued by other states? <input type="checkbox"/> Yes <input type="checkbox"/> No	

16. Will this subgrant target, or make specific efforts to reach, an underserved population:

YES - the subgrant will target the following underserved population classifications.
(If you checked YES, please check all underserved populations that you make a special effort to target and/or serve)

NO- proceed to question #17

A. Geographic Location:

- Rural area
- Tribal area
- Underserved urban area

B. Ethnic Population:

- African-American
- Asian-American
- Hispanic
- Native American
- _____

C. Primary Languages spoken, other than English: _____

D. Special Needs:

- Physical Disability
- Elderly
- Children and Youth
- Sexual Orientation
- Migrant Farm Worker

Other: _____

17 A. Type of Criminal Justice Agency Funded:
(check one, if applicable)

- Law Enforcement
- Prosecution
- Courts
- Probation, Parole, or other Correctional Agency

Also check here **if** Victim Service Unit/Office within agency is a funding recipient

17 B. Other Type of Implementing Agency Funded: (check one only if 17A. does not apply)

- Nonprofit, Nongovernmental Victim Services
- Government Victim Services
- State Administrative Agency
- Tribal Government
- Multi disciplinary Team
- Professional Association
- Other: _____

18. Direct Beneficiaries -- the people attending or using the funded program:
(Check all that apply)

- Victims
- Offenders
- Children or Youth
- General Public
- Prosecution
- Judiciary (Judges, Magistrates, Clerks, etc...)
- Law Enforcement
- Probation, Parole, and other Corrections
- Private Victim Service Providers
- Health Care Providers
- Other Service Providers (e.g., mental health, housing, social service providers, etc...)
- Other: _____

<p>19. Purpose Area of Subgrant: (Please check all that apply)</p> <p><input type="checkbox"/> Training for Law Enforcement</p> <p><input type="checkbox"/> Training for Prosecution</p> <p><input type="checkbox"/> Specialized units for law enforcement</p> <p><input type="checkbox"/> Specialized units for prosecution</p> <p><input type="checkbox"/> Development of law enforcement policies, protocols, orders and services</p> <p><input type="checkbox"/> Development of prosecution policies, protocols, orders and services</p> <p><input type="checkbox"/> Development of new data and communication systems</p> <p><input type="checkbox"/> Expansion of current data and communication systems <input type="checkbox"/></p> <p>Establishment of new victim services</p> <p><input type="checkbox"/> Expansion or enhancement of existing victim services</p> <p><input type="checkbox"/> Programs addressing stalking</p> <p><input type="checkbox"/> Programs addressing Indian tribes</p> <p><input type="checkbox"/> Other: _____</p>	<p>21. Type of Service provided by Subgrant (Please check all that apply):</p> <p>21 A. Direct Services to the Public:</p> <p><input type="checkbox"/> Direct services for victims focused on helping them through the criminal and civil justice systems or other systems such as financial aid, housing, employment, etc. (Can include advocacy, forensic examinations, other collection of evidence, etc.)</p> <p><input type="checkbox"/> Direct services for victims designed to meet personal needs through counseling, therapy, safety planning, shelter, education/awareness, etc.</p> <p><input type="checkbox"/> Treatment, supervision, or other direct services to offenders.</p> <p><input type="checkbox"/> Education, counseling, or other services to children or youth.</p> <p><input type="checkbox"/> Public education or awareness.</p> <p><input type="checkbox"/> Other: _____</p> <p>21 B. Expanding Agency Capacity:</p> <p><input type="checkbox"/> Increase staffing</p> <p><input type="checkbox"/> Purchase of equipment or supplies</p> <p><input type="checkbox"/> Development of resource materials (e.g., notice of victims' rights or services, officers' or prosecutors' handbook, benchbook, materials translated into another language, etc.)</p> <p><input type="checkbox"/> Development of data/communications within an agency (e.g., a police department, prosecution office, victim service agency, court, etc.)</p> <p><input type="checkbox"/> Policy development for agencies <u>other</u> than law enforcement, prosecution, or victim services (e.g., courts, corrections, health, child welfare, other social services, etc.)</p> <p><input type="checkbox"/> Training for agencies <u>other</u> than law enforcement, prosecution, or victim services (e.g., courts, etc.)</p> <p><input type="checkbox"/> Development of special units for agencies <u>other</u> than law enforcement, prosecution, or victim services (e.g., courts, etc.)</p> <p><input type="checkbox"/> Other: _____</p> <p>21 C. Enhancing Systemwide Capacity in the Community or State:</p> <p><input type="checkbox"/> Needs or resource assessment, planning</p> <p><input type="checkbox"/> Provision of technical assistance to other agencies</p> <p><input type="checkbox"/> Development of data/communications across agencies (e.g., developing a protection order registry accessed by courts and law enforcement, or developing a method of linking together single-agency databases currently in existence)</p> <p><input type="checkbox"/> Enhancing coordination on a larger system-wide basis, across disciplines (e.g., establishing a multi disciplinary community task force, doing cross-training, or hiring a case manager to coordinate interventions of various agencies)</p> <p><input type="checkbox"/> Evaluation of STOP subgrant activities</p> <p><input type="checkbox"/> Other: _____</p>
<p>20 A. How will the effectiveness of the project be evaluated: (Check all that apply)</p> <p><input type="checkbox"/> Collection and analysis of statistical systems data</p> <p><input type="checkbox"/> Pre questionnaires for direct program participants (professionals, victims, offenders, or others)</p> <p><input type="checkbox"/> Post questionnaires for direct program participants (professionals, victims, offenders, or others)</p> <p><input type="checkbox"/> Follow-up surveys with direct participants</p> <p><input type="checkbox"/> Follow-up interviews with direct participants</p> <p><input type="checkbox"/> Follow-up focus groups with direct participants</p> <p><input type="checkbox"/> Follow-up surveys with other beneficiaries besides direct participants</p> <p><input type="checkbox"/> Follow-up interviews with other beneficiaries besides direct participants</p> <p><input type="checkbox"/> Follow-up focus groups with other beneficiaries besides direct participants</p> <p><input type="checkbox"/> Review of subgrantee reports, phone contacts, and or site visits for monitoring purposes</p> <p><input type="checkbox"/> Other: _____</p> <p>20 B. Who will be conducting the evaluation: (Check all that apply)</p> <p><input type="checkbox"/> State agency awarding subgrant</p> <p><input type="checkbox"/> Subgrantee agency personnel</p> <p><input type="checkbox"/> Independent evaluators</p>	



U.S. DEPARTMENT OF JUSTICE
OFFICE OF JUSTICE PROGRAMS

Subgrant Statistical Summary (SSS)

Subgrant Activities Through 12/31/96

STOP VIOLENCE AGAINST WOMEN FORMULA GRANT PROGRAM

PART I: GENERAL INFORMATION

1. Subgrant Number: _____ State: _____ Date: ____/____/____

Subgrant project title: _____

2. Subgrant project period: Start Date ____/____/____ End Date: ____/____/____

For this report, please provide information on subgrant activities from the project start date through December 31, 1996. If the subgrant project started in 1997, stop here and return this report.

3. \$ _____ Total amount of STOP Subgrant Award, combining both new and continuation award amounts.

3.A. Allocation of Funds by Focus Areas
(If the subgrant has multiple focuses, please indicate the amount of STOP funds used to address each, combining both New and Continuation award amounts.):

\$ _____ Domestic Violence

\$ _____ Sexual Assault

\$ _____ Stalking

\$ _____ Other. Identify: _____

4. Total amount of all funds, excluding STOP funds, that were (are) being used for this project:

\$ _____ Federal Funds. Identify: _____

\$ _____ State Funds. Identify: _____

\$ _____ Local Government Funds. Identify: _____

\$ _____ In-kind cash equivalent

\$ _____ Private Funds

5. Direct Services to the Public Supported by Subgrant (check all that apply):

Direct services for victims focused on helping them through the criminal and civil justice systems or other systems such as financial aid, housing, employment, etc. (Can include advocacy, forensic examinations, other collection of evidence, etc.)

Direct services for victims designed to meet personal needs through counseling, therapy, safety planning, shelter, education/awareness, etc.

Treatment, supervision, or other direct services to offenders

Education, counseling, or other services to children or youth

Public education or awareness

Other: _____

6. Expanding Agency Capacity Supported by Subgrant (check all that apply):

- Increased/enhanced service capacity (e.g. increased staffing, trained staff, bought equipment and/or supplies, developed resource materials, etc.).
- Developed data/communications system within an agency (e.g., a police department, prosecution office, victim service agency, court, etc.).
- Developed policy for agencies other than law enforcement or prosecution (e.g., victim services, courts, corrections, health, child welfare, other social services, etc.).
- Training for agencies other than law enforcement or prosecution (e.g., victim services, courts, etc.).
- Developed special units for agencies other than law enforcement or prosecution (e.g., victim services, courts, etc.).
- Other: _____

7. Enhancing Systemwide Capacity in the Community or State Supported by Subgrant (check all that apply):

- Needs or resource assessment, planning.
- Providing technical assistance to other agencies.
- Developing data/communications systems across agencies (e.g., developing a protection order registry accessed by courts and law enforcement, or developing a method of linking together single agency databases currently in existence).
- Enhancing coordination on a larger systemwide basis, across or within disciplines (e.g., establishing a multidisciplinary community task force, coordinating prosecution policies across the state, doing cross training, or hiring a case manager to coordinate interventions of various agencies).
- Other: Identify: _____

7.A. Evaluating STOP subgrant activities. (Check all that apply.)

- Evaluating immediate impact before recipients leave the training/service site.
- Evaluating longer-term impact on victims.
- Evaluating longer-term impact on trainees, systems, etc.
- Other: Identify: _____

PART II: TRAINING

8. Did the project develop or conduct training? Yes No
If Yes, go to #9, if No, proceed to PART III.

9. Total number of personnel trained under subgrant through 12/31/96: _____

10. Enter the number of people trained, by type of agency. The total of these categories, including "unknown," should equal the total number of people trained in #9.

___ Law enforcement ___ Prosecution ___ Courts
 ___ Corrections ___ Health care ___ Public sector victim services
 ___ Private, non-profit victim services ___ Other social service providers ___ Unknown
 ___ Other. Identify: _____

11. How many training sessions were conducted through

12. Other training activities (check all that apply):
 New training materials were/are being developed.
 Previous training materials were/are being revised or expanded.
 Other. Identify: _____

PART III: SPECIAL UNITS

13. Did this project support a special unit?
 Yes No
If Yes, go to #14, if No, proceed to PART IV.

14. What is the special unit's starting date? ___/___/___

15. A. Identify where the unit is administratively located (check all that apply, or write an explanation under Other.):

- Law enforcement agency Prosecution office Court Corrections
- Private non-profit victim services agency Public sector victim services office Health care agency
- Other social service agency Other. Identify: _____

15. B. Identify how many personnel of each type staff the special unit (in full time equivalents):

- ___ Law enforcement officers ___ Prosecutors ___ Court personnel
- ___ Correctional personnel ___ Victim advocates/counselors ___ Health care providers
- ___ Administrative staff
(i.e. secretaries, etc.) ___ Other. Identify: _____

Please complete PART IX for all cases handled by the Special Unit through 12/31/96

PART IV: POLICIES

16. Did this project address policy, procedure, protocol, administrative order, or service development? Yes No
If Yes, go to # 18, if No, proceed to PART V.

17. Policy development activities (check all that apply):

- A new policy, procedure, protocol, administrative order, or service development, was (is being) developed.
- A previous policy, procedure, protocol, administrative order, or service development, was (is being) revised or expanded.
- Other. Identify: _____

18. How many agencies of each type received a copy of the policy, procedure, protocol, administrative order, or service development, through 12/31/96? (Provide numbers for each type of agency that applies.):

- ___ Law enforcement ___ Prosecution ___ Courts ___ Corrections
- ___ Private non-profit victim services ___ Public sector victim services
- ___ Health care ___ Other social services ___ Other. Identify: _____

19. How many agencies of each type adopted the policy, procedure, protocol, administrative order, or service development, through 12/31/96? (Provide numbers for each type of agency that applies.):

- ___ Law enforcement ___ Prosecution ___ Courts ___ Corrections
- ___ Private non-profit victim services ___ Public sector victim services
- ___ Health care ___ Other social services ___ Other. Identify: _____

20. What subject area(s) does the policy, procedure, protocol, administrative order, or service development, address (check all that apply):

20. A. Law Enforcement:

- Policies on training standards and requirements, including roll-call, in-service, and academy training for officers and other personnel.
- How to enforce applicable laws including authority and decision making skills on the scene, when to make an arrest, removing weapons, enforcing orders of protection, and checking records for prior incidents and warrants.
- Collection of evidence by photographing injuries, identifying and interviewing secondary witnesses such as neighbors, etc.
- Procedures to promote officer safety.
- How to serve victims and witnesses better, including notifying victims of their rights, available services, progress on their case, and protecting children on the scene.
- What to do when an officer is involved in domestic violence, including any special procedures such as calling a supervisor to the scene and follow-up contacts.
- Other. Please Explain: _____

20. B. Prosecution:

- Aggressive prosecution, including how to prosecute cases vigorously, pro-prosecution policies, charging and plea bargaining practices, and prosecuting without the victim's testimony.
- How to structure prosecution offices and manage caseloads, including vertical prosecution, special domestic violence and/or sexual assault units, coordinated case management techniques, and case tracking systems.
- How special court structures work, such as family courts or specialized domestic violence courts.
- How to serve victims and witnesses better, including role of victim/witness staff.
- Other. Please explain: _____

20 C. Other Agencies:

- Courts. Please describe: _____
- Corrections. Please describe: _____
- Victim services. Please describe: _____
- Health care. Please describe: _____
- Other social services. Please describe: _____
- Other: _____

PART V: DATA COLLECTION

21. Did this project support data collection/communications? Yes No

If Yes, go to # 23; if No, proceed to PART VI.

22. A. Indicate what type of agency has primary responsibility for the data collection/communications system (check one):

- Law enforcement Prosecution
 Courts Corrections
 Health care Other social services
 Private non-profit victim services
 Public sector victim services
 Other. Please explain: _____

22. B. What other agencies (will) use or access the system (check all that apply):

- None. Only one agency uses or accesses this system.
 Law enforcement Prosecution
 Courts Corrections
 Health care Other social services
 Private non-profit victim services Public sector victim services
 Other. Please explain: _____

23. Where are the agencies that use the system (check only one):

- All within the same city, county or community.
 Spread across a larger region of the state.
 Spread across the entire state
 Not applicable. Only one agency uses the system.

24. What type of database or tracking system does the subgrant support? (Check all that apply.)

- Sex offender registry Criminal history
 Victim notification 911 calls
 Case tracking system Hotline calls
 Protection/restraining order tracking system
 Forms development or standardization
 Other. Please explain: _____

If any victims received direct services under this project, please complete PART IX for all victims served through 12/31/96

PART VI: VICTIM SERVICES

25. Did this project support victim services? Yes No

If Yes, go to #2; if No, proceed to PART VII.

26. What type of victim service program was funded? (Check all that apply.)

- Shelter In-person crisis counseling In-person non-emergency supportive counseling
 Court advocacy Community education Hotline/telephone counseling
 Planning, coordination, or technical assistance Other. Identify: _____

27. Were subgrant funds used to provide services directly to victims? Yes No

If Yes, continue to #28 and #29, AND complete Part IX for all victims served through 12/31/96. If No, proceed to Part VII.

28. What types of services did victims receive? (Check all that apply.)

- New types of services not previously available to them.
 Improved or enhanced versions of services already available.
 More of the same services already available.

29. What victims were served under this project? (Check all that apply.)

- The same victims already receiving services.
 New victims who would not have been served without this project.

PART VII: STALKING

30. Did this project address stalking?
 Yes No
 If Yes, go to #31; if No, proceed to PART VIII.

31. Indicate whether subgrant funds were used to (check all that apply):
 Provide direct services to the public.
 Provide training, policy development, or other professional support services
 Other: _____

If any victims received direct services under this project, please complete PART IX for all victims served through 12/31/96

PART VIII: INDIAN POPULATIONS

32. Did this project address the needs of Indian populations?
 Yes No
 If Yes, proceed to # 33; if No, proceed to PART IX.

33. Indicate whether subgrant funds were used to (check all that apply):
 Provide direct services to Indians.
 Provide training, policy development, or other professional support services
 Other: _____

If any victims received direct services under this project, please complete PART IX for all victims served through 12/31/96

PART IX: CHARACTERISTICS OF VICTIMS RECEIVING DIRECT SERVICES

34. Report the total number of victims served by this project. Then report the number of those victims served by each Purpose Area through 12/31/96.

Total Victims	Special Units	Data Collection	Victim Services	Stalking	Indian Populations

35. Report by gender the total number of victims served by this project. Then give the number of those victims served by each Purpose Area through 12/31/96.

NOTE: The total number of victims listed in #35 must add up to the number listed in #34.

Gender	Total Victims	Special Units	Data Collection	Victim Services	Stalking	Indian Populations
Male						
Female						
Unknown						

36. Report by age grouping the total number of victims served by this project. Then give the number of those victims served by each Purpose Area through 12/31/96.

NOTE: The total number of victims listed in #36 must add up to the number listed in #34.

Age	Total Victims	Special Units	Data Collection	Victim Services	Stalking	Indian Populations
0-12						
13-17						
18-25						
26-40						
41-60						
61+						
Don't Know						

37. Report the total number of victims from underserved populations that were served through 12/31/96 and the total for each Purpose Area. Then report the total numbers of these victims by each of the underserved categories in #37.A. to E., and further break those totals down into the number of victims served in each of the Purpose Areas. (If a victim represents more than one underserved category below, then count him/her in each applicable category.)

Category	Total Victims	Special Units	Data Collection	Victim Services	Stalking	Indian Populations
Total Victims from Underserved Populations						
37.A Geographic Location						
Rural						
Native American reservation						
Underserved urban						
Other Identify:						
37.B. Racial Populations						
Black						
Asian/Pacific Islander						
American Indian/Aleut/ Eskimo						
Multiracial						
Other Identify:						
37.C. Ethnic Populations						
Hispanic						
Other Identify:						
37.D. Primary Language Spoken by Victims Served, Other than English						
Spanish						
Asian languages						
Other Identify:						
37.E. Other Special Needs Populations						
Mental/ emotional disability						
Physical/ medical disability						
Migrant farm worker						
Other Identify:						

38. Identify by type of victimization the total number of victims served through 12/31/96. Within the type of victimization, give the number of victims served by the Purpose Area. (If some victims suffered multiple types of victimization, please include them under each appropriate category. In this case, these categories may add up to more than the total number of victims served.)

Type of Victimization	Total Victims	Special Units	Data Collection	Victim Services	Stalking	Indian Populations
38.A. Sexual assault (rape and non-rape sex offenses between strangers, acquaintances, friends, intimates, or family members)						
38.B. Domestic violence - physical crimes (physical assault, battery, malicious wounding, assault with a weapon, attempted murder)						
38.C. Domestic violence non-physical crimes, (violation of protective order, destruction of property, arson, harassment, telephone misuse, credible threats)						
38.D. Stalking						
38.E. Other Identify:						
Don't Know						

39. Report by victim/offender relationship the number of victims served through 12/31/96, both total and by Purpose Area. NOTE: The total of all victims in #39.A. should add up to the total number of victims reported in #34.

Category	Total Victims	Special Units	Data Collection	Victim Services	Stalking	Indian Populations
39.A Relationship of offender to victim:						
Related						
Non-related						
Don't Know						
39.B Now identify, as best you can, the number of victims in each type of related victim-offender relationship:						
Married to offender						
Separated from offender						
Divorced from offender						
Offender's parent/step-parent						
Offender's child/step-child						
Offender's partner's child						
Offender's sibling						

Member of offender's extended family						
Other relative. Identify:						
39.C. Now identify, as best you can, the number of victims in each type of non-related victim-offender relationship:						
Never married to offender but in a relationship						
Never married to offender but formerly in a relationship						
Friend of offender						
Offender's roommate, landlord, tenant, neighbor						
Offender's schoolmate, coworker, or known through work						
Stranger to offender						
Other non-relative. Identify:						

CHAPTER 6

CHOOSING AN EVALUATION DESIGN³

This chapter describes the major evaluation choices available to you and discusses the factors you need to consider in picking an approach for your evaluation. It covers the three types of evaluation described in Chapter 1—impact evaluation, process evaluation, and performance monitoring. Following an overview of evaluation designs, the chapter briefly describes a variety of data sources and data collection strategies, including quantitative and qualitative ones. You may use one or more types of evaluation, and one or more data collection strategies, in evaluating your program. Once you have an overview of options in evaluation design, the chapter helps you figure out which level of evaluation might be right for your program, what evaluation activities to select, and how to proceed when your project is already operating and now you want to introduce an evaluation. The chapter also includes a discussion of how to safeguard victim rights and well-being when you are thinking of collecting data from the individuals you serve. The chapter ends with a list of additional reading about evaluation design and basic research methods.

Parts of this chapter are important for everyone to read. These include the section on choosing the right level of evaluation for your program (among impact and process evaluations, and performance monitoring), and the section on informed consent and data security. Other sections of this chapter may seem quite technical to some readers in their discussion of evaluation design choices and the data collection methods that might go along with them. Therefore, read those sections if you intend to get seriously involved in working with an evaluator to shape the design of an evaluation, or if you intend to design your own evaluation (in the latter case, you probably will also want to consult some of the books listed in the Addendum at the end of this chapter).

Readers who are not going to be deeply involved in evaluation design can still interact effectively with evaluators without committing the more technical parts of this chapter to memory. The most important thing for these readers is to become familiar with the general evaluation options and choices (impact or process evaluation, performance monitoring, and what is meant by a comparison group and what makes a good one) by skimming the more technical sections to get a general idea of design options. Exhibit 6.1, which will be found at the end of the material on evaluation design options, gives you a quick graphic overview and summary of the technical material described in the impact evaluation design portions of this chapter. The exhibit is a “decision tree” showing the design options for impact evaluations available to you under different conditions.

³ This chapter presents an overview of evaluation options and issues, but it cannot take the place of the many books and texts that exist on evaluation design, research methods, and data collection strategies. Readers who desire these more detailed treatments may want to look at one or more of the books listed in the Addendum to this chapter.

IMPACT EVALUATION DESIGNS

Our description of impact evaluations begins with the least demanding design and moves to more elaborate designs. The following sections present the key elements of each design and variations you can consider. The strengths and limitations of each are summarized as are the general requirements of each in terms of resources such as budget and staff. As you move through these choices, the budget increases as does the extent to which you produce scientifically convincing results. However, as noted below, the best choice is often driven by a consideration of the audience for your results—who wants to know, when do they need to know, what issues do they care about, and what types of information will convince them?

We have tried to avoid unnecessary jargon in describing evaluation methods, but we do use the traditional evaluation terms to describe the people from whom you will be collecting data. Project participants are called “the treatment group” and the services they receive are called “the treatment.” Those who do not receive services are called “the control group” (if people are randomly assigned to treatment and control groups) or “the comparison group” (if some method other than random assignment is used to select this group).

Non-Experimental Impact Evaluations

Key Elements. Non-experimental impact evaluations examine changes in levels of risk or outcomes for project participants, or groups that may include project participants (e.g., all women in a particular neighborhood). Non-experimental designs do not compare the outcomes for participants to individuals or groups who do not get services.

Design Variations. You can choose from four primary types of non-experimental design: (1) comparisons of groups before and after treatment; (2) time series designs; (3) panel studies; and (4) cross-sectional comparisons after a treatment has been delivered.

The first two designs are based on analysis of aggregate data—that is, data for *groups*, *not for individuals*. In a **before and after comparison**, outcomes for *groups* of participants that enter the project at a specific time and progress through it over the same time frame are measured before and after an intervention. Your assessment of program impact is inferred from the differences in the average score for the group before and after the services. This simple design is often used to assess whether knowledge, attitudes, or behavior of the group changed after exposure to an intervention. For example, a project focused on training might ask whether the average score on knowledge about domestic violence policies increased for your group of participating police, prosecutors, or others after the training compared to the baseline score measured at the start of training. Similarly, you could measure public attitudes or beliefs before and after a public safety campaign.

A time series design is an extension of the before and after design that takes measures of the outcome variables several times before an intervention begins (e.g., once a month for the six months before an intervention starts) and continues to take measures several times after the intervention is in place (e.g., once a month for six months after the intervention). The evaluation tests whether a statistically significant change in direction or level of the outcome occurs at or shortly after the time of the intervention. For example, a project trying to increase community collaboration could begin collecting information on the number of cross-agency referrals and other collaborative actions every month for the six months before intensive collaboration development efforts begin, and for every month of the two years following the initiation of collaborative work. You could then trace the development of collaborative activity and tie it to events in the community (including the timing of stepped-up efforts to promote collaboration).

Time series measures may be collected directly from project participants. However, people also use a time series design based on information from larger groups or units that include but are not restricted to project participants. For example, rates of reported violent offenses against women for neighborhoods in which special police patrols are introduced might be used to assess reductions in violence. A time series design using publicly available data (such as the rate of violent offenses just suggested) should be considered when it is difficult to identify who receives project services, or when the evaluation budget does not support collection of detailed data from project participants. Although new statistical techniques have strengthened the statistical power of these designs, it is still difficult to rule out the potential impact of non-project events using this approach.

The next two designs examine data at the individual level (that is, data come from *individuals, not just from groups*). **Cross-sectional comparisons** are based on surveys *of project participants* that you conduct after the project is completed. You can use the data collected with this design can be used to estimate correlations between the outcomes experienced by individuals and differences in the duration, type, and intensity of services they received. This will let you draw some conclusions about plausible links between outcomes and services within your treatment group. However, you can not draw definitive conclusions about what caused what, because you do not have any type of comparison group that would let you say “it happened for those who got services, but not for those who did not get services.” **Panel designs** use repeated measures of the outcome variables for individual participants in a treatment. In this design, outcomes are measured for the same group of project participants, often starting at the time they enter the project and continuing at intervals over time. The design is similar to the “time series” design described earlier, but the data come from individuals, not from groups, and data collection rarely starts *before* the individuals enter the program or receive the intervention.

Considerations/Limitations. Correctly measuring the services received by project participants is critical in non-experimental evaluations. Because the inferences about project impact are based on response to services, differences in the type and amount of service received are critical. The

key variations in services need to be spelled out carefully in developing your logic model. Several limitations to non-experimental designs should be noted:

- ❑ First, the cross-sectional and panel designs provide only a segment of the “dose-response curve.” That is, they only give estimates of the *differences* in impact related to *differences* in the services received. These designs *cannot estimate the full impact of the project compared to no service at all*, unless estimates can be based on other information on the risks of the target population.
- ❑ Second, the designs that track participants over time (before and after, panel, and time series) cannot control for the effects of changes that would have occurred anyway without services, or for the effects of other events outside the project’s influence.
- ❑ Third, the extent to which the results can be assumed to apply to other groups or other settings is limited, because this design provides no information for assessing the extent to which outcomes for those who participated differ from those who might be eligible for services but chose not to participate. For example, if those who came to your training program were those interested in learning more about domestic violence, they might show greater attitude changes after training than a group of trainees required to come by their supervisor. Alternatively, if the project provides intensive counseling only for sexual assault victims who are still showing severe psychological symptoms one or more years after the assault, their gains in mental health might be slower than those of victims whose symptoms (e.g., of fear, anxiety, or depression) abated more quickly.

Practical Issues/Data Collection. Non-experimental designs have several practical advantages. They are relatively easy and inexpensive to conduct. Data from individuals for cross-sectional or panel analyses are often collected routinely by the project at the end (and sometimes beginning) of project participation. When relying on project records, the evaluator needs to review the available data against the logic model to be sure that adequate information on key variables is already included. If some key data are missing, the evaluator needs to set up procedures for collecting additional data items.

When individual project records are not available, aggregate statistics may be obtained from the project or from other community agencies that have information on the outcomes you care about. The primary problem encountered in using such statistics for assessing impacts is that they may not be available for the specific population or geographic area targeted by the project. Often these routinely collected statistics are based on the general population or geographic areas served

by the agency (e.g., the police precinct or the clinic catchment area). The rates of negative outcomes for the entire set of cases included may well be lower than rates for your target group, if you are trying to serve those with the most severe cases or history of violence. The larger the population or geographical area covered by the statistics, the greater the risk that any effects on program participants will be swamped by the vastly larger number of nonparticipants included in the statistics.

A more expensive form of data collection for non-experimental evaluations is a survey of participants some time after the end of the project. These surveys can provide much needed information on longer term outcomes such as rates of employment or earnings for battered women after leaving the battering situation, or psychological health for sexual assault victims one or more years after the assault. As in any survey research, the quality of the results is determined by response rate rather than by overall sample size, and by careful attention to the validity and reliability of the questionnaire items.

There are a variety of data collection strategies for use in non-experimental and other types of evaluation designs. We describe a number of them later in this chapter, after the review of evaluation designs is completed.

Quasi-Experimental Designs

Key Elements. Quasi-experimental evaluations compare outcomes from project participants to outcomes for comparison groups that do not receive project services. The critical difference between quasi-experimental and experimental designs is that the decision on who participates in the program is not random. Comparison groups are made up of individuals as similar as possible to project participants on factors that could affect the selected outcomes you want to measure. Statistical techniques are then used to control for remaining differences between the groups.

Usually, evaluators use existing groups for comparison—victims (or police officers) in the same or similar neighborhoods of the city who did not receive services (or training), or those who have similar cases in other neighborhoods. In some situations, selected staff (or precincts or court dockets) try a new “treatment” (approach to services) while others do not. When selecting a comparison group, you need to be sure that the comparison group is indeed similar to the treatment group on critical factors. If victims are to be served or officers are to be trained, those receiving new services should be similar to those who get the existing services.

Design Variations. As just described, the main way to define a comparison group is to find an existing group as similar as possible to the treatment group. The most common variation to the “whole group” approach is called “matching.” In matching, the researcher *constructs* a comparison “group” by matching individuals who do not receive treatment to individuals in the

treatment group on a selected set of characteristics. This process for constructing a comparison group runs two relatively serious threats to validity. The first is that the groups, while similar at the time of selection, may change over time due to pre-existing characteristics. As a result, changes over time may reflect factors other than the “treatment.” The second is that the researcher may have failed to use key variables influencing outcomes the matching process. These variables, which differed between the two groups at the outset, may still cause matched groups to differ on outcomes for reasons other than the treatment. To do the best possible job on selecting critical variables for matching, you should refer to the background factors which your logic model identifies as likely to influence outcomes. These factors should be used in the match.

Quasi-experimental designs vary in the frequency and timing of collecting data on outcome measures. One makes decisions about the frequency and timing of measurements after assessing the potential threats posed by competing hypotheses that cannot be ruled out by the comparison methodology. In many situations, the strongest designs are those that collect pre-project measures of outcomes and risk factors and use these in the analysis to focus on within-individual changes that occur during the project period. These variables are also used to identify groups of participants who benefit most from the services. One design variation involves additional measurement points (in addition to simple before and after) to measure trends more precisely. Another variation is useful when pre-project data collection (such as administering a test on knowledge or attitudes) might “teach” a sample member about the questions to be asked after the project to measure change, and thus distort the measurement of project impact. This variation involves limiting data collection to the end of the project period for some groups, allowing their post-project answers to be compared with the post-project answers of those who also participated in the pre-project testing.

Considerations/Limitations. Use of non-equivalent control group designs requires careful attention to procedures that rule out competing hypotheses regarding what caused any observed differences on the outcomes of interest.

A major threat in STOP evaluations may be that known as “history”—the risk that unrelated events may affect outcomes. The rapid change in laws, services, and public awareness of violence against women may affect the policies and services available to treatment and comparison groups alike. Changes may occur suddenly in large or small geographic areas, jurisdictions, or service catchment areas. For example, if one court begins using a victim advocate successfully, other nearby courts may adopt the practice or even undertake a more comprehensive project with similar goals. The same is true of prosecution strategies or law enforcement approaches. If your comparison group came from the courts or offices that leapt on the bandwagon shortly after you drew your sample, your “comparison group” has just become a treatment group.

A second threat to validity is the process of “selection”—the factors that determine who is eligible for, or who chooses to use, services. Some of these factors are readily identified and could be

used in selecting the comparison sample, or could be included in the statistical models estimating project impact. For example, if victims who do not speak English are excluded from services either formally or informally, the comparison of outcomes needs to consider English language proficiency as a control variable. Such differences may not be easy to measure during the evaluation.

Practical Issues/Data Collection. It is a challenge to build defenses or “controls” for threats to validity into evaluation designs through the selection of comparison groups and the timing of outcome observations. Even when the comparison group is carefully selected, the researcher cannot be sure that all relevant group differences have been identified and measured accurately. Statistical methods can adjust for such problems and increase the precision with which project effects can be estimated, but they do not fully compensate for the non-random design. Findings need to be interpreted extremely cautiously, and untested alternative hypotheses need to be considered carefully.

Plans for quasi-experimental evaluations need to pay close attention to the problem of collecting comparable information on control group members and developing procedures for tracking them. You may be able to collect data and provide contact information for treatment group members relatively easily because the program and cooperating agencies have continuing contacts with clients, other agencies, and the community, and have a stake in the outcome of your evaluation. Collecting comparable data and contact information on comparison groups can be difficult. If you collect more complete information for your treatment group than for your comparison group or lose track altogether of more comparison than treatment group members, not only will the evaluation data be incomplete, it will be biased—that is, it will provide distorted and therefore misleading information on project impact. The best way to avoid bias from this problem is to plan tracking procedures and data collection at the start of the evaluation, gathering information from the comparison group members on how they can be located, and developing agreements with other community agencies, preferably in writing, for assistance in data collection and sample member tracking. These agreements are helpful in maintaining continuing contact with your sample in the face of staff turnover at the agencies involved.

Quasi-experimental designs may employ a variety of quantitative and qualitative approaches to gather the data needed to draw conclusions about a project and its impact. Data collection strategies are described below, once we have reviewed all of the options for evaluation design.

Experimental Designs

Key Elements. Experimental designs are considered the “gold standard” in impact evaluation. Experiments require that individuals or groups (e.g., trainees, police precincts, courtrooms, or victims) be assigned at random (by the flip of a coin or equivalent randomizing procedure) to one

or more groups prior to the start of project activities. A “treatment” group receives particular services designed to achieve clearly specified outcomes. If several new services are introduced, the experiment can compare multiple treatment groups. A “control” group continues to receive the services in existence prior to the introduction of the new project (either no services or already existing services). The treatment group outcomes are compared to outcomes for alternative treatment groups and/or to a control group to estimate impact. Because chance alone determines who receives the project services, the groups can be assumed to be similar on all characteristics that might affect the outcome measures. Any differences between treatment and control groups, therefore, can be attributed with confidence to the effects of the project.

Design Variations. One design variation is based on a random selection of time periods during which services are provided. For example, new services may be offered on randomly chosen weeks or days. A version of this approach is to use “week on/week off” assignment procedures. Although not truly random, this approach closely approximates random assignment if client characteristics do not vary systematically from week to week. It has the major advantage that project staff often find it easier to implement than making decisions on project entry by the flip of a coin on a case-by-case basis. A second design variation is a staggered start approach in which some members of the target group are randomly selected to receive services with the understanding that the remainder will receive services at a later time (in the case of a school or classroom, the next month, semester, or year). One disadvantage of the staggered start design is that the observations of outcomes are limited to the period between the time the first group completes the project and the second group begins. As a result, it is generally restricted to assessing gains made during participation in relatively short-term projects.

Limitations/Considerations. Although experiments are the preferred design for an impact evaluation on scientific grounds, random assignment evaluations are not always the ideal choice in real life settings. Some interventions are inherently impossible to study through randomized experiments for legal, ethical, or practical reasons. Laws cannot be enforced selectively against a randomly selected subset of offenders or areas in a community. Access to legal protections cannot be curtailed. For example, protection orders cannot be issued to victims only during selected weeks. Essential services should not be withheld. However, it may be possible to randomly assign alternative services or responses if the relative merits of the alternatives are unknown.

You need to ask yourself whether the results that are likely to be obtained justify the investment. Experiments typically require high levels of resources—money, time, expertise, and support from project staff, government agencies, funders, and the community. Could the answers to evaluation questions—and subsequent decisions on project continuation, expansion, or modification—be based on less costly, less definitive, but still acceptable evaluation strategies? The answer is often “yes.”

Practical Issues/Data Collection. Experimental designs run the most risk of being contaminated because of deliberate or accidental mistakes made in the field. To minimize this danger, there must be close collaboration between the evaluation team and the project staff in identifying objectives, setting schedules, dividing responsibilities for record-keeping and data collection, making decisions regarding client contact, and sharing information on progress and problems. Active support of the key project administrators, ongoing staff training, and communication via meetings, conference calls, or e-mail are essential.

Failure to adhere to the plan for random assignment is a common problem. Staff are often intensely committed to their clients and will want to base project entry decisions on their perceptions of who needs or will benefit most from the project—although these judgments may not be supported by later research. Thus it is important that the evaluator, not project staff, remain in charge of the allocation to treatment or control group.

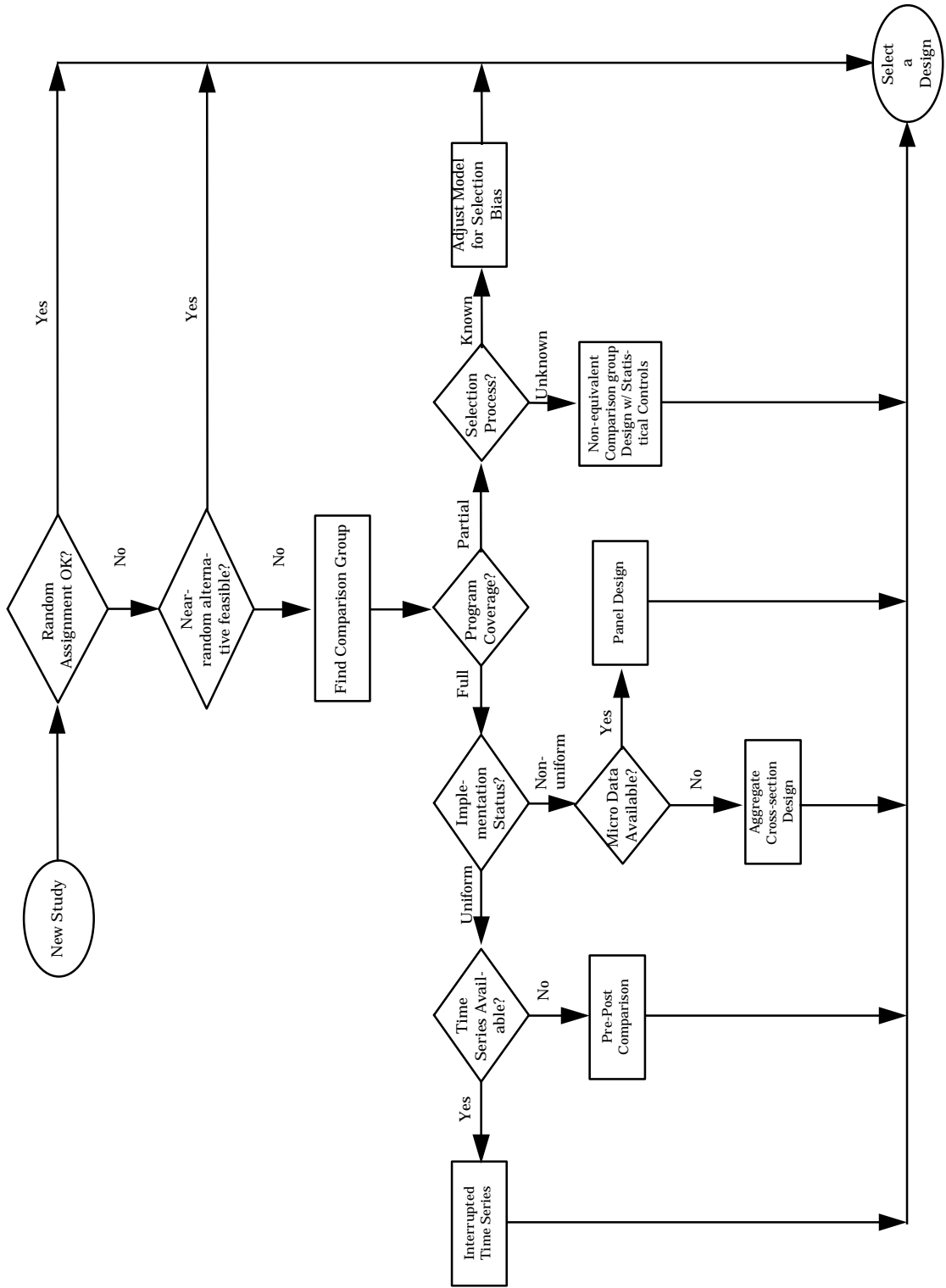
As in quasi-experimental evaluations, lack of comparable information for treatment and control group members can be a problem. Experiments generally use both agency records and data collected from individuals through questionnaires and surveys. To assure access to these individuals, quasi-experimental evaluations need to plan for data collection and tracking of sample members at the start of the project and get agreements with agencies and consent procedures with individuals in place early in the process.

Along with all other types of impact evaluation, quasi-experimental designs often combine quantitative data with qualitative information gathered through process evaluation in order to understand more about the program when interpreting impacts on participants. Another issue is documenting what parts of the program each participant received. If the project services and content change over time, it may be difficult to determine what level or type of services produced the outcomes. The best strategy is to identify key changes in the project and the timing of changes as part of a process evaluation and use this information to define “types of project” variations in the project experience of different participants for the impact analysis.

THE IMPACT EVALUATION DESIGN “DECISION TREE”

Exhibit 6.1 is a “decision tree” taken from Harrell (1996), organized around a set of questions to which the program wanting to conduct an impact evaluation answers “yes” or “no.” With each answer the program advances closer to a decision about the type of impact evaluation most appropriate for its circumstances and resources. This decision tree is a quick graphic way of summarizing the foregoing discussion about alternative impact evaluation designs and their requirements. If your program is ready for impact evaluation, the “decision tree” may help you to think about the type of evaluation that would best suit your program.

EXHIBIT 6.1
 Process for Selecting Impact Evaluation Designs



PROCESS ANALYSIS

Key Elements

Process evaluations rarely vary in basic design. Most involve a thorough documentation and analysis of activities of the program. A good process analysis design is guided by a set of core questions: Is the project model being implemented as specified and, if not, how do operations differ from those initially planned? Does the program have unintended consequences and unanticipated outcomes and, if so, what are they and who is affected? What is the view of the project from the perspectives of staff, participants, and the community? The answers to these questions are useful in providing guidance to policy makers and project planners interested in identifying key project elements and in generating hypotheses about project impact that can be tested in impact analyses.

Design Variations

Process evaluations vary in the number of projects or sites included. Most process evaluations focus on a single project or site. However, some undertake comparative process analysis. Comparative process analysis requires that observations, interviews, and other data collection strategies be structured in advance around a set of questions or hypotheses about elements of implementation believed to be critical to project success. Comparative process analysis allows the evaluation to make assessments about alternative strategies and is useful in generalizing the findings to other settings or jurisdictions. This strategy is used to assess which approach is most successful in attaining goals shared by all when competing models have emerged in different locations. It requires purposely selecting sites to represent variations in elements or types of projects, careful analysis of potential causal models, and the collection of qualitative data to elaborate the causal links at each site.

Most design uncertainties in process evaluation involve deciding what information will be collected, from whom and how. Process evaluation can be based solely on qualitative data. However, qualitative data are usually combined with quantitative data on services produced, resources used, and outcomes achieved. Qualitative data collection strategies used in process evaluation include semi-structured interviews with those involved in project planning and operations; focus groups with project planners, staff, or participants; and researcher observations of project activities. Data collection strategies for use with all types of evaluation are described below, following the presentation of performance monitoring.

Practical Issues

In a process evaluation, it is often difficult to decide on what information is truly key to describing program operations and what information is simply extraneous detail. In selecting relevant data and posing questions about program operations, the evaluator needs to refer carefully to the logic model prepared at the start of the project, although it is permissible and important in process evaluation to revise the original logic model in light of findings during the evaluation.

Analysis of qualitative data requires considerable substantive knowledge on the part of the evaluator. The evaluator needs to be familiar with similar projects, respondents, and responses, and the context in which the project is operating. Your evaluator will need to be able to understand the project's historical and political context as well as the organizational setting and culture in which services are delivered. At the same time, the evaluator needs to maintain some objectivity and separation from project management in order to be able to make an unbiased assessment of whether responses support or refute hypotheses about the way the project works and the effects it has.

Collecting qualitative data also requires skilled researchers who are experienced in interviewing and observing. Data must be carefully recorded or taped. Notes on contextual factors and interim hypotheses need to be recorded as soon as possible after data collection. When using interview guides or semi-structured interview protocols, interviewers must be trained to understand the intent of each question, the possible variety of answers that respondents might give, and ways to probe to ensure that full information about the issues under investigation is obtained.

PERFORMANCE MONITORING

Key Elements

Performance monitoring is used to provide information on (1) key aspects of how a system or project is operating; (2) whether, and to what extent, pre-specified project objectives are being attained (e.g., numbers of women served by a shelter, increases in cases prosecuted, improved evidence collection); and (3) identification of failures to produce project outputs (this kind of data can be used in managing or redesigning project operations). Performance indicators can also be developed to (4) monitor service quality by collecting data on the satisfaction of those served; and (5) report on project efficiency, effectiveness, and productivity by assessing the relationship between the resources used (project costs and other inputs) and the output and outcome indicators.

If conducted frequently enough and in a timely way, performance monitoring can provide managers with regular feedback that will allow them to identify problems, take timely action, and subsequently assess whether their actions have led to the improvements sought. Performance measures can also stimulate communication about project goals, progress, obstacles, and results among project staff and managers, the public, and other stakeholders. They focus attention on the specific outcomes desired and better ways to achieve them, and can promote credibility by highlighting the accomplishments and value of the project.

Performance monitoring involves identification and collection of specific data on project outputs, outcomes, and accomplishments. Although they may measure subjective factors such as client satisfaction, the data are numeric, consisting of frequency counts, statistical averages, ratios, or percentages. **Output** measures reflect *internal* activities: the amount of work done within the project or organization. **Outcome** measures (immediate and longer term) reflect progress towards project goals. Often the same measurements (e.g., number/percent of women who filed for a protection order) may be used for both performance monitoring and impact evaluation. However, unlike impact evaluation, performance monitoring does not make any rigorous effort to determine whether these outcomes were caused by project efforts or by other external events.

Design Variations

When projects operate in a number of communities, the sites are likely to vary in mission, structure, the nature and extent of project implementation, primary clients/targets, and timeliness. They may offer somewhat different sets of services, or have identified somewhat different goals. In such situations, it is advisable to construct a “core” set of performance measures to be used by all, and to supplement these with “local” performance indicators that reflect differences. For example, some victim service projects will collect detailed data on the needs of women or the history of domestic violence, while others will simply have data on the number provided with specific services. Performance indicators need to be constructed so that results can be compared across projects in multi-site projects.

Considerations/Limitations

Indicators of outcomes should be clearly differentiated from elaborate descriptions of the population served. For example, there is a tendency of funders to ask for nitty-gritty details about program clients, when they should be asking what the program has done for these women. Take the case of victim services programs under VAWA. The governing legislation specifies only a few victim characteristics as the information that must be reported. This is quite different from the more important information of what the programs did for victims (services provided), and whether the victims benefited from the services. We probably need only basic information about

victims, and might do better to concentrate our evaluation effort on understanding the short- and long-term outcomes that the program has helped them achieve. Chapter 7 lays out what these might be, from a sense of being heard and understood, to living in safety and peace of mind.

In selecting performance indicators, evaluators and service providers need to consider:

- ❑ **The relevance of potential measures to the mission/objective of the project.** Do process indicators reflect project strategies/activities identified in mission statements or written agreements between agencies? Do outcome indicators cover the full range of identified objectives? Do indicators capture the priorities of project planners and staff?
- ❑ **The comprehensiveness of the set of measures.** Does the set of performance measures cover inputs, outputs, and service quality as well as outcomes? Does it include relevant items of customer feedback?
- ❑ **The project's control over the factor being measured.** Does the project have influence/control over the outputs or outcomes measured by the indicator? If the project has limited or no influence over the outputs or outcomes being measured, the indicator may not fairly reflect project performance.
- ❑ **The validity of the measures.** Do the proposed indicators reflect the range of outcomes the project hopes to affect? Are the data free from obvious reporting bias?
- ❑ **The reliability and accuracy of the measures.** Can indicators be operationally defined in a straightforward manner so that supporting data can be collected consistently over time, across data gatherers, and across communities? Do existing data sources meet these criteria?
- ❑ **The feasibility of collecting the data.** How much effort and money is required to generate each measure? Should a particularly costly measure be retained because it is perceived as critically important?

Practical Issues

The set of performance indicators should be simple, limited to a few key indicators of priority outcomes. Too many indicators burden the data collection and analysis and make it less likely that managers will understand and use reported information. At the same time, the set of indicators should be constructed to reflect the informational needs of stakeholders at all levels—community members, agency directors, and national funders. Most importantly, the

performance indicators should reflect key activities defined as central to the project in the logic model.

Regular measurement, at least quarterly, is important so that the system provides the information in time to make shifts in project operations and to capture changes over time. However, pressures for timely reporting should not be allowed to sacrifice data quality. For performance monitoring to take place in a reliable and timely way, the evaluation should include adequate support and plans for training and technical assistance for data collection. Routine quality control procedures should be established to check on data entry accuracy and missing information. At the point of analysis, procedures for verifying trends should be in place, particularly if the results are unexpected.

The costs of performance monitoring are modest relative to impact evaluations, but still vary widely depending on the data used. Most performance indicator data come from records maintained by service providers. The added expense involves regularly collecting and analyzing these records, as well as preparing and disseminating reports to those concerned. This is typically a part-time work assignment for a supervisor within the agency. The expense will be greater if client satisfaction surveys are used to measure outcomes. An outside survey organization may be required for a large-scale survey of past clients; alternatively, a self-administered exit questionnaire can be given to clients at the end of services. In either case, the assistance of professional researchers is needed in preparing data sets, analyses, and reports.

DATA COLLECTION STRATEGIES

Quantitative Strategies

There are many types of quantitative data collection strategies and sources. The interested reader can pursue more details through the references provided at the end of this chapter. Here we present only the briefest descriptions of the most common types of data:

- ❑ **Data from and records of public agencies, including the program being evaluated** (on services received, cases filed, case outcomes, recidivism, injury, etc.). These have the advantage of already existing, so no new data collection is needed, and often some parts of the data are published in public documents. However, they have several disadvantages for which your design might want to try to compensate (1) they only report on people who contacted public agencies; (2) they rarely contain all of the information you want; (3) they usually cover a geographical area that contains many more people than you will be serving, thereby making it very hard to see an effect; (4) if you want to use actual records rather than

published reports, they are not always easy to obtain from the agency (both in terms of getting permission and in terms of extracting information from files). Published data will never be at the individual level, but can still be useful in characterizing the experiences of districts, neighborhoods, or cities. Data from files *can* be accessed at the individual level, but it takes a lot of work. Further, when relying on public records you need to consider whether the treatment group is more closely monitored than the comparison group—which could lead to an error in interpretation. For example, an intensive supervision probation program involves very careful monitoring of offenders. This careful monitoring may mean that probation officers know more about what offenders are doing than regular probation officers do, and may result in higher rates of probation violation because of that knowledge, even if the two groups of offenders actually commit violations at the same rate.

- **Systematic data collection directly from individuals**—through questionnaires, exit surveys, telephone or mail surveys, or in-person interviews. Any type of evaluation can use this data source, and many do so. Going directly to the people served (and to comparison group members) is often required and useful for a number of reasons (1) you can get the data you want, rather than the data that an agency has collected for other purposes; (2) you can sometimes get a more honest and complete reading on people’s experiences than they would tell an official agency; (3) you will not be burdening program staff with the extra work of data collection; (4) if you have a comparison or a control group, you have no other source of information about them other than to go to them directly (that is, they are not participants in the program, so the program cannot tell you anything about them). Further, collecting information in different ways from the treatment and the comparison group can introduce considerable measurement error.

Surveys need to have careful sampling plans that describe who is included in the study and how eligibility is defined. You need to know who is represented by the sample so it will be clear to what groups the findings can be generalized. Other issues in survey design may require considerable expertise. You need questions that are well constructed, easily understood by the respondent, and unambiguously interpreted. They need to be phrased in neutral ways to avoid a response bias. You need to consider who is administering the survey and in what setting. The willingness of respondents to give complete and accurate answers may depend on their

confidence in the interviewer and the protections of privacy built into the survey procedures.

Qualitative Strategies

Qualitative data collection strategies are extremely useful. They can stand by themselves, as they do in certain types of process evaluation or case studies. Or, they can be used in combination with quantitative methods as part of virtually any of the designs described in this chapter. As with the quantitative strategies just described, the interested reader can pursue more details through the references provided at the end of this chapter. Qualitative strategies include:

- ❑ **Semi-structured interviews**, which contain specific questions about particular issues or project practices. The “semi” aspect of these discussion guides refers to the fact that a respondent may give as long, detailed, and complex a response as he or she desires to the question—whatever conveys the full reality of the project’s experience with the issue at hand. If some issues have typical categories associated with them, the protocols will usually contain probes to make sure the researcher learns about each category of interest.
- ❑ **Focus groups**, which seek to understand attitudes through a series of group discussions guided by one researcher acting as a facilitator, with another researcher present to take detailed notes. Five or six general questions are selected to guide open-ended discussions lasting about an hour and a half. The goals of the discussions may vary from achieving group consensus to emphasizing points of divergence among participants. Discussions are tape-recorded, but the primary record is the detailed notes taken by the researcher who acts as recorder. Less detailed notes may also be taken publicly, on a flip-chart for all to see, to try to achieve consensus or give group members the chance to add anything they think is important. Soon after a particular focus group, the recorder and facilitator summarize in writing the main points that emerged in response to each of the general questions. When all focus groups are completed, the researchers develop a combined summary, noting group differences and suggesting hypotheses about those differences.
- ❑ **Observations** of project activities, which may be guided by structured or semi-structured protocols designed to ensure that key items reported in interviews are verified and that consistent procedures for rating project performance are used across time and across sites. In addition, process

evaluation includes notes maintained by researchers on meetings and activities they see while collecting data on the project.

THE PLANNER'S QUESTIONS

1. Who is the audience for the evaluation? Who wants to know, what do they want to know, when do they need the information, and what types of data will they believe?
2. What kinds of evaluation should be included? Impact evaluation, process evaluation, performance monitoring, or all three?
3. What does the logic model indicate about the key questions to be asked?
4. What kinds of data can be collected, from whom, by whom, and when?
5. What levels of resources—budget, time, staff expertise—are required? What are available?

ADDITIONAL CONSIDERATIONS IN PLANNING YOUR EVALUATION

What Level of Evaluation to Use

Every project can do performance monitoring, regardless of whether you can find a good comparison group or whether you undertake a full-fledged process or impact evaluation. Collecting data to describe clients served can show you changes over time, and whether you are meeting certain goals, such as increasing the proportion of your clients who come from underserved populations or reaching clients earlier in their process of deciding to leave a batterer. Routinely collecting data on which services you have given people and who gets them allows you to track whether everyone who needs certain services gets them, which are your most and least frequently used services, what types of services are least likely to be available, and so on. For police and prosecution agencies, such tracking can also document where people get “stuck” in the system, and perhaps help you unplug important bottlenecks.

In addition to performance monitoring, most projects can benefit from some level of process evaluation, in which you compare your processes to your logic model and see where the problems lie. A good process evaluation can help you improve your program, and can also get you to the point where conducting an impact evaluation will be worth the investment.

Designs for Projects Already in Progress

Many projects cannot begin evaluating “at the beginning,” because they are already operating at full strength when the evaluation begins. You need not let this stop you. You can still construct meaningful comparison groups in a number of ways, and you can certainly begin collecting data on your own clients as soon as you know the evaluation is going to proceed.

For comparison groups, you can use participants in other programs that do not have the type of intervention you are doing (i.e., collect data from participants in a program similar to yours but across town, or in the next county, which does not have the upgraded services your STOP grant provides), or you can use participants in your own program who predate the enhanced services (i.e., collect follow-up data on women who went through your program before your STOP grant started). If necessary when doing this, you can also collect information on participant characteristics that your old intake forms did not include.

Even without a comparison group, you can do performance monitoring for clients beginning as soon as (or even before) you get your STOP money. As described above, you can learn a lot from performance monitoring, and it can be of great help by providing feedback to help you improve your program. In addition, you can institute “exit interviews” or other exit data collection, through which you can get important feedback about client perceptions of and satisfaction with services. Chapter 7 offers for some ideas about what to measure at these interviews, and how to do it.

Informed Consent, Follow-Up Arrangements, and Confidentiality/Data Security

Ethical considerations dictate a careful review of the risks and benefits of any evaluation design. The risks to victims, project staff, and offenders need to be enumerated and strategies to minimize them should be developed. Studies of violence against women need to be particularly sensitive to avoiding “secondary victimization” through data collection procedures that could cause psychological or emotional trauma, place the victim (particularly in family violence cases) at risk from the offender, or reveal private information including the woman’s status as a victim.

A review of whether the evaluation procedures meet acceptable standards for the protection of the individuals and agencies being studied should be conducted before work begins. Many funders require a formal review of the research design by a panel trained in guidelines developed to protect research participants. Even when such review is not required, explicit consideration of this issue is essential. Two considerations should be part of this review—informed consent, and confidentiality/data security.

Informed consent refers to what you tell people about what you want from them, the risks to them of participating in the research/evaluation, the benefits that might accrue to them from participating, and what you intend to do to protect them from the risks. With respect to women victims of violence from whom you wish to collect data on impacts at some later time, it also involves establishing permission for follow-up and procedures for recontact that will safeguard the woman. You owe it to your evaluation participants to think these matters through and write out a clear and complete statement of risks and protections. Then, before you gather any information from women, share this disclosure with them and get their consent to continue. Some funders will require that you get this consent in writing. Informed consent is relevant not only with evaluation participants who have been victims, but is also necessary with *anyone* you gather information from, including agency employees, volunteers, and members of the general public.

If you want to collect follow-up information on impacts, you will need **permission to recontact**, and will also need to set up **safe procedures** for doing so. Even if you have no immediate plans to conduct follow-up data collection, if you are beginning to think about doing an evaluation and realize that you *might* need to recontact women in the future, consider setting up a permission procedure *now*. Set up a form for this, that includes the text of the appeal you will make (see below), plus spaces to note agreement or refusal and, if agreement, the woman's name and contact information. With every woman who receives help from your agency, say something like:

“We are very interested in improving our services to help women more. To do this, it would be very helpful if we could contact you at some future time to learn about what has happened to you, whether you think our efforts helped, and what more we could have done to assist you. Would you be willing to have us contact you again, if we can work out a safe way to do so? [if yes...] What would be the safest way for us to contact you in the future?”

Work out acceptable arrangements with the woman and write down the particulars on the form. If possible, it would also be best to have her sign the form to indicate affirmative consent to follow-up.

How you will assure the **confidentiality and data security** of the information they give you is the final thing you need to tell people as part of informed consent. You need to think through risks to victims of telling their stories, and risks to agency employees of answering questions about how things “really” work. Then you need to develop procedures to guard their data so the risks do not materialize (that is, you need to ensure that they do not suffer repercussions should they be identified as a source of information that reflects negatively on an agency). Above all, you should tell research participants **up front** how you intend to handle the data they give you—will you cite them by name, will you disguise the source of your information, or will you report only grouped data that does not identify individuals. They can then decide for themselves how much they want to reveal. Whatever you tell them, that's what you must do, or you will have violated

the understanding under which they were willing to share their perceptions, opinions, and information with you. If you promised not to cite them in any way that would make them identifiable, then don't break your promise. If you want to be able to cite them, then tell them so up front. If you have data with people's names attached to it and you have promised to keep their information confidential, then you will have to develop security procedures to maintain that confidentiality (e.g., keeping it in locked file cabinets, putting only an ID number on the data, and keeping the key that links ID number and name in a separate, locked drawer, limiting access to the data to those people who have committed themselves to respect the conditions of confidentiality you have promised).

ADDENDUM: EVALUATION AND BASIC RESEARCH METHODS

General Issues

Bausell, R. (1986). *A Practical Guide to Conducting Empirical Research* (1st edition). New York: Harper & Row.

Eichler, M. (1988). *Nonsexist Research Methods*. Boston: Allen & Unwin.

Wholey, J., Hatry, H.P. and Newcomer, S. (1994) *Handbook of Practical Program Evaluation*. San Francisco, CA: Jossey-Bass.

Harrell, A., with M. Burt, H. Hatry, S. Rossman, J. Roth, and W. Sabol. (1996). *Evaluation Strategies for Human Service Programs: A Guide for Policymakers and Providers*. Washington, DC: The Urban Institute Press.

Kosecoff, J. and Fink, A. (1982). *Evaluation Basics: a Practitioner's Manual*. Newbury Park, CA: Sage

Schmidt, R., Scanlon, W., and Bell, S. (1979). Evaluability Assessment: Making Public Programs Work Better. *Human Services Monograph Series, 14*, 4-5, Washington, DC: Urban Institute.

Wholey, J. (1994). Assessing the Feasibility and Likely Usefulness of Evaluation. In Joseph S. Wholey, Harry P. Hatry, and Katherine E. Newcomer, *Handbook of Practical Evaluation*. pp. 15-39. San Francisco: Jossey-Bass.

Leigh, J.W. (1997). *Communicating for Cultural Competence*. Boston: Allyn & Bacon.

Defining Your Research Issue

Smith, H.W. (1991). *Strategies of Social Research: The Methodological Imagination*. Ft. Worth, TX: Harcourt Brace College Publishers.

Weiss, C.H. (1995). "Nothing is as practical as a good theory: Exploring theory-based evaluation for comprehensive communities initiatives for children and families." In Connell, J.P., Kubisch, A.C., Schorr, L.B., and Weiss, C.H. (Eds.), *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*, Pp. 65-92. Washington, DC: Aspen Institute.

A Review of Research Designs

Bailey, K.D. (1987). *Methods of Social Research*. New York: The Free Press.

Mitchell, M. and Jolley, J. (1995). *Research Design Explained*. Orlando, FL: Harcourt Brace College Publishers.

Rebovich, D.J., and Panagopoulos, J. (1996). *Measuring Impact: a Guide to Program Evaluation for Prosecutors*. Alexandria, VA: American Prosecutors Research Institute.

Weisberg, H.F., Krosnick, J.A., and Bowen, B.D. (1996). *An Introduction to Survey Research, Polling, and Data Analysis*. Thousand Oaks, CA: Sage.

Case Study, Implementation Assessment, and Other Qualitative Methods

Kirchner, R.A., Przybylski, R. and Cardella, R.A. (1994). *How to Assess Program Implementation*. Thousand Oaks, CA: Sage.

Krueger, R. (1994) *Focus Groups: A Practical Guide for Applied Research* (2nd Edition). Thousand Oaks, CA: Sage.

Miles, M.B. and Huberman, A.M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook* (2nd edition). Mill Valley, CA: Sage.

Patton, Michael Q. (1987). *How to Use Qualitative Methods in Evaluation* (2nd edition). Mill Valley, CA: Sage.

Patton, Michael Q. (1996). *Utilization-Focused Evaluation* (3rd edition). Mill Valley, CA: Sage.

Robert K. Yin. (1994). *Case Study Research: Design and Methods* (2nd edition). Newbury Park, CA: Sage.

Surveys and Questionnaires

Fowler, F.J., Jr. (1993). *Survey Research Methods* (revised edition). Newbury Park, CA: Sage.

Rossi, P., Wright, J. and Anderson, A. (1985). *Handbook of Survey Research*. New York: Academic Press.

Dillman, D.A. (1978). *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley. (a good “bible” for all the practical details of actually doing surveys and achieving acceptable response rates)

Experimental and Quasi-Experimental Design

Dennis, M.L. and Boruch, R.F. (1994). Improving the Quality of Randomized Field Experiments: Tricks of the Trade. *New Directions for Program Evaluation*.

Cook, T.D. and Campbell, D.T. (1979). *Quasi-Experimentation*. Chicago: Rand-McNally.

Causal Modeling—Regression as a General System, for (Almost) Any Type of Data

Dunn, O. (1987). *Applied Statistics: Analysis of Variance and Regression*. New York: Wiley.

Glantz, S.A. and Shanker, B.K. (1990). *Primer of Applied Regression and Analysis of Variance*. New York: McGraw Hill.

Blalock, H.M., Jr. (1985). *Causal Models in Panel and Experimental Design*. Hawthorne, NY: Aldine de Gruyter.

Kerlinger, F.N. (1986). *Foundations of Behavioral Research*. Orlando, FL: Harcourt Brace College Publishers.

Cost Analyses

Levin, Henry M. (1983). *Cost-Effectiveness: A Primer*. Beverly Hills, CA: Sage.

Performance Monitoring

Hatry, H. and Kopczynski, M. (1997). *Guide to Program Outcome Measurement for the U.S. Department of Education*. Working Document. U.S. Department of Education, Planning and Evaluation Service. Washington, DC: The Urban Institute Press.

INTRODUCTION TO THE RESOURCE CHAPTERS

The remaining chapters in this *Guidebook* provide resources to help you measure and evaluate your program(s). The first six chapters (Chapters 7 through 12) focus on **the types of outcomes** you may need to measure. As explained below, any given program may need to draw on resources from several chapters to get a complete picture of what the program has accomplished. The next two chapters (Chapter 13 on training and Chapter 14 on data system development) describe evaluation issues and measurement approaches to two fairly complex activities that can be funded with STOP grants. The last chapter offers some background and critical contextual information about conducting evaluations of programs on Indian tribal lands, as these pose some unique challenges of both program development and evaluation.

Most STOP-funded projects will need to draw on at least one of these resource chapters; many projects will need to incorporate the suggestions of several chapters into their evaluation design. The following brief chapter descriptions repeat some information from the Preface and Chapter 1, but augment it by indicating the types of projects that would benefit from reading each chapter:

- ❑ Chapter 7 focuses on immediate and long-term outcomes for women victims of violence. You would want to look at the measures in this chapter for any project that ultimately wants to make a difference for women's short- or long-term outcomes. That probably includes most projects funded under STOP.
- ❑ Chapter 8 focuses on indicators of the scope, variety, and organization of victim services, from within a given agency up to a network that encompasses the whole community. Projects devoted to victim services would want to read this chapter, as would projects trying to achieve changes within the justice systems and/or increased community collaboration.
- ❑ Chapter 9 focuses on changes that might occur within specific civil and criminal justice agencies as they work to improve their handling of cases involving violence against women. Any project for which these are immediate or ultimate outcomes (e.g., training, special units, policies and procedures, data systems) would want to look at the measures in this chapter.
- ❑ Chapter 10 focuses on changes that might occur within the service network of an entire community (including victims services, law enforcement, prosecution, the judiciary, and other elements of the network) as it attempts to develop more cooperative and collaborative arrangements for handling situations involving violence against women. Projects of any type that have a collaborative or system-building aspect would want to look at this chapter.

- ❑ Chapter 11 focuses on community-wide issues such as attitudes toward violence against women, levels of violence against women, and signs that businesses and other community elements with low historical involvement in the issue are beginning to get involved. Projects trying to bring more elements of the community into partnership in reducing violence against women, and any project that has the reduction of violence against women as its ultimate goal, might want to look at this chapter.
- ❑ Chapter 12 focuses on women’s perceptions that the civil and criminal justice systems do or do not treat them fairly. Since we hope that system changes (as indicated by measures from Chapters 9 and 10) improve the fairness of the system, the measures suggested in this chapter might be used to get a reading on whether women perceive that the fairness of their treatment has increased.
- ❑ Chapter 13 describes the issues to address when conducting evaluations of training projects, including four levels of evaluation: reaction, learning, behavior change, and impact on violence against women. The chapter contains its own logic model for a training project. It also includes some concrete measures of immediate outcomes, especially attitude change on the part of trainees.
- ❑ Chapter 14 describes the issues to address when conducting evaluations of projects that are developing data systems for justice agencies (mostly), although the same issues pertain to data systems for victim services. It also contains its own logic model. In addition to projects that directly address the data systems purpose area, any project that is *developing or using* a data system to track project activities should take a look at this chapter. This might include a victim services project that uses a client tracking system, and special unit project that has a case tracking data system, and so on.
- ❑ Chapter 15 lays out factors that are important to understand when conducting evaluations of violence against women projects on Indian tribal lands. Many of the factors discussed in this chapter could be considered background or antecedent variables in a logic model framework—things that need to be understood and documented to provide the context for the project’s eventual success or failure. Also included in the chapter is some advice for working with people who are involved with these projects, so that issues of cultural competence do not cloud the progress of the evaluation.

The remainder of this introduction presents several logic models, starting with simpler ones and progressing to the more complex. For each element in these logic models, we refer to one or more of the chapters in this resource section, to give you an idea how you might use the material

in these chapters to construct a full evaluation design. For more examples of complex program models, you might also want to look at the logic models for training and for data system development included in Chapters 13 and 14, respectively.

EXAMPLE 1: COUNSELING SERVICES

Exhibit LM.1 shows the logic underlying an evaluation of a relatively simple counseling program. The basic service, counseling, is shown in Column B, which also indicates that client case records are expected to be the source of data to document the types and amounts of counseling provided to clients. A variety of possible immediate and longer-term outcomes for clients (effects of counseling) appear in Column D; the primary data source would be client interviews, constructed to include some of the measures found in Chapter 7. The simplest evaluation of this program would involve collecting data relevant only to Columns B and D, on services received and victim outcomes. Client interviews would be required to obtain these outcome data.

You can complicate the simple evaluation considerably, which will mean more work but probably also more knowledge about what really makes a difference if you do. You can measure background factors—in this case, pertinent characteristics of the women coming for counseling—which appear in Column A. You would use information about client characteristics to help you understand what types of women are helped most by which types of counseling. Another complication is external factors that might increase or decrease the likelihood that counseling will produce the desired outcomes. These are shown in Column C, and include the availability of other supportive services from the same agency or from other agencies in the community (see Chapters 8, 9, and 10); other stressors in each woman's life (see Chapter 7); and justice system actions in each woman's case. The expected sources for both background and external factors in this model include intake forms, assessment forms, and/or research interviews with clients.

Remember that whether or not you measure them, background and external factors are always present. You may leave them out of your data collection plans, but you can't leave them out of your thinking. They should be represented in your logic model so you will be sure to consider (1) what you will miss if you leave them out, and (2) the limits on your ability to interpret findings because you have limited the scope of the variables available for analysis.

EXAMPLE 2: SPECIAL PROSECUTION UNIT

Exhibit LM.2 shows the logic underlying an evaluation of a special prosecution unit. The activities of the unit itself are shown in Column B, while the immediate and longer-term outcomes

are shown in Column D. Outcomes include some that apply to women victims of violence (see Chapter 7) and some that apply to criminal justice system changes (see Chapter 9). As with Example 1, the simplest evaluation one could do on a special prosecution unit would examine outcomes (Column D) in relation to inputs (Column B). Greater complexity could (should) be introduced by including background factors (Column A, in this case characteristics of the sexual assault and/or domestic violence cases handled by the unit, and their similarity to or difference from the totality of cases handled before the unit began operations, such as whether the unit gets only the tough cases). One might also include external factors in the analysis; Column C suggests a number of external factors pertinent to a woman's decision to pursue a case (e.g., her level of danger, the quality of her support system, or other stresses in her life). Measures for these factors can be found in Chapter 7. Other external factors may also be relevant to particular evaluation situations. Exhibit LM.2 indicates that quite a variety of data sources may be necessary for a full treatment of this logic model, and that one might want to look at Chapters 7, 9, 13, and 14 in the process of planning the data collection.

EXAMPLE 3: COURT ADVOCACY PROGRAM

Exhibit LM.3 describes a logic model for a court advocacy program to help women coming to civil court for a protection/restraining/stay away order. Column B shows the direct activities of the program, which one would document through case records and process analysis, including observations. Column D shows a variety of outcomes, including some that are personal to the woman and some that would also be considered system outcomes (following through to a completed permanent order). Column A suggests some background factors of the women seeking orders that might affect both their use of the advocacy services and their own ultimate outcomes. Column C indicates some external realities of court accommodation to the program (or failure to do so) that might make a difference for service efficacy. Data sources would be case records, possible data system data, process analysis, and client interviews.

These brief examples as guides to the resource chapters and their connection to logic models will hopefully provide you with a practical basis for jumping into the resource chapters themselves. They do not answer all questions—for example, they do not give you variables to use in describing clients (although you could use the SSS Part IX data fields as a start for this, and add much more that your own program wants to know). Nor do the chapters give you a system for describing and counting services, as this is much too variable across programs. However, the resource chapters do offer a rich array of ideas and measures covering topics and issues that most programs will need to include in an evaluation. Their specifics should be helpful to you. In addition, just reading the chapters may give you some practice in the way that an evaluator might think about measuring things, and this practice may help you develop the details of your own evaluation.

Exhibit LM.1

COUNSELING SERVICES

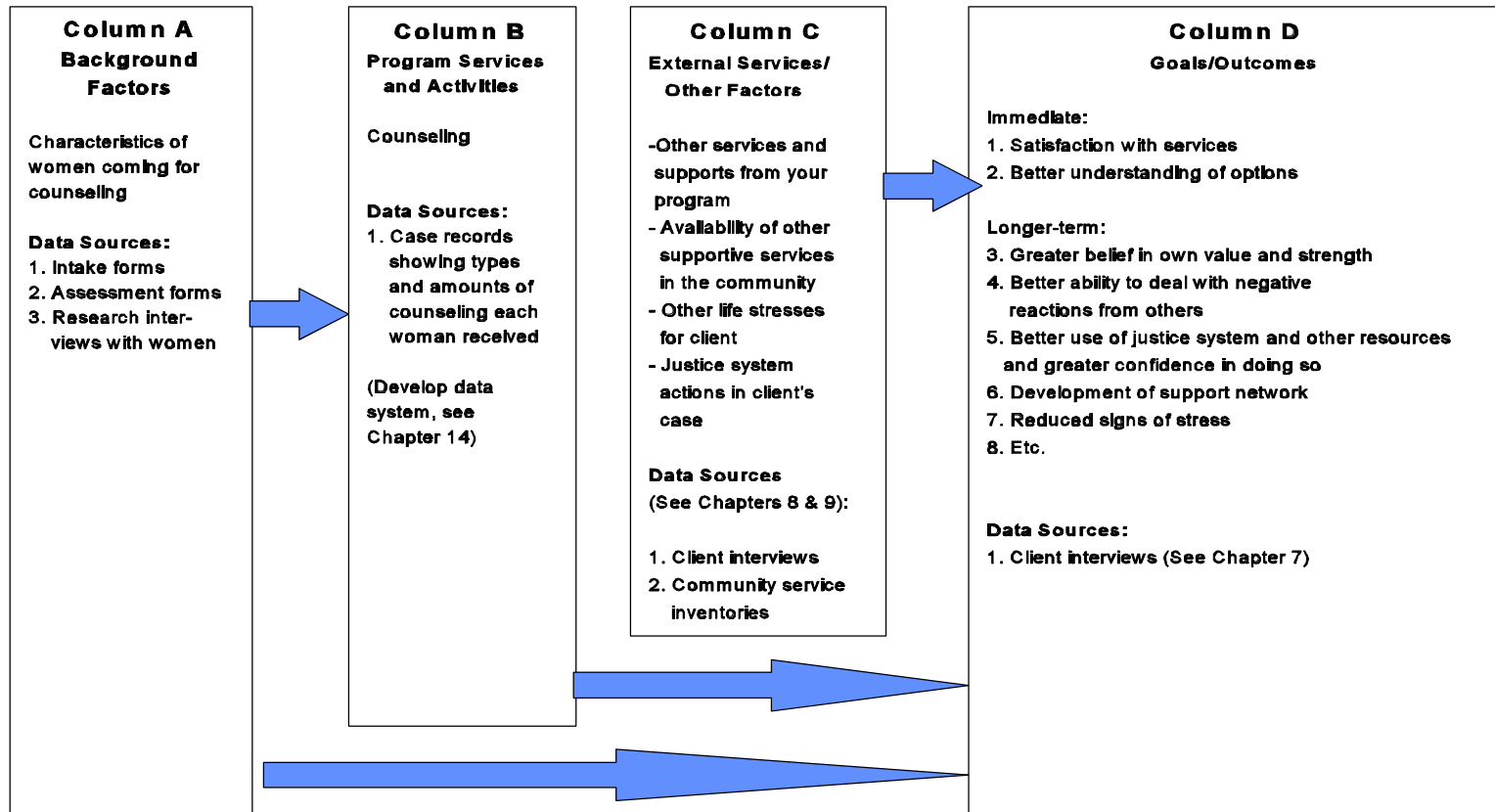


Exhibit LM.2

SPECIAL PROSECUTION UNIT

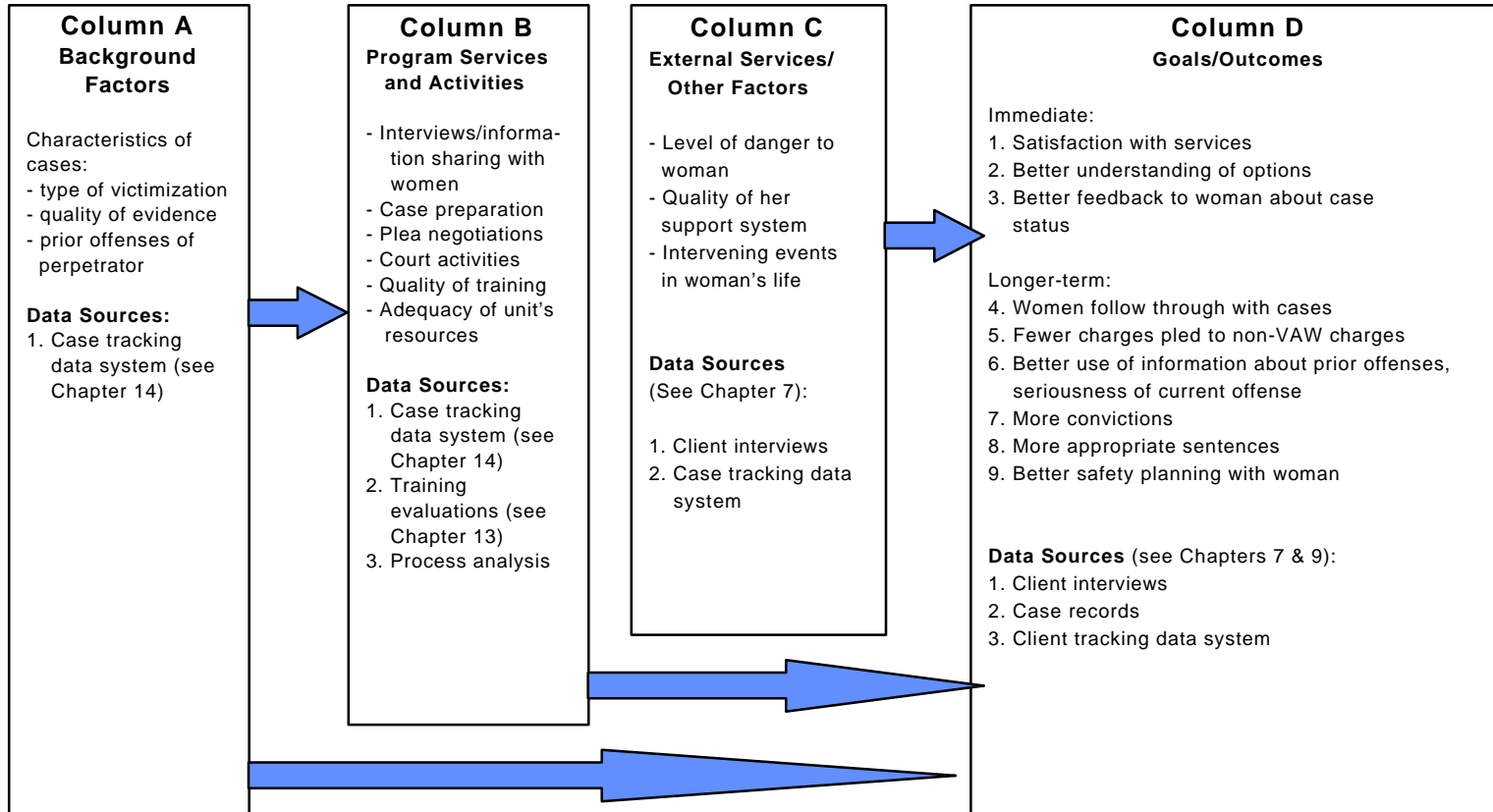
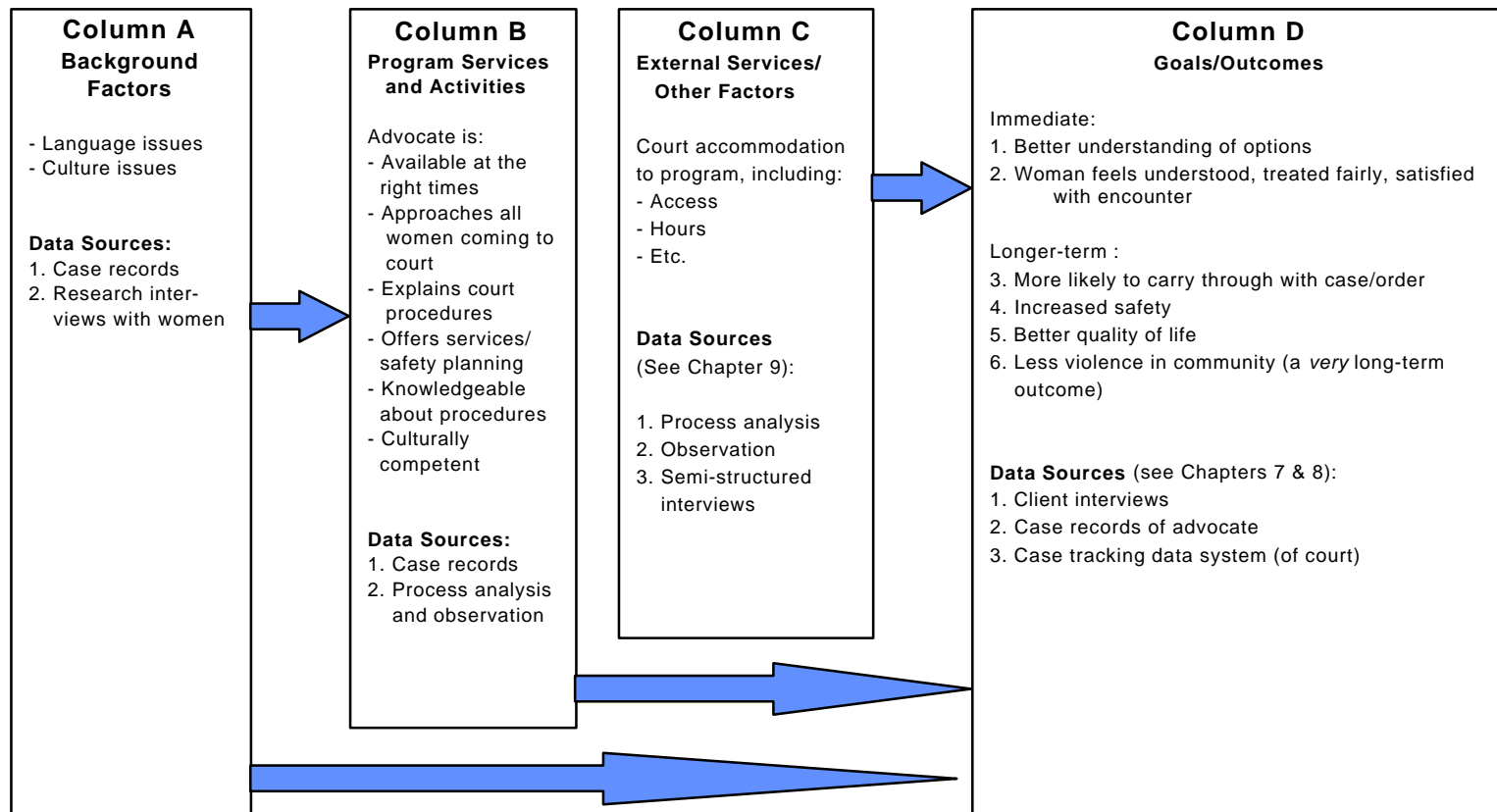


Exhibit LM.3

COURT ADVOCACY PROGRAM



CHAPTER 7

VICTIM SAFETY AND WELL-BEING: MEASURES OF SHORT-TERM AND LONG-TERM CHANGE

By Cris M. Sullivan⁴

Many STOP projects have goals that involve making changes in victims' lives. This chapter offers suggestions about ways to document such changes. It discusses the need to identify and measure changes that occur in the short run, and also changes that are more likely to take a significant period of time to develop. Once this distinction between short- and long-term impacts has been described, the chapter offers specific instruments (scales, questions, formats) that measure both short- and long-term changes.

WHAT IS SHORT-TERM AND WHAT IS LONG-TERM CHANGE?

Short-term changes are those more immediate and/or incremental outcomes one would expect to see quickly, and that will eventually lead to desired long-term changes. Optimal long-term changes might include (1) freedom from violence, (2) decreased trauma symptoms, and/or (3) increased physical, psychological, economic, and/or spiritual well-being. However, we would not expect these outcomes to occur quickly as the direct or sole result of a new or improved community-based program. Rather, effective programs would ideally result in some degree of measurable, immediate, positive change in women's lives, with this change ultimately contributing to long-term safety and well-being. For example, a hospital-based medical advocacy project for battered women might be expected to result in more women being correctly identified by the hospital, more women receiving support and information about their options, and increased sensitivity being displayed by hospital personnel in contact with abused women. Or, a SANE (sexual assault nurse examiner) program for treating sexual assault victims might be expected to produce many of these same outcomes, as well as better evidence collection. These short-term changes might then be expected to result in more women accessing whatever community resources they might need to maximize their safety (e.g., shelter, personal protection order) and/or help them cope with emotional issues (e.g., counseling, hotlines), which ultimately would be expected to lead to reduced violence and/or increased well-being (long-term outcomes). However, it would be unrealistic to expect to see a change in the level of violence in women's lives or their full psychological healing immediately or even shortly after receipt of medical

⁴ Cris M. Sullivan is Associate Professor of Community Psychology at Michigan State University, where she has conducted numerous studies on the course of women's recovery from domestic violence and the effects of services on women's outcomes. She serves on the Advisory Group for the National Evaluation of the STOP Formula Grants under the Violence Against Women Act.

advocacy offered to battered women or women who have been sexually assaulted. Rather, programs should measure the **short-term** changes they expect to impact. In these examples, that might include (1) the number of women correctly identified in the hospital as survivors of domestic abuse or sexual assault; (2) the number of women with full, complete, and secure evidence collected; (3) women's satisfaction with information and support received from the program; (4) women's increased knowledge of available resources post-intervention; (5) victims' perceptions of the effectiveness of the intervention in meeting their needs; and (6) hospital personnel's attitudes toward victims of domestic violence and sexual assault.

There are two critical points to make here:

- ❑ *Most programs using STOP grants should focus on the measurement of short-term, not long-term, change.*
- ❑ *Direct service delivery programs that are “victim”-focused should not be expected to produce decreased violence in women’s lives. Victim-based direct service programs **can** provide support, information, assistance, immediate safety for women and/or psychological counseling, but they are not designed to decrease the perpetrators’ abuse or end the risk of rape. A coordinated community response that holds perpetrators accountable for their behaviors is necessary to decrease the risk of continued abuse and sexual assault.*

Once you have decided which program outcomes you want to measure, you will need to choose or develop measuring instruments that are sensitive enough to detect whether desired changes have occurred. It is preferable to use a well-established instrument whenever possible to maximize the odds of detecting change and to increase confidence in your findings. We include in this chapter a number of instruments measuring victim safety and well-being over time, all of which are currently being used in research related to violence against women. You should not try to develop a new instrument specifically for your project unless you cannot find any existing instruments in the literature.

MEASURES OF SHORT-TERM CHANGE

In order to measure **short-term change**, answers must be provided to questions such as:

- ❑ **What specifically** did survivors receive from this program/service/intervention?
- ❑ **How much** did survivors receive from this program/service/intervention (i.e., how much time, how many units of service)?

- How effective** did survivors feel this intervention was in meeting their needs?
- How satisfied** were survivors with the various components of this intervention?
- If this program/service/intervention was designed to result in any tangible, measurable change in survivors' lives (e.g., change of residence, new financial resources), **did this change occur?**

Unlike constructs such as depression or social support, which can be measured by standardized instruments, short-term changes are generally assessed using questions you create yourself, such as:

How effective was having a court advocate in helping you understand the court process?

- 1 = very ineffective
- 2 = somewhat ineffective
- 3 = somewhat effective
- 4 = very effective

How effective was the (protection/restraining/stay away order) clinic in helping you obtain a protection/restraining/stay away order?

- 1 = very ineffective
- 2 = somewhat ineffective
- 3 = somewhat effective
- 4 = very effective

How effective was the SANE nurse in helping you understand and feel comfortable with the evidence collection process?

- 1 = very ineffective
- 2 = somewhat ineffective
- 3 = somewhat effective
- 4 = very effective

How supported did you feel by the staff of this program?

- 1 = very unsupported
- 2 = somewhat unsupported
- 3 = somewhat supported
- 4 = very supported

Evaluation Guidebook

Did the hotline counselor help you understand the feelings you were having after you were raped?

- 1 = did not help at all
- 2 = helped a little
- 3 = helped somewhat
- 4 = helped a lot

Did shelter staff help you devise a safety plan while you were here?

- 1 = yes
- 2 = no
- 0 = I did not request or want a safety plan

How satisfied are you with the services you received from this program?

- 1 = very dissatisfied
- 2 = somewhat dissatisfied
- 3 = somewhat satisfied
- 4 = very satisfied

While it is often important to ask open-ended questions, such as “What did you like?” and “What can we improve?” these types of questions should be asked in addition to, not instead of, more quantitative questions (closed-ended questions with forced options, such as those just presented). It is much easier to describe effects and to detect change with closed-ended questions, which assign a number to each answer and force the respondent to select one and only one answer. Thus, if you want to know what clients liked, ask specific questions and use specific answer categories. For example, you could ask:

How much did you like (or “how satisfied were you with,” or “how helpful was”):

- ...the promptness of our response,
- ...the specific information we provided about resources,
- ...the way people in our program interacted with you,
- ...the advice we gave you,
- ...the ways we helped you think about your situation.

For these questions, you could use answer categories such as: 1=did not like at all, 2=liked somewhat, 3=liked a lot. Or the categories could be 1=not at all satisfied/helpful, 2=somewhat satisfied/helpful, 3=very satisfied/helpful.

It is important to remember that *the wording of questions influences the responses received*. If, for example, a legal advocacy project is designed to help survivors make the best informed legal

decisions they can for themselves, based on complete and accurate information, it would be inappropriate to ask women whether they did or did not participate in pressing charges against their assailants to determine program outcome, as pressing charges may not be the best option for all women. Rather, questions might be used such as the following:

What decision did you make regarding participating in pressing charges against (perpetrator's name)?

- 1 = yes, will participate in pressing charges
- 2 = undecided
- 3 = no, will not participate in pressing charges

What factors influenced your decision?	YES	NO
prior experience with the criminal justice system	1	2
information received from legal advocate	1	2
information received from prosecutor	1	2
fear of perpetrator (etc., etc.)	1	2
other (specify: _____)	1	2

Do you think this decision will result in (perpetrator's name) being:

- 1 = less violent against you in the future
- 2 = no change in violence
- 3 = more violent against you in the future
- 7 = don't know

Did you find our legal advocacy program to be:

- 1 = not at all helpful in making this decision
- 2 = somewhat helpful in making this decision
- 3 = very helpful in making this decision

Again, short-term change is generally measured by examining *what* the consumer received, *how much* she received, *how effective* she found the service, *how satisfied* she was with the service, and whether *short-term, incremental change* occurred.

MEASURES OF LONG-TERM CHANGE: VICTIM SAFETY AND WELL-BEING

Although the majority of programs receiving STOP grant funding will not be in a position to evaluate longer-term outcomes, it is sometimes feasible and appropriate to measure whether victims' level of safety and/or quality of life improves over time as a result of an intervention. This level of evaluation generally requires additional time and financial resources, but when such an effort is warranted, there are a number of standardized instruments available from which to

choose. The remainder of this chapter presents brief critiques of various measures previously used in research pertaining to violence against women. The first section describes eight instruments developed to measure physical, psychological, and/or sexual abuse. These measures can be used to examine whether such violence increases, decreases, or remains the same over time. The second section pertains to measures of well-being, including depression, post-traumatic stress, overall quality of life, and self-esteem. The last section describes instruments that measure correlates of well-being—access to community resources, stressful life events, and level of social support.

Criteria Used to Select Instruments

Numerous instruments have been developed that measure some aspect of safety and/or well-being, and not all could be included in this *Guidebook*. The instruments described in this chapter were chosen because they met the following criteria:

- (1) easily administered by laypersons;
- (2) relatively few items (less than 30 wherever possible);
- (3) history of use in research on violence against women;
- (4) acceptable reliability and validity; and
- (5) easily interpreted by laypersons.

Each scale presented has demonstrated at least some degree of adequate reliability and validity as a field instrument; therefore we do not detail the properties of each instrument unless they are noted as a concern.⁵ Sample items follow each scale description, as do references to articles where you can find detailed information about how each instrument behaves. Complete instruments can be ordered from the STOP TA Project (800 256-6883 or 202 265-0967 in the Washington, D.C. area) unless they are under copyright. Copyrighted instruments that must be obtained by the publisher or author are noted.⁶

⁵ *Reliability* refers to certain characteristics of a scale related to *sameness*: that all the items correlate highly with each other; and that if the scale is used at two different points in time and nothing has changed between those times, the scale will produce approximately the same score both times. *Validity* refers to a number of characteristics of a scale related to *correctness*: that all the items on the scale seem on their face to be measuring the concept of interest; that each item on Scale A “likes” the other items on Scale A better than it “likes” the items that comprise Scale B, C, etc.; and that the scale scores are strongly associated with concepts you would expect them to be associated with, and not associated with things you would not expect them to be related to.

⁶ Unquestionably, additional scales measuring safety and well-being and meeting the above criteria either have been overlooked by the author or were excluded due to space limitations. The inclusion or exclusion of an instrument in this chapter does not necessarily reflect an endorsement or lack thereof on the

A Note of Caution

The majority of the instruments that follow were developed for research purposes. Their strength lies in their ability to characterize *groups* of people, and they should not be used by laypersons as individualized assessment instruments or as diagnostic tools for individual clients. For example, although the *Sexual Experiences Survey* classifies respondents into one of four categories (nonvictimized, sexually coerced, sexually abused, or sexually assaulted), such classifications are made based on large aggregated data sets. It would be problematic and unethical to inform a particular woman of her “classification,” based on her responses to this survey. A woman who did not self-identify in the same manner as her classification on this measure could lose faith in the intervention program designed to assist her and/or could feel misunderstood or even revictimized. The following instruments should be used for descriptive purposes of the sample as a whole and to examine group, not individual, differences.

Measures of Victim Safety

The following measures of physical, psychological, and/or sexual violence are presented alphabetically in this section.

Measures of Physical Abuse by Intimates

Abusive Behavior Inventory [also measures psychological abuse] (Shepard & Campbell, 1992)
Conflict Tactics Scales (Revised) (Straus et al., 1996)
Danger Assessment (Campbell, 1986)
Index of Spouse Abuse [also measures psychological abuse] (Hudson & McIntosh, 1981)
Severity of Violence Against Women Scales (Marshall, 1992)

Measures of Psychological Abuse

Index of Psychological Abuse (Sullivan, Parisian, & Davidson, 1991)
Psychological Maltreatment of Women Inventory (Tolman, 1989)

Measure of Sexual Violence

Sexual Experiences Survey (Koss & Oros, 1982)

Intimate Physical Abuse

ABUSIVE BEHAVIOR INVENTORY - PARTNER FORM

Citation Shepard, M.F., & Campbell, J.A. (1992). The Abusive Behavior Inventory: A measure of psychological and physical abuse. *Journal of Interpersonal Violence*, 7(3), 291-305. Copyright © 1992 by Sage Publications, Inc. Reprinted by Permission of Sage Publications, Inc.

Description Drawing from both feminist theory and educational curriculum with batterers, the authors designed this 29-item measure of psychological (18 items) and physical (11 items) types of abuse. A self-report measure, it is simple to administer and would generally take no more than 5 minutes to complete. The scale's strengths are that it was found to successfully differentiate between abusers and non-abusers, and it was designed to tap power and control issues within the relationship (for example: "checked up on you," and "stopped you or tried to stop you from going to work or school"). Its weaknesses, conceded by the authors, are that (1) there is no attention to injuries or to medical attention needed, which could approximate severity of the violence; and (2) the reliability and validity of the measure were based on a sample of inpatient, chemically-dependent men and women.

Sample items:

Copyright restrictions prohibit electronic distribution of scale contents.

Reference Petrik, N.D. (1994). The reduction of male abusiveness as a result of treatment: Reality or myth? *Journal of Family Violence*, 9(4), 307-316.

Intimate Physical Abuse

THE REVISED CONFLICT TACTICS SCALES (CTS2)

Citation Straus, M.A., Hamby, S.L., Boney-McCoy, S., & Sugarman, D.B. (1996). The Revised Conflict Tactics Scales (CTS2): Development and preliminary psychometric data. *Journal of Family Issues*, 17, 283-316. Copyright © 1996 by Sage Publications, Inc. Reprinted by Permission of Sage Publications, Inc.

Description The first instrument designed to measure conflict and violence between intimate partners was the original version of this scale (CTS1). The CTS1 has been both widely used and widely criticized. It was an 18-item scale of relationship conflict tactics, with the latter 10 items measuring violent strategies. Its strengths were that it was used successfully in many settings and with many populations, and that it was short and easy to administer. Its weaknesses included (1) measuring partner violence only within the context of conflict, while domestic violence is about power and control, (2) ignoring many common types of woman-battering, including symbolic gestures of violence as well as tactics of power, control, and intimidation, (3) rating some acts of violence as more severe than others outside of the context of the event (i.e., slapping is rated as “mild,” although a hard slap can cause severe injury), (4) ignoring whether an act was committed in self-defense, and (5) excluding injuries sustained or medical attention needed to approximate severity of the violence.

The revised CTS2 has added more items to include some additional types of conflict tactics, has added a section on injuries, and now has 39 items. The other weaknesses remain even with the CTS2. Both the CTS and CTS2 are still very good instruments to use because a great deal is known about their strengths and weaknesses. However, you may want to compensate for remaining weaknesses with additional questions. There is no standardized instrument in general use to help you in making these compensations, so this is one place where you may have to make up some questions of your own to cover the issues that the CTS2 omits.

Sample items:

Copyright restrictions prohibit electronic distribution of scale contents.

Reference Straus., M.A. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics (CT) Scales. *Journal of Marriage and the Family*, 75-88.

Evaluation Guidebook

Straus, M.A., & Gelles, R.J. (1986). Societal change and change in family violence: Violence from 1975 to 1985 as revealed by two national surveys. *Journal of Marriage and the Family*, 48, 465-479.

Intimate Physical Abuse

THE DANGER ASSESSMENT

- Citation** Campbell, J.C. (1986). Nursing assessment for risk of homicide with battered women. *Advances in Nursing Science*, 8, 36-51. Copyright © 1981 by the National Council on Family Relations, 3989 Central Ave., NE, Suite 550, Minneapolis, MN55421. Used by permission. The instrument is available from Jacquelyn Campbell on request.
- Description** This 11-item assessment tool was created to assist women with abusive partners in assessing their danger of homicide. The author recommends this tool be used as part of a nursing assessment of domestic violence, and that nurses and patients complete the tool together. This instrument was created with the input of battered women, shelter workers, law enforcement officials, and other experts on battering. Due to the singularities of each woman's situation, however, the author stresses that *no actual prediction of lethality* be made based upon a woman's score. The score (summed affirmative responses) should be shared with the woman who has completed the Danger Assessment so she can determine her own risk.

Sample items:

Response categories:

0=no
1=yes

1. Has the physical violence increased in frequency over the past year?
4. Is there a gun in the house?
7. Does he threaten to kill you and/or do you believe he is capable of killing you?

- References** McFarlane, J., & Parker, B. (1994). Preventing abuse during pregnancy: An assessment and intervention protocol. *American Journal of Maternal/Child Nursing*, 19, 321-324.
- McFarlane, J., Parker, B., & Soeken, J. (1995). Abuse during pregnancy: Frequency, severity, perpetrator and risk factors of homicide. *Public Health Nursing*, 12(5), 284-289.

Intimate Physical Abuse

INDEX OF SPOUSE ABUSE (ISA)

Citation Hudson, W.W., & McIntosh, S.R. (1981). The assessment of spouse abuse: Two quantifiable dimensions. *Journal of Marriage and the Family*, 43, 873-888.

Description This 30-item self-report instrument takes about 5 minutes to administer, and measures both physical (15 items) and non-physical (15 items) types of intimate abuse. The primary drawback of this scale is related to one of its strengths. Because the authors note that some types of violence are more severe than other types, they have assigned weights to the items, which results in slightly more complexity in computing scale scores for respondents. However, Hudson & McIntosh (1981) describe, in simple terms, the computation required to obtain the two scores.

The 15 items measuring non-physical abuse are quite inclusive of numerous psychologically and emotionally abusive behaviors. The “physical abuse” items, however, do not include such behaviors as kicking, restraining, burning, choking, pushing, or shoving. Women who have experienced these types of abuse without experiencing punches would receive artificially minimized scores on this measure. Further, 7 of the 15 “physical abuse” items only *imply* physical abuse. An example of such an item is: ‘My partner becomes abusive when he drinks.’ An additional drawback is that two items refer to abuse occurring when the perpetrator has been drinking, which could also artificially minimize the scores of women whose abusers do not drink. To compensate for aspects of domestic violence that this scale omits, you might want to make up some questions of your own.

Sample items:

Response categories:

1=never

2=rarely

3=occasionally

4=frequently

5=very frequently

[physical abuse]

7. My partner punches me with his fists.

[non-physical abuse]

1. My partner belittles me.

References Campbell, D.W., Campbell, J., King, C., Parker, B., & Ryan, J. (1994). The reliability and factor structure of the Index of Spouse Abuse with African-American women. *Violence and Victims*, 9(3), 259-274.

Intimate Physical Abuse

SEVERITY OF VIOLENCE AGAINST WOMEN SCALES

Citation Marshall, L.L. (1992). Development of the Severity of Violence Against Women Scales. *Journal of Family Violence*, 7(2), 103-121. Copyright © 1992 by Plenum Publishing Corp. Used with permission.

Description This 46-item instrument was specifically created to measure “threatened, attempted, and completed behaviors likely to cause injury or pain” (Marshall, 1992: 105). The nine dimensions of violence measured by this scale are: symbolic violence; threats of mild, moderate, and serious violence; acts of mild, minor, moderate, and severe violence; and sexual violence. A strength of this scale is that it captures symbolic violence—behaviors often used by perpetrators to frighten and intimidate women. Weaknesses of this instrument are its weighting system, and therefore the categories of threats and violence that it produces. Items were weighted for severity based on ratings provided by samples of women who were not necessarily abused themselves. For example, the item “held her down, pinning her in place” was rated as “mild,” and the item “bit her” was rated as “minor.” This illustrates the difficulty of rating behaviors out of context, as these acts can of course also be very serious. However, the 46 empirically derived items are excellent examples of intimate male violence against women and can be used without the weighting system discussed in Marshall (1992).

Sample items:

Response categories (referring to the prior 12 months):	[symbolic violence]
1=never	1. Hit or kicked a wall, door or furniture.
2=once	[threats]
3=a few times	7. Shook a fist at you.
4=many times	[physical violence]
	35. Choked you.
	[sexual violence]
	41. Demanded sex whether you wanted to or not.

References Vitanza, S., Vogel, L.C.M., & Marshall, L.L. (1995). Distress and symptoms of posttraumatic stress disorder in abused women. *Violence and Victims*, 10(1), 23-34.

Psychological Abuse

INDEX OF PSYCHOLOGICAL ABUSE

Citation Sullivan, C.M., Parisian, J.A., & Davidson, W.S. (1991). Index of Psychological Abuse: Development of a measure. Presented at the 99th annual convention of the American Psychological Association, San Francisco, CA.

Description This 33-item instrument was designed to measure the common types of psychological abuse reported by battered women: criticism, ridicule, isolation, withdrawal, and control. It is easy to administer and generally takes less than 5 minutes to complete. This measure was originally developed on two samples: women who were exiting a domestic violence shelter, and dating college students. It has since been validated with a Korean sample of abused and non-abused women (Morash et al., in preparation).

Sample items:

Response categories:

- | | |
|---|---|
| 1=never | 2. Accused you of having or wanting other sexual relationship(s). |
| 2=rarely | 7. Tried to control your activities. |
| 3=sometimes | 20. Criticized your intelligence. |
| 4=often | |
| 8=not applicable (i.e., no children, no pets) | |

References Morash, M., Hoffman, V., Lee, Y.H., & Shim, Y.H. (in preparation). Wife abuse in South Korea.

Sullivan, C.M., Tan, C., Basta, J., Rumptz, M., & Davidson W.S. (1992). An advocacy intervention program for women with abusive partners: Initial evaluation. *American Journal of Community Psychology*, 20(3), 309-332.

Tan, C., Basta, J., Sullivan, C., & Davidson, W.S. (1995). The role of social support in the lives of women exiting domestic violence shelters. *Journal of Interpersonal Violence*, 10(4), 437-451.

Psychological Abuse

THE PSYCHOLOGICAL MALTREATMENT OF WOMEN INVENTORY

- Citation** Tolman, R.M. (1989). The development of a measure of psychological maltreatment of women by their male partners. *Violence and Victims*, 4(3), 159-178. Copyright © 1994 by Springer Publishing Company, Inc. Used by permission.
- Description** This 58-item scale was derived from clinical observations, the clinical literature, and nonphysical abuse items from the Conflict Tactics Scales and the Index of Spouse Abuse (described under “Measures of Physical Abuse” in this chapter). This self-report measure, which was piloted on batterers as well as battered women, takes 10-15 minutes to administer. The author notes, however, that batterers underestimate their own abusiveness and cautions against using this scale as a pre-intervention evaluation tool for batterer intervention programs.
- Sample items:**
- | | |
|----------------------|--|
| Response categories: | 2. My partner insulted me or shamed me in front of others. |
| 1=never | 10. My partner called me names. |
| 2=rarely | 42. My partner restricted my use of the phone. |
| 3=sometimes | |
| 4=frequently | |
| 5=very frequently | |
- References** Dutton, D.G. (1995). A scale for measuring propensity for abusiveness. *Journal of Family Violence*, 10(2), 203-221.
- Kasian, M. (1992). Frequency and severity of psychological abuse in a dating population. *Journal of Interpersonal Violence*, 7(3), 350-364.

Sexual Abuse

THE SEXUAL EXPERIENCES SURVEY

Citation Koss, M.P., & Oros, C.J. (1982). The Sexual Experiences Survey: A research instrument investigating sexual aggression and victimization. *Journal of Consulting and Clinical Psychology*, 50(3), 455-457. Copyright © 1982 by the American Psychological Association. Reprinted with permission.

Description This 13-item instrument, which can be asked of victims or slightly reworded to be administered to perpetrators, is the only measure currently used to detect victims of rape as well as unidentified offenders in the general population. Women can be classified into one of the following four categories based on their responses to this survey: nonvictimized, sexually coerced, sexually abused, and sexually assaulted. “Sexually assaulted” includes having experienced those acts that would legally be classified as criminal sexual assault, including gross sexual imposition and attempted rape. The authors note that rape survivors responded consistently to this instrument whether it was administered privately or by an interviewer. This is important, given the reticence many women have in admitting to having been sexually victimized.

Sample items:

Response categories:

1=yes

2=no

4. Have you ever had sexual intercourse with a man even though you didn't really want to because he threatened to end your relationship otherwise?
8. Have you ever been in a situation where a man tried to get sexual intercourse with you when you didn't want to by threatening to use physical force (twisting your arm, holding you down, etc.) if you didn't cooperate, but for various reasons sexual intercourse did not occur?
12. Have you ever been in a situation where a man obtained sexual acts with you such as anal or oral intercourse when you didn't want to by using threats or physical force (twisting your arm, holding you down, etc.)?

References Koss, M.P., & Gidycz, C.A. (1985). Sexual Experiences Survey: Reliability and validity. *Journal of Consulting and Clinical Psychology*, 53(3), 422-423.

Koss, M.P. (1985). The hidden rape victim: Personality, attitudinal, and situational characteristics. *Psychology of Women Quarterly*, 9(2), 193-212.

Measures of Psychological Well-Being

In addition to examining level of violence as a long-term outcome, it is sometimes useful to examine the effects of a program on a woman's overall well-being. Prior research in the area of violence against women has often focused on the following constructs when examining well-being: depression, post-traumatic stress symptoms, perceived quality of life, and self-esteem. It is important to remember, when measuring the level of distress in women's lives, that *normative distress does not equal psychological dysfunction*. It is typically "normal" and adaptive to be distressed after a physical or sexual assault. Therefore, the following instruments should not be used by laypersons for diagnostic purposes, and high scores on distress scales should be cautiously interpreted.

Depression

Beck Depression Inventory (Beck, Rush, Shaw, & Emery, 1979)
CES-D (Radloff, 1977)

Post-Traumatic Stress Disorder

Posttraumatic Stress Scale for Family Violence (Saunders, 1994)
Trauma Symptom Checklist (Briere & Runtz, 1989)

Quality of Life

Quality of Life Scale (Sullivan, Tan, Basta, Rumptz, & Davidson, 1992)
Satisfaction with Life Scale (Pavot & Diener, 1993)

Self-Esteem

Coopersmith Self-Esteem Inventory (Coopersmith, 1967)
Rosenberg Self-Esteem Inventory (Rosenberg, 1965)

Growth Outcomes and Coping Strategies

"How I See Myself Now" (Burt & Katz, 1987)
"Changes that Have Come from Your Efforts to Recover" (Burt & Katz, 1987)
"How I Deal With Things" (Burt & Katz, 1987, 1988)

Depression

BECK DEPRESSION INVENTORY

Citation

Description This 21-item inventory is a clinically derived measure originally designed for use with psychiatric populations. The measure can be administered verbally or in writing, and takes approximately 10-15 minutes to complete. Individuals endorse one of four choices per item (on a 0-3 scale), to receive a final summed score between 0 and 63. A score of 0-9 indicates no depression, while a score of 10-15 indicates mild depression and a score of 16-23 suggests moderate depression. Scores above 23 indicate severe clinical depression. This is a popular inventory that has been used across numerous clinical and non-clinical samples, including college samples, women with abusive partners, and survivors of rape. It is relatively simple to administer and is easily interpretable. However, it requires more deliberation and attention to detail from study participants than does the CES-D (described on next page). Within each item, the respondent must decide which of four statements most accurately describes her or his feelings, a task requiring more concentration and reflection than may be necessary, depending on the study.

Sample items:

1. I do not feel sad.
 I feel sad.
 I am sad all the time and I can't snap out of it.
 I am so sad or unhappy that I can't stand it.
2. I am not particularly discouraged about the future.
 I feel discouraged about the future.
 I feel I have nothing to look forward to.
 I feel that the future is hopeless and that things cannot improve.

References

Beck, A.T., Rush, A.J., Shaw, B.F., & Emery, G. (1979). *Cognitive Therapy for Depression*. New York: Guilford.

Robinson, B.E. (1996). Concurrent validity of the Beck Depression Inventory as a measure of depression. *Psychological Reports*, 79(3), 929-930.

Depression

CENTER FOR EPIDEMIOLOGICAL STUDIES-DEPRESSION (CES-D) SCALE

Citation Radloff, L.S. (1977). The CES-D Scale: A self report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385-401. Copyright © 1977 by West Publishing Company/Applied Psychological Measurement, Inc. Used with permission.

Description This is a 20-item self-report checklist of depressive symptoms found within the general population. It can be administered verbally or in written format, and takes 5-10 minutes to complete. Individuals rate the degree to which they have been bothered by each symptom, within the prior week, on a 0 (*rarely or never*) to 3 (*most or all the time*) scale. For determining level of depression, scores are summed (*four are reverse coded*). A score of 0 - 15.5 indicates no depression, a score of 16-20.5 indicates mild depression, and scores 31 and higher indicate severe depression. This checklist was originally validated on a nonclinical, general population sample. It has since been administered to numerous study samples, including women with abusive partners and incarcerated women. It is simple to administer and easily understood by study participants. Having been validated for research in the general population, it is appropriate to use for community samples of women who are survivors of violence.

Sample items:

Response categories (referring to the past week):

- 0=rarely or never
- 1=some or a little
- 2=occasionally
- 3=most or all the time

1. I was bothered by things that usually don't bother me.
2. I did not feel like eating; my appetite was poor.

References Lin, N., Dean, A., & Ensel, W. (1986). *Social Support, Life Events, and Depression*. FL: Academic Press.

Martin, S.L., Cotten, N.U., Browne, D.C., Kurz, B., & Robertson, E. (1995). Family violence and depressive symptomatology among incarcerated women. *Journal of Family Violence, 10*(4), 399-412.

Post-Traumatic Stress

POSTTRAUMATIC STRESS SCALE FOR FAMILY VIOLENCE

Citation Saunders, D.G. (1994). Posttraumatic stress symptom profiles of battered women: A comparison of survivors in two settings. *Violence and Victims, 9*(1), 31-44. Copyright © 1994 by Springer Publishing Company, Inc. Used by permission.

Description This 17-item scale measures level of posttraumatic stress symptoms reported by women no longer experiencing intimate partner violence. It would be important to include a number of additional items if this scale were used in a research study: length and severity of abuse, length of time between cessation of abuse and presence of symptomatology, and length of time PTSD symptoms were/have been experienced. Also, although the author preferred to use the number of times a problem was experienced instead of the more subjective responses of “never,” “rarely,” “sometimes,” and “often,” there are a number of concerns with this strategy. First, it can be difficult if not impossible to remember an exact number of times one experiences a troubling feeling or thought. Second, the response ranges are not equivalent (i.e., 1-2 times vs 51-100 times), and the author created scale scores by summing the midpoints between each response category. This atypical scaling strategy was not well justified. However, the scale does seem to capture symptoms of posttraumatic stress in women who have experienced intimate abuse.

Sample items:

As a result of any of your partner’s verbal or physical abuse of you, please circle how many times you had each of the following problems:

- never
- 1-2
- 3-11
- 12-24
- 25-36
- 37-50
- 51-100
- over 100 times

- 1. Unpleasant memories of the abuse you can’t keep out of your mind.
- 5. Trying to avoid thoughts or feelings associated with the abuse.
- 14. Difficulty concentrating.

References Woods, S.J., & Campbell, J.C. (1993). Posttraumatic stress in battered women: Does the diagnosis fit? *Issues in Mental Health Nursing, 14*, 173-186.

Post-Traumatic Stress

THE TRAUMA SYMPTOM CHECKLIST

Citation Briere, J., & Runtz, M. (1989). The Trauma Symptom Checklist (TSC-33): Early data on a new scale. *Journal of Interpersonal Violence*, 4(2), 151-163. Copyright © 1989 by Sage Publications, Inc. Reprinted by Permission of Sage Publications, Inc.

Description This 33-item scale was designed to examine the impact of childhood abuse on later adult functioning. It contains five subscales: dissociation [6 items], anxiety [9 items], depression [9 items], hypothesized post-sexual-abuse-trauma [6 items], and sleep disturbances [4 items] (1 item appears in two different subscales). The instrument can be administered verbally or in written form, and it has been found to discriminate between women who were and were not sexually abused as children. However, the authors caution against using this measure as a “litmus test” for the presence of childhood sexual victimization, as symptomatology varies across individuals.

Sample items:

Copyright restrictions prohibit electronic distribution of scale contents.

References Briere, J., & Runtz, M. (1987). Post-sexual abuse trauma: Data and implications for clinical practice. *Journal of Interpersonal Violence*, 2, 367-379.

Briere, J., & Runtz, M. (1988). Symptomatology associated with childhood sexual victimization in a non-clinical adult sample. *Child Abuse and Neglect*, 12, 51-59.

Quality of Life

QUALITY OF LIFE SCALE

Citation Sullivan, C.M., Tan, C., Basta, J., Rumpitz, M., & Davidson W.S. (1992). An advocacy intervention program for women with abusive partners: Initial evaluation. *American Journal of Community Psychology*, 20(3), 309-332. Copyright © 1992 by Plenum Publishing Corp. Used with permission.

Description This scale first contained 25 items, modified from Quality of Life scales developed by Andrews & Withey (1976), designed to measure how satisfied respondents were with their overall quality of life. This instrument was pilot tested on women exiting a domestic violence shelter program, and was re-administered every six months over a two year period. Items were so highly intercorrelated that the top 9 items were then chosen for a shorter, final scale. The measure is easy to administer and sensitive to detecting change over time.

Sample items:

Response categories:

- 1=extremely pleased
- 2=pleased
- 3=mostly satisfied
- 4=mixed (equally satisfied and dissatisfied)
- 5=mostly dissatisfied
- 6=unhappy
- 7=terrible

- 1. First, a very general question. How do you feel about your life in general?
- 4. How do you feel about the amount of fun and enjoyment you have?
- 9. How do you feel about the way you spend your spare time?

References Andrews, F., & Withey, S. (1976). *Social Indicators of Well-being: American's Perceptions of Life Quality*. New York: Plenum Press.

Atkinson, T. (1982). The stability and validity of quality of life measures. *Social Indicators Research*, 10, 113-132.

Sullivan, C.M., Campbell, R., Angelique, H., Eby, K.K., & Davidson, W.S. (1994). An advocacy intervention program for women with abusive partners: Six month followup. *American Journal of Community Psychology*, 22(1), 101-122.

Quality of Life

SATISFACTION WITH LIFE SCALE

Citation Pavot, W., & Diener, E. (1993). Review of the Satisfaction With Life Scale. *Psychological Assessment, 5*(2), 164-172. Copyright © 1993 by the American Psychological Association. Reprinted with permission.

Description The 5-item Satisfaction With Life Scale assesses the positive aspects of people's lives. It can detect change over time, and has been translated into French, Russian, Korean, Hebrew, Mandarin Chinese, and Dutch. Because it is short, the entire scale is reproduced below. No permission is needed to use this instrument.

Items:

Response categories:

- 1=strongly disagree
- 2=disagree
- 3=slightly disagree
- 4=neither agree nor disagree
- 5=slightly agree
- 6=agree
- 7=strongly agree

1. In most ways my life is close to my ideal.
2. The conditions of my life are excellent.
3. I am satisfied with my life.
4. So far I have gotten the important things I want in life.
5. If I could live my life over, I would change almost nothing.

References Diener, E., Emmons, R.A., Larsen, R.J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment, 49*, 71-75.

Diener, E., Sandvik, E., Pavot, W., & Gallagher, D. (1991). Response artifacts in the measurement of well-being. *Social Indicators Research, 24*, 35-56.

Larsen, R.J., Diener, E., & Emmons, R.A. (1985). An evaluation of subjective well-being measures. *Social Indicators Research, 17*, 1-18.

Pavot, W., Diener, E., Colvin, R., & Sandvik, E. (1991). Further validation of the Satisfaction With Life Scale: Evidence for the cross-method convergence of well-being measures. *Journal of Personality Assessment, 57*(1), 149-161.

Self-Esteem

COOPERSMITH SELF-ESTEEM INVENTORY

Citation Coopersmith, S. (1967). *The Antecedents of Self-esteem*. San Francisco: W.H. Freeman & Co. From: THE ANTECEDENTS OF SELF-ESTEEM by Coopersmith © by W.H. Freeman and Company. Used with permission.

Description This 25-item self-administered form takes approximately 10 minutes to complete. It was originally tested on school children but has since been administered to adults across numerous settings as well. Participants indicate whether each statement is “like me” or “unlike me.” Raw scores, which range from 0 to 25, are multiplied by 4, for a final score ranging from 0 to 100. Higher scores indicate higher self-esteem.

Sample items:

Items are answered either “like me” or “unlike me.” The high-esteem response is indicated in parentheses after each item.

1. I often wish I were someone else. (unlike me)
5. I’m a lot of fun to be with. (like me)

References Getzinger, S., Kunce, J., Miller, D., & Weinberg, S. (1972). Self-esteem measures and cultural disadvantage. *Journal of Consulting and Clinical Psychology*, 38, 149.

Silvern, L., Karyl, J., Waelde, L., Hodges, W.F., Starek, J., Heidt, E., & Min, K. (1995). Retrospective reports of parental partner abuse: Relationships to depression, trauma symptoms and self-esteem among college students. *Journal of Family Violence*, 10(2), 177-202.

Ziller, R., Hagey, J., Smith, M.D., & Long, B. (1969). Self-esteem: A self-social construct. *Journal of Consulting and Clinical Psychology*, 33, 84-95.

Self-Esteem

ROSENBERG SELF-ESTEEM INVENTORY

Citation Rosenberg, (1965). Self-Esteem Scale. Rights reserved by Princeton University Press. Used with permission.

Description This 10-item scale is easy to administer verbally or in written form. It has been used extensively in the past with varied populations, including women with abusive partners and rape survivors.

Sample items:

Response categories:

- 1=Strongly agree
- 2=Agree
- 3=Disagree
- 4=Strongly disagree

- 1. I feel that I am a person of worth, at least on an equal basis with others.
- 2. I feel that I have a number of good qualities.

References Fleming, J.E., & Courtney, B.E. (1984). The dimensionality of self-esteem: II. Hierarchical facet model for revised measurement scales. *Journal of Personality and Social Psychology*, 46, 404-421.

Myers, M.B., Templer, D.I., & Brown, R. (1984). Coping ability of women who become victims of rape. *Journal of Consulting and Clinical Psychology*, 52, 73-78.

Saunders, D.G. (1994). Posttraumatic stress symptom profiles of battered women: A comparison of survivors in two settings. *Violence and Victims*, 9(1), 31-44.

Growth Outcomes and Coping Strategies

“HOW I SEE MYSELF NOW”

Citation Burt, M.R., & Katz, B.L. (1987). Dimensions of recovery from rape: Focus on growth outcomes. *Journal of Interpersonal Violence*, 2 (1), 57-81. Copyright © 1987 by Sage Publications, Inc. Reprinted by Permission of Sage Publications, Inc.

Description This 43-item self-administered instrument measures various components of self-concept. It focuses on dimensions of a woman’s life that are challenged or “shaken up” by a rape. It has six subscales: angry/needy/lonely; independence/competence; mental health; trust; help; and guilt/blame. A respondent rates the degree to which she feels that each adjective describes her. The scale also deals with the respondent’s satisfaction with her status on each dimension. This measure was one of three growth outcome instruments created by Burt and Katz and used in their 1987 study as part of a questionnaire given to adult rape victims 1 to 14 years after their rape.

Response categories (referring to the present time)

1=almost never	u=unsatisfied
2=rarely	m=mixed feelings
3=sometimes	s=satisfied
4=half the time	
5=often	
6=usually	
7=almost always	

Sample items: guilty willing to ask for help
independent doing well
self-respecting lonely
competent deserving blame

References Katz, B.L. & Burt, M.R. (1988). Self-Blame in recovery from rape: Help or hindrance? Chapter 10 in Ann W. Burgess, (ed.), *Rape and Sexual Assault, Vol. II*. New York: Garland Press.

Janoff-Bulman, R., & Frieze, I.H. (1983). A theoretical perspective for understanding reactions to victimization. *Journal of Social Issues*, 39 (2), 195-221.

Veronen, L., & Kilpatrick, D. (1983). Rape: A precursor to change. In E. Callahan & K. McCluskey (Eds.), *Life-span Development Psychology: Nonnormative Life Events*. New York: Academic Press.

Growth Outcomes and Coping Strategies

“CHANGES THAT HAVE COME FROM YOUR EFFORTS TO RECOVER”

Citation Burt, M.R., & Katz, B.L. (1987). Dimensions of recovery from rape: Focus on growth outcomes. *Journal of Interpersonal Violence*, 2 (1), 57-81. Copyright © 1987 by Sage Publications, Inc. Reprinted by Permission of Sage Publications, Inc.

Description This 28-item measure assesses respondents' perceptions of how they are now compared to how they were before being raped. It draws its items from anecdotal accounts of ways that women say they have become stronger as a result of coping with and trying to recover from a rape. The resulting scale yields three subscales: self-value, positive actions, and interpersonal skills. Each has acceptable reliability and important elements of construct validity. Most women interviewed 2 to 14 years postrape could report several ways in which they liked themselves better or saw themselves as having grown stronger during the process of recovery.

Sample items:

Response categories (referring to the present time):

- 1=much less than before the rape
- 2=somewhat less
- 3=a little less
- 4=the same as before the rape
- 5=a little more
- 6=somewhat more
- 7=much more than before the rape

- 2. I am able to talk with my family about all types of important things.
- 8. I'm confident in my judgments of people.
- 16. I value myself.
- 18. I can handle people who try to boss or control me.
- 23. I feel like I'm in control of my life.

References Veronen, L., & Kilpatrick, D. (1983). Rape: A precursor to change. In E. Callahan & K. McCluskey (Eds.), *Life-Span Development Psychology: Nonnormative Life Events*. New York: Academic Press.

Davis, R.C., Kennedy, J. & Taylor, B.G. (1997). Positive changes following sexual assault. New York: Victim Services Agency.

Growth Outcomes and Coping Strategies

“HOW I DEAL WITH THINGS”

Citation Burt, M.R., & Katz, B.L. (1987). Dimensions of recovery from rape: Focus on growth outcomes. *Journal of Interpersonal Violence*, 2 (1), 57-81. Copyright © 1987 by Sage Publications, Inc. Reprinted by Permission of Sage Publications, Inc.

Description This self-report instrument contains 33 items describing different behaviors often used when coping with a rape (or with other traumatic events). It yields five subscales measuring different coping tactics: avoidance, expressive behaviors, nervous/anxious behaviors, cognitive approaches, and self-destructive behaviors. It may be used to refer to the present time only, or, as in its original use, respondents can be asked to give two sets of answers—one for how they dealt with the rape in the weeks and months immediately after it occurred, and a second set for how they deal with the rape now. As with the other growth outcome scales, this measure was validated on adult rape victims and is still used on the same population.

Sample items:

Response categories (referring to both “now” and “in the first several months after the rape occurred”):

- 1=never
- 2=rarely
- 3=sometimes
- 4=half the time
- 5=often
- 6=usually
- 7=always

- 1. Trying to rethink the situation and see it from a different perspective.
- 10. Directly showing your feelings when you are with others--actually crying, screaming, expressing confusion, etc.
- 20. Drinking a lot of alcohol or taking other drugs more than usual.

References Burt, M.R. & Katz, B.L. Coping strategies and recovery from rape. (1988). *Annals of the New York Academy of Sciences*, 528, 345-358.

Burgess, A.W., & Holmstrom, L.L. (1978). Recovery from rape and prior life stress. *Research in Health and Nursing*, 1(4), 165-174.

Horowitz, M. (1976). *Stress Response Syndrome*. New York: Jason Aronson, Inc.

Other Correlates of Well-Being

When examining change over time, it is important to consider other factors that might be influencing survivors' safety and well-being. Especially when evaluating a program that has no control (or even comparison) group, caution must be taken when interpreting findings. For example, a domestic violence program that implemented a shelter-based counseling project might be interested in seeing if the counseling decreased women's depression. Administering a depression measure to a woman entering the shelter and again when the woman exited the shelter might seem logical, and in fact would likely show the desired effects (decreased depression). However, women's depression would be expected to decrease *regardless* of a specific component of the shelter program. Rather, such a decrease would be due to what is called a *time effect*. In this example, the passage of time between the crisis that precipitated entrance into the shelter (the first interview) and leaving the shelter program (second interview) could itself account for the decrease in depression. Similarly, one might choose to interview women six months after they have been sexually assaulted in order to evaluate their level of fear, anxiety, stress, depression, and/or quality of life. How a woman is feeling at any point in time, however, will be influenced by a variety of factors. Is she unemployed, or was she just promoted at her job? Has she had a recent death in the family, or did she just have a baby? Is she dealing with custody issues? What is her level of social support from family and friends? Although a woman's experience of violence would be expected to contribute to her overall feelings of well-being, it is only one of many factors simultaneously influencing her adjustment and overall life satisfaction.

Factors that have been found to correlate with survivors' well-being include (1) their level of social support, (2) their level of daily stress and hassles, and (3) their access to community resources. Instruments designed to measure the level and intensity of these variables are described in the next section:

Access to Community Resources

Effectiveness in Obtaining Resources (Sullivan, Tan, Basta, Rumptz, & Davidson, 1992)
Facility Availability, Usage, and Quality (Coulton, Korbin, & Su, 1996)

Social Support

Multidimensional Scale of Perceived Social Support (Zimet, Dahlem, Zimet, & Farley, 1988)
Social Support Scale (Sullivan, Tan, Basta, Rumptz, & Davidson, 1992)

Stresses and Hassles

Perceived Stress Scale (Cohen, Kamarck, & Mermelstein, 1983)
Survey of Recent Life Experiences (Kohn & Macdonald, 1992)

Access to Community Resources

EFFECTIVENESS IN OBTAINING RESOURCES

Citation Sullivan, C.M., Tan, C., Basta, J., Rumptz, M., & Davidson W.S. (1992). An advocacy intervention program for women with abusive partners: Initial evaluation. *American Journal of Community Psychology*, 20(3), 309-332. Copyright © 1992 by Plenum Publishing Corp. Used with permission.

Description This 11-item scale was created to measure battered women's effectiveness in obtaining desired resources from their communities. For each area in which a woman reported needing help or assistance (legal assistance, housing, employment, education, child care, resources for her children, health care, material goods, finances, social support, "other"), she was asked how effective she was in obtaining that resource. Scale scores were created by taking the mean score across all valid areas. Therefore, a woman who worked on ten different issues did not receive an artificially inflated score when compared to a woman who only needed to work on two issues. This instrument was created specifically to evaluate a post-shelter advocacy intervention program and may need to be modified to effectively evaluate other interventions.

Sample items:

[Following an affirmative response to whether she needed help or assistance in an area]:

How effective have your efforts been in accomplishing your goals in this area? Would you say:

- 1=very ineffective
- 2=somewhat ineffective
- 3=somewhat effective
- 4=very effective

Access to Community Resources

FACILITY AVAILABILITY, USAGE, AND QUALITY

Citation Coulton, C.J., Korbin, J.E., & Su, M. (1996). Measuring neighborhood context for young children in an urban area. *American Journal of Community Psychology*, 24(1), 5-32. Copyright © 1996 by Plenum Publishing Corp. Used with permission.

Description This 13-item scale was created to differentiate between high-risk and low-risk neighborhoods for children with regard to crime, safety and child abuse. It can, however, be modified to ask about the availability of community resources more applicable to women with abusive partners, such as hospitals and clinics, legal aid offices, city hall, community college, etc.

Sample item:

3a. Is there a recreation center in your neighborhood?

1=yes

2=no

b. How would you rate its quality?

very bad 1 2 3 4 5 6 7 8 9 10 excellent

c. Have you used this in the past 2 months?

1=yes

2=no

References Coulton, C.J., Korbin, J.E., Su, M., & Chow, J. (1995). Community level factors and child maltreatment rates. *Child Development*, 66, 1262-1276.

Korbin, J., & Coulton, C. (1994). Neighborhood impact on child abuse and neglect. Final report of Grant No. 90-CA1494. Washington, DC: National Center on Child Abuse and Neglect, Department of Health and Human Services.

Social Support

MULTIDIMENSIONAL SCALE OF PERCEIVED SOCIAL SUPPORT

Citation Zimet, G.D., Dahlem, N.W., Zimet, S.G., & Farley, G.K. (1988). The Multidimensional Scale of Perceived Social Support. *Journal of Personality Assessment*, 52(1), 30-41. Copyright © 1993 by Laurence Earlbaum Associates. Reprinted with permission.

Description This 12-item scale measures perceived social support in the area of (a) friends, (b) family, and (c) significant other. It was validated on college students but has also been used by Barnett et al. (1996) with a sample of women with abusive partners. This instrument is simple to use and interpret, and can be administered verbally or in written format.

Sample items:

Response categories:

- 1=strongly disagree
- 2=disagree
- 3=neither disagree or agree
- 4=agree
- 5=strongly agree

[significant other]

1. There is a special person who is around when I am in need.

[family]

4. I get the emotional help and support I need from my family.

[friends]

12. I can talk about my problems with my friends.

References Barnett, O.W., Martinez, T.F., & Keyson, M. (1996). The relationship between violence, social support, and self-blame in battered women. *Journal of Interpersonal Violence*, 11(2), 221-233.

Dahlem, N.W., Zimet, G.D., & Walker, R.R. (1991). The Multidimensional Scale of Perceived Social Support: A confirmation study. *Journal of Clinical Psychology*, 47(6), 756-761.

Social Support

SOCIAL SUPPORT SCALE

Citation Sullivan, C.M., Tan, C., Basta, J., Rumpitz, M., & Davidson W.S. (1992). An advocacy intervention program for women with abusive partners: Initial evaluation. *American Journal of Community Psychology*, 20(3), 309-332. Copyright © 1992 by Plenum Publishing Corp. Used with permission.

Description This 9-item scale was modified from Bogat et al.'s (1983) Social Support Questionnaire, and measures the amount and quality of respondents' social support. This scale has been used extensively with women who are or who have been involved with abusive partners. The instrument has been found to be sensitive to change over time (Tan et al., 1995), and is easy to administer and interpret. In creating a scale score, items should be reverse-scored and summed, so that a higher score indicates higher social support.

Sample items:

Response categories:

- 1=extremely pleased
- 2=pleased
- 3=mostly satisfied
- 4=mixed (equally satisfied and dissatisfied)
- 5=mostly dissatisfied
- 6=unhappy
- 7=terrible

Thinking about the people in your life, family and friends:

- 3. In general, how do you feel about the amount of advice and information that you receive?
- 8. In general, how do you feel about the quality of emotional support that you receive?

References Bogat, G.A., Chin, R., Sabbath, W., & Schwartz, C. (1983). *The Adult's Social Support Questionnaire*. Technical report. East Lansing, MI: Michigan State University.

Tan, C., Basta, J., Sullivan, C.M., & Davidson, W.S. (1995). The role of social support in the lives of women exiting domestic violence shelters: An experimental study. *Journal of Interpersonal Violence*, 10(4), 437-451.

Stresses and Hassles

PERCEIVED STRESS SCALE

Citation Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24, 385-396. Copyright © 1983 by the American Sociological Association. Reprinted with permission.

Description This 14-item instrument measures the degree to which individuals find their lives to be stressful. While some scales measure *objective* stressors, such as whether someone has recently lost his or her job, experienced the death of a loved one, etc., this scale measures the respondents' *perceived* stress. The items were designed to assess the degree to which people find their lives to be unpredictable, uncontrollable, and overloading. The scale was validated on college students and members of a smoking-cessation program, but has also been used by Tutty et al. (1993) with a sample of women who have abusive partners.

Sample items:

Response categories:

- 0=never
- 1=almost never
- 2=sometimes
- 3=fairly often
- 4=very often

- 1. In the last month, how often have you been upset because of something that happened unexpectedly?
- 4. In the last month, how often have you dealt successfully with irritating life hassles?
- 14. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?

References Hewitt, P.L. (1992). The Perceived Stress Scale: Factor structure and relation to depression symptoms in a psychiatric sample. *Journal of Psychopathology and Behavioral Assessment*, 14(3), 247-257.

Tutty, L.M., Bidgood, B.A., & Rothery, M.A. (1993). Support groups for battered women: Research on their efficacy. *Journal of Family Violence*, 8(4), 325-344.

Stresses and Hassles

SURVEY OF RECENT LIFE EXPERIENCES (SRLE)

Citation Kohn, P.M., & Macdonald, J.E. (1992). The Survey of Recent Life Experiences: A decontaminated hassles scale for adults. *Journal of Behavioral Medicine*, 15(2), 221-236. Copyright © 1992 by Plenum Publishing Corp. Used with permission.

Description This 51-item instrument was designed to measure individuals' exposure to hassles over the prior month. It contains six subscales: social and cultural difficulties [11 items], work [7 items], time pressure [8 items], finances [6 items], social acceptability [5 items], and social victimization [4 items]. This instrument was initially validated with a sample of visitors to the Ontario Science Centre, who were primarily young and highly educated. It has since been validated with a Dutch population (de Jong et al., 1996), but its generalizability to lower-income individuals and/or women who have experienced violence has yet to be determined.

Sample items:

Response categories (referring to "how much a part of your life" each experience has been in the prior month):

- 1=not at all
- 2=only slightly
- 3=distinctly
- 4=very much

- 1. Disliking your daily activities.
- 9. Too many things to do at once.
- 27. Financial burdens.

References de Jong, G.M., Timmerman, I.G.H., & Emmelkamp, P.M.G. (1996). The Survey of Recent Life Experiences: A psychometric evaluation. *Journal of Behavioral Medicine*, 19(6), 529-542.

Kohn, P.M., Hay, D.H., & Legere, J.J. (1994). Hassles, coping styles, and negative well-being. *Personality and Individual Differences*, 17(2), 169-179.

CHAPTER 8

DESCRIBING VICTIM SERVICES AND SUPPORT SYSTEMS

More than a third of STOP funds are being used to support a wide variety of victim services and other activities to help women victims of violence. The minimum goal of these projects is simply to make more victim services available, on the assumption that such services promote better victim outcomes (as detailed in Chapter 7). An additional goal of many projects is to change something about the ways that victim services are delivered. Some projects focus on changing the numbers or types of women they serve, which may also entail changing some aspects of service delivery. Some projects are trying to develop better integrated services within a full-service advocacy context. Other projects help victims during their involvement with a criminal justice system agency (e.g., Victim Witness Assistance Programs), and may have goals pertaining to system as well as victim outcomes (e.g., faster case processing, as well as better victim understanding of what is happening). Finally, some projects use advocacy to influence the ways that legal and other systems treat women victims of violence. Advocacy may occur at the case-specific level (i.e., helping an individual woman negotiate the system) or at the system level (i.e., trying to improve treatment of women in the system as a whole).

To know what has changed, one needs to know what has been. With respect to victim services and systems, this means we need to measure agencies using set of dimensions that encompasses their most important aspects—the dimensions across which programs serving and supporting victims may differ the most. This chapter offers a set of dimensions and some suggestions for how to assess them. The set is not exhaustive, but it does include quite a number of critical program and system characteristics. The examples it provides should give you some ideas about how to develop additional measures for yourself.

INSTITUTIONAL AFFILIATION, ORGANIZATIONAL STRUCTURE, AND LINKAGES

Activities supporting women victims of violence occur in many venues and are organized in many ways. As with changes to the civil and criminal justice systems described in Chapter 9, there are not many “tried and true” ways to measure the structure and activities of the programs and systems designed to help women. Table 8.1 offers a few of the most important dimensions, plus suggestions for possible measures. We cover venue, the structure of activities within an agency, and linkages to other agencies.

Venue

“Venue” refers to both the type of agency that houses a project and to its location. The *type of agency* offering the activity or service may be a justice system agency, another governmental agency, or a nongovernmental agency. Common agency types for STOP grants are law

Evaluation Guidebook

enforcement, prosecution, judicial, and nonprofit, nongovernmental victim service agencies, but others are also possible. You might want to start with the list in the SAR, items 17A and 17B (see Chapter 5), and expand it as necessary. Among the things you will want to examine are the following:

- Whether the activities within a particular type of agency expand,
- Whether some victim service agencies that existed prior to STOP begin receiving a larger (or smaller) proportion of the total funds available to support victim service activities,
- Whether new types of agencies come into the picture (or try to) that have not previously been involved in victim services, and
- Whether agencies change the mix of services they provide.

The *location* of STOP projects is also of great interest, given the focus of the VAWA on reaching underserved populations. Critical dimensions of location include:

- Whether the program site is inner city, other urban, suburban, or rural,
- Whether the program site is in or near major concentrations of women who would be considered underserved (e.g., women from urban racial, ethnic, cultural, or language minority communities), and
- Whether the program is located in an agency with a history of serving women from designated underserved populations.

Structure of Activities

“Structure of activities” refers to several things, including vertical integration, duration of the problem integration, horizontal integration, and non-case-based advocacy. *Vertical integration*, in its ideal form, means that any woman coming for services is aided by the same staff person throughout the course of her case/situation/time with the agency. Thus one prosecutor does everything with a woman involved in prosecution, from the initial interview to the case preparation to the court appearance to the follow-up.

Duration of the problem integration means, again ideally, that the same agency, usually a victim support agency, offers a woman support from her first call through the entire duration of the violent conduct affecting her, the risk to her safety, and her psychological recovery. In many

instances, this support extends well beyond the period during which a criminal case is in progress. This is especially important to note with respect to sexual assault, as most rape victims using hotlines, counseling, and other supportive services are never involved in a criminal case. It is also relevant to many battered women who come for counseling but are never involved in either a criminal or a civil case.

Horizontal integration, in its ideal form, means that an agency offers advocacy and assistance in all areas of a woman's life that the woman asks for or is interested in receiving. Such assistance may range from shelter and food to employment, housing, income support, education, child care and parenting, health screens, help with substance abuse problems, legal advocacy, psychological adjustment and recovery, counseling for the woman and sometimes other family and friends, safety planning, and so on. Help may be available through the program itself, or by linkages with other agencies. Linkages are especially likely for the more conventional benefits such as housing, income support, and child care.

Non-case-based advocacy is activity focused on system reform to enhance the safety and autonomy of women, prevent their further victimization by systems that are supposed to help them, see that offenders are held accountable, and increase the resources for helping women victims of violence. It also includes general advocacy to keep the spotlight focused on ending violence against women. Sometimes advocacy can also be case-based, as when an agency helps a particular woman deal with whatever part of the system she must negotiate.

Linkages

“Linkages” refers to the nature and variety of agencies, services, and activities with which your program has connections. “Connections,” as used here, does not simply mean that you give out their phone number and tell a client to call. Rather, it means that you have some type of active arrangement with another program or agency to send your clients over, to help your clients get services, and so on. It is different from the situation in which your own agency staff goes with a client to help her negotiate the nitty-gritty of another agency such as the welfare office, the court, etc.—this situation falls under the “advocacy” category.

SERVICES, ACTIVITIES, AND CLIENTS

A major goal of programs operating under the victim services purpose area is to increase the variety and volume of services available to women. Another important goal of many STOP projects (not just those designated as “victim services”) is to reach more women from the populations designated by the VAWA as underserved, including women from racial, ethnic, cultural, language, and sexual orientation minority communities, and rural women (who may also

share other aspects of “underservedness”). Table 8.2 shows several objectives that are part of reaching these goals, including more and better services and advocacy; more women using services, especially those from underserved populations; and increased cultural competence of agency staff.

More and Better Services and Advocacy

You will want to document the type of services that have been developed or expanded, the level of service use, the quality of services, and their appropriateness to women’s needs. Type of service to be expected will vary by type of agency, so there is no uniform list that will work for all STOP projects. Here are some examples of service types that might be relevant to different types of agencies:

- For a victim-witness assistance program (VWAP), services might/should include interviewing a woman who has a case pending with the agency to gather relevant data, informing her of what to expect from the agency, developing a safety plan with the woman, keeping her informed about the progress of her case and the location/situation of the offender (if he is in custody), preparing her for testimony/court/other events where she must be present, offering information and linkages to sexual assault/domestic violence services in the community.
- For a sexual assault nurse examiner (SANE) program, services might/should include providing a private space for the interview and examination, timely and courteous treatment, information and counseling, explaining the steps in the exam, performing the evidentiary exam, providing medical treatment as necessary, providing support for any accompanying family or friend, offering follow-up contact and/or information and linkages to sexual assault services in the community, secure transfer of evidence, court testimony.
- A hotline program might be expected to provide information and referral, telephone counseling, offers of in-person counseling or other services, callback or follow-up if the caller wants it or accepts it when it is offered. The content of telephone counseling might be further differentiated into concrete information, help thinking about safety, help anticipating the future, help dealing with feelings, help dealing with reactions of others, and so on.
- A domestic violence shelter might be expected to provide food and lodging; safety planning; assessment and case management; links to public programs such as income maintenance, housing, or child care; help with parenting; court advice and advocacy for both civil and criminal cases; and so on.

The variety and volume of services or activities are relatively easy to document or measure. Documentation of variety will come from published agency documents, and can also be verified through interviews with program staff and clients. Service volume or use levels have a clear meaning and lend themselves to counting (e.g., number of hotline calls last month, number of nights of shelter provided). However, it can be a laborious task to establish the data system that will let you record service use levels, and an even more strenuous exercise to get program staff to enter the necessary data into the system. See Chapter 14 on developing data systems for thoughts on these issues.

The quality and appropriateness of services are considerably more difficult to conceptualize and measure. Quality can be measured in any number of ways, including timeliness, supportiveness of interactions between program and client, completeness in relation to scope of the client's issues and problems, and appropriateness. Appropriateness can be defined for this purpose as "Did she get what she needed, and did she get it when she needed it?" We suggest several sources of information on quality and appropriateness, each of which has its advantages and disadvantages. These sources are client feedback, expert ratings, and process feedback from key informants in the community who interact with the project and its clients.

Client feedback means asking clients whether they feel the services they received were what they needed, when they needed it. They can be asked whether they were satisfied with the services in general, and with various aspects of particular services. They can be asked whether there was anything they needed that they did not get, or whether they were treated in any ways that put them off. Nothing can replace this type of direct feedback from clients, and most programs find it invaluable in defining what they do well and what they could do better.

However, clients are generally inclined to be satisfied with services, especially when the questions asked are very broad. In addition, they usually do not know the full scope of what they might have gotten in an ideal full-service one-stop shopping agency that puts client needs uppermost. So it helps to get a reading on service quality from other sources in addition to clients. Therefore we include expert ratings and key informant feedback, as described below.

Expert ratings are one other source for evaluating service quality. Experts could be given a sample (random, of course) of case records containing client need assessments and documentation of agency contacts and service use. For each case they are asked to rate the thoroughness of the need assessment, the appropriateness of the help given to the needs revealed by the assessment, and areas where help might have been important but where it was not given. If they are familiar with the community, they might also be asked to identify opportunities for assistance that exist in the community but were missed in the case at hand.

An alternative approach to expert ratings can be used for the many agencies that do not keep written case records. In this alternative approach a meeting is set up with agency staff and the

Evaluation Guidebook

outside experts. The experts bring sample cases, usually starting with very simple ones and proceeding to ones of greater complexity. Staff describe how they typically handle each case, what resources they would call upon, how long they would follow the case, what they would look for as complicating issues, and so on. The experts can then reach a judgment about the general comprehensiveness and appropriateness of staff efforts with respect to a variety of cases.

Key informant feedback is a process analysis technique used to obtain perceptions of the program and its services from knowledgeable people in the community. These qualitative data can give you an idea of whether other agencies perceive the referrals they receive from the project being evaluated as appropriate, whether adequate documentation comes with the referrals, whether project staff are easy to work with on individual cases, and so on. Even if an adversarial relationship sometimes exists between the key informant's agency and the project being evaluated (due to advocacy work or other past interactions), you still can often get valuable information about the respect with which the project is perceived and the quality of the supports it offers to women. In fact, it can be extremely useful to know that the police department, prosecutor's office, or judge perceives the victim support program to be an effective advocate for its clients and sees it as having stimulated some changes in the system.

More Clients/Users, More Underserved Clients/Users

Your intake forms and other case records can give you information about important client characteristics. You should include all the characteristics requested by the Subgrant Statistical Summary (see Chapter 5), especially those related to characteristics that identify underserved populations. You will probably want to add more characteristics to your own forms, depending on the information you have found to be essential to helping women victims of violence.

VAWGO identifies women from racial, ethnic, cultural, language, and sexual minority groups, and rural women, as "underserved." Projects engaged in special outreach efforts to make services available to women from these groups in culturally competent ways should count their activities, and the women they reach, as relevant to the "underserved" criterion. However, if your agency *routinely* serves, or deals with, women from these groups just because that has been the nature of your business, and nothing has changed about the way you do it, then you do not meet the "underserved" criterion in VAWGO's terms. That is, if you are a VWAP in a police department, and serve mostly African-Americans because your jurisdiction has mostly African-American residents and that's who the department sees, this is your "traditional" population. You serve them all the time; they are not "underserved" by your department. If you make a special effort to reach members of another population, or if you do something substantially different so as to become more culturally competent and acceptable to the African-American community, then you can count the new and/or better-served women as "underserved" in VAWGO's terms.

Increased Cultural Competence

A very large agency might think about assessing cultural competence by comparing the demographic makeup of its staff to that of the community it serves. However, in a small agency, the percentage figures this would produce will not have much meaning. It may be more important to ask whether staff have had adequate training on cultural competence issues; to assess staff attitudes toward women from underserved populations, both in general and with regard to violence against women; and to identify the presence or absence of language competencies. In addition, it is critical to ask clients about their experiences with staff in relation to cultural competence. You could also ask the community at large about their perceptions of the agency as being hospitable to community members. You can do this through focus groups, interviews with key informants from the target underserved communities, or community surveys of the general public.

Table 8.1
Institutional Affiliation, Organizational Structure, and Linkages

Objectives	Specific Measures	Data Collection Procedures	Caveats
<p>Expanded venues—more agencies, in more places, offer new or expanded victim services.</p> <p>Type of agency Location of agency VAW focus</p>	<p>Categorical data, specifying agency type and location. Use categories from SAR, or develop expanded categories.</p> <p>VAW focus = agency whose primary mission is offering service to women victims of violence. May be further differentiated into DV, SA, or both. If the agency’s primary mission is not VAW, this variable would be scored zero.</p>	<p>Review of agency documents, grant application.</p> <p>Observation.</p>	<p>When identifying nonprofit, nongovernmental victim service agencies, be sure to use separate categories for generic victim services (e.g., VOCA, crisis hotline) and agencies that have a specific focus on violence against women.</p>
<p>Structure of services and activities increase in vertical, duration of problem, and horizontal integration,, and in non-case-based advocacy.</p>	<p>Duration of contact = average amount of time your agency is in contact with a client.</p> <p>Amount of vertical integration = number of people at your agency who deal with the average client during that client’s contact with your agency.</p> <p>Amount of duration-of-problem integration = proportion of time-from-beginning-to-full-resolution that your agency offers support to clients.</p> <p>Amount of horizontal integration = proportion of life issues relevant to clients that your agency routinely provides.</p> <p>Amount of advocacy = percentage of staff time, or of agency budget, devoted to non-case-based system advocacy.</p>	<p>You will need a “measuring stick” of what “full service” means for your clients—what is the complete package of assessment, services, etc. that would meet their needs. Then you have to examine your agency’s activities to see what proportion of that need you address (horizontality). For duration-of-problem, the measuring stick should relate to time-from-beginning-to-resolution (as defined by the woman), and how much of that time your agency is available to the client.</p> <p>Also case records, agency budgets and staffing patterns, process analysis of agency activities, time and staff allocations.</p>	<p>There is no standard “measuring stick.” You may have to do quite a few complete client needs assessments to get the whole picture, even if your agency has no capacity to meet many of these needs.</p>

Table 8.1
Institutional Affiliation, Organizational Structure, and Linkages

Objectives	Specific Measures	Data Collection Procedures	Caveats
<p>Agencies establish more linkages with important supports for women throughout community.</p>	<p>Number and type of agencies/services/activities with whom your agency commonly shares clients.</p> <p>Nature of linkages—formal agreements between agencies; informal arrangements between line workers, co-location of staff or joint action in response to VAW incidents, degree of shared understanding, cross-training, etc.</p>	<p>Case records of your and other agencies.</p> <p>Client interview data reporting services received from other agencies and help received from your agency in accessing those services.</p> <p>See also Chapter 10 for measures of collaboration.</p>	

Table 8.2
Characteristics of Services, Activities, and Clients

Objectives	Specific Measures	Data Collection Procedures	Caveats
<p>More and better services, activities, and advocacy will be available to women.</p> <p>Types of services/advocacy Amounts/Use levels Appropriateness and Quality</p>	<p>Number of different types of services, activities, advocacy; number given by program, number given by referral.</p> <p>Number of women using each type of service, help.</p> <p>Number of units given, for each type of service or help.</p> <p>Proportion of clients who receive some advocacy assistance in addition to direct services.</p> <p>Proportion of clients who got what they needed, when they needed it.</p> <p>Proportion of clients satisfied with help received.</p>	<p>Types of services/activities documented by content analysis of agency documents.</p> <p>Numbers of services, use levels, and source of service documented from agency records, including case records/client tracking system.</p> <p>Client ratings of own service, from interviews; expert ratings of sampled cases, using victim service advocates from outside the agency being evaluated; process feedback from key informants.</p>	<p>Expected/desirable types of services/activities will vary by agency. You will need to develop a “standard” for your type of agency and then assess your own activities against it.</p>
<p>More women, and more women from previously underserved populations, will use services/advocacy.</p>	<p>Number of women using each type of service or help, categorized by “underserved” characteristics.</p> <p>Number of women with other characteristics of importance to your project.</p>	<p>Client/case records and interviews. Client databases.</p>	<p>Be sure to define “underserved clients” as women from a group that has not usually received services and whether your project has a specific focus on reaching this population.</p>

Table 8.2
Characteristics of Services, Activities, and Clients

Objectives	Specific Measures	Data Collection Procedures	Caveats
Cultural competence of staff will increase.	<p>Percent of staff trained in cultural competence.</p> <p>Variety of resource materials available in relevant languages.</p> <p>Staff demographic profile compared to community demographic profile.</p> <p>Staff attitudes are supportive, accepting, positive.</p> <p>Client ratings of staff cultural sensitivity.</p>	<p>Examine agency training records.</p> <p>Examine resource materials.</p> <p>Compare agency personnel records with census data on community population. Use personnel records to assess staff language competencies.</p> <p>Use standard attitude measures.</p> <p>Client interviews.</p>	

CHAPTER 9

EVALUATING CRIMINAL AND CIVIL JUSTICE AGENCY CHANGES

One of the most important long-range goals of the STOP program is to change permanently the way justice system agencies do business. Change might take many forms, from law enforcement officers making more arrests, to prosecutors pursuing cases more vigorously, to courts imposing more victim-oriented conditions and sentences, to probation and parole officers supervising offenders more closely. Whatever type of change your STOP-funded project aims to produce, it is important to document the project's effects on how agencies work so you will know what you have accomplished and what is left to be done. This chapter provides some guidance in how you can go about anticipating and measuring changes in agency functioning.

Many different types of STOP projects can potentially impact how justice system agencies do their jobs. Projects falling under the training, special units, and policy development purpose areas clearly intend to improve the performance of justice agency personnel by providing them with more or better knowledge, resources, or work guidelines. Data system development projects may also be designed to help justice personnel do better work by providing them with improved information for making arrests, prosecuting cases, and sentencing and supervising offenders. Projects in the victim service area may include a collaboration component to help justice agencies serve victims better or develop improved training or policies. Projects on stalking or for Indian tribes might include any of a wide variety of activities, many of which may help justice agencies work better.

Even projects that are not explicitly intended to change how an agency works might bring about some changes along the way. For example, a private nonprofit victim service agency may be funded to provide in-court support to victims. Having these staff available may free up the victim/witness staff in the prosecutor's office to put more time into preparing victim impact statements, which may help judges make more informed sentencing decisions that provide better protection and justice to victims. If you predict any changes in how an agency works or is structured when you specify your project's logic model (see Chapter 2), whether these changes are the end goal, an intermediary step, or simply a by-product, you should consider how to measure these changes so you will have a thorough understanding of what your project has accomplished and how it made things happen.

You should also think about documenting or measuring agency changes that may happen completely independently of your project. These may affect your project and help explain its impact. But don't overdo it. Justice agencies can be laden with long-standing traditions and bureaucracies that resist change. When you consider what changes might come about and how you would measure them, be realistic about what you can reasonably expect given your STOP project's activities, timeline, and the general flexibility or inflexibility of the system you're working in. Also bear in mind the time frame of your evaluation and the resources you'll have

available for evaluation activities when you decide what are the most worthwhile evaluation issues to pursue.

The rest of this chapter presents some ideas about what sorts of changes might occur in justice agencies and how to measure them, whether they are the goals of STOP projects or not. Tables are provided for each of the four major types of justice agencies—law enforcement (Table 9.1), prosecution (Table 9.2), judges and courts (Table 9.3), and corrections (Table 9.4). Each table provides some suggestions for how the agency might change, how you could measure each type of change, and where to get that information. These are only suggestions for the most general or obvious types of changes that might occur. You should consider what changes your project is likely to produce when you select from among these or devise your own.

This chapter is concerned with changes in tangible behaviors and agency resources rather than changes in individuals' attitudes, knowledge, or beliefs. Since behavioral change is not the sort of thing that lends itself to standardized measures, it's a bit more complicated than putting together a questionnaire and handing it out. In some cases the information you need may be available in an agency's case files or, better yet, in its automated records. In other cases you may wish to interview concerned parties to get richer, more in-depth information on how and why things have (or haven't) changed. Some creativity may be required in figuring out the least intrusive and most efficient way to get the information you need, and how to get the agency concerned to be willing to make any necessary changes in the way records are kept. Thinking through the issues involved in maintaining confidentiality will stand you in good stead here. When people feel they can trust you to keep their names separate from any of the things they tell you, they will tell you a lot more. If, on the other hand, they feel they cannot trust you and that information they give you may come back to haunt them, they will be a good deal more close-mouthed and you will not learn what you need to know.

These variables and measures can be used to make a number of different types of comparisons, depending on your research design. If you are using a pre/post design, for example, you can compare interview and evidence collection techniques before and after the officers in the law enforcement agency were trained or the new policy was implemented. Or, as another example, you can compare prosecutors' charging practices across jurisdictions if you are using a non-equivalent comparison group design. The critical point is to select the appropriate variable reflecting agency change, decide how best to measure it, and then measure it as consistently as possible across the time periods or sites your research design includes.

Also please note: there is no need to treat the suggestions in the following tables as gospel. They are merely suggestions of objectives and some measures that might be relevant to them. Certainly you should use these suggestions to stimulate your own thinking about some measures that you feel would offer convincing evidence that you have met your goals. Modify our suggestions to fit your circumstances, and add your own.

DON'T FORGET BACKGROUND AND EXTERNAL FACTORS

Tables 9.1 through 9.4 identify outcomes of interest and their potential measures within justice system agencies. Collecting data on these outcomes lies at the heart of any evaluation design. They are not, however, the only data you need to collect. Be sure to construct a logic model of your project, in which you note not only the services you offer and their outcomes, but also *the background and external factors that could make your job harder or intervene to nullify the impact of the services you offer*. You will need to include plans for collecting data on these factors as part of your evaluation design. Data on background factors (characteristics of clients, of staff, etc.) may be available from intake interviews, case records, personnel records, or research interviews. Data on external factors are most likely to come from process analysis, using any of the methods described in Chapter 6.

Table 9.1: Impacts on Law Enforcement Agencies

Objectives	Specific Measures	Data Collection Procedures	Caveats
<p>Improve dispatchers' identification of calls for service as VAW.</p>	<p>Percent of correct classifications = number of calls which both dispatchers and responding officers classified as either VAW or not VAW, divided by the total number of VAW calls to which officers were dispatched.</p>	<p>Dispatchers must record, for every call, classification made (VAW vs. non-VAW) and whether officers were dispatched. Officers must record, for every call they're dispatched to, classification made (VAW vs. non-VAW).</p>	<p>This assumes that the classification of responding officers is the correct one, and judges dispatchers' classifications against them. Multiple calls on the same incident should be counted as a single event.</p>
<p>Improve dispatchers' response to calls for service by prioritizing VAW calls and always dispatching officers as soon as possible; accessing and providing officers with information on the history of the perpetrator; and providing the victim or other caller with crisis intervention and/or safety planning.</p>	<p>Priority rating for VAW calls = average priority assigned (sum of priority assigned to each VAW call divided by the total number of VAW calls).</p> <p>Dispatching rate = number of VAW calls to which officers were dispatched to respond divided by the total number of VAW calls.</p> <p>Timeliness of dispatching = number of minutes between time an officer becomes available to respond (or time of call if officers are available immediately) and the time he or she is dispatched.</p> <p>Provision of information to officers = number of calls in which the dispatcher provides any information on the perpetrator's criminal history, Brady indicators, protection order history, firearms possession or use history, or history of injuring victims, divided by the total number of calls for which such information is available.</p> <p>Services to victims/callers = proportion of calls in which dispatchers attempt to help the victim or caller meet immediate safety needs while waiting for the officers to arrive.</p>	<p>Dispatchers must record priority rating assigned to every VAW call, whether officers were dispatched to respond, time of call, time an officer became available, time officer was dispatched, provision of perpetrator history information to officers, and response to victim/caller.</p>	<p>Multiple calls on a single incident should be counted as one event.</p>

Table 9.1: Impacts on Law Enforcement Agencies

Objectives	Specific Measures	Data Collection Procedures	Caveats
<p>Improve officers' response to dispatchers' instructions by increasing their response rates and timeliness.</p>	<p>Response rate = number of VAW calls to which officers responded on the scene divided by the number of VAW calls to which they were dispatched.</p> <p>Timeliness of response = number of minutes between the time an officer is dispatched and the time he or she arrives at the scene.</p>	<p>In addition to dispatchers' records, above, officers must record whether they responded to the scene and, if yes, their time of arrival.</p>	<p>Any delays in response that look unusually long should be assessed by examining the officers' location when dispatched and how much distance and traffic they had to get through to reach the scene of the call.</p>
<p>Improve interviewing, evidence collection, and report writing by responding officers and detectives to whom cases are referred for follow-up work. Incident reports and other relevant documents should be available to victims, prosecutors, and victim services programs more quickly.</p>	<p>Quality of on-scene and follow-up interviewing practices with victims, perpetrators, and witnesses = average ratings of interviewing practices. Good practices include on-scene separation of parties before interviewing (if DV), conducting interviews in a private setting, showing respect and courtesy, asking appropriate questions, and making thorough notes (including "excited utterances" that may be admissible in court without the victim's corroboration).</p> <p>Quality of documentation of physical evidence = average ratings of documentation practices. Good practices include collection of clothes or objects damaged in the assault, use of technology (e.g., photographing injuries or crime scene, DNA testing, referral for SANE or other rape kit examination, etc.), follow-up evidence collection (e.g., returning to photograph injuries several days after the assault when bruises have appeared), securing the chain of custody of evidence, and other means to improve casework.</p> <p>Percent of calls for which written report completed even if no arrest is made.</p> <p>Average number of days between incident and availability of report to victims, prosecutors, and victim service programs.</p>	<p>Quality of interviews can be assessed with surveys of victims, perpetrators, and witnesses. Criteria can include whether interviews were conducted in a private and secure context; interviewees' ratings of officer demeanor; and content of questions. Case files can also be examined to assess interview setting; content of questions; and whether notes were logical and thorough.</p> <p>Quality of evidence collection can be measured by reviewing case files and comparing them against standards of practice and resources available in the jurisdiction. For example, if victim reports injuries or damage and cameras are available to officers, there should be photographic evidence.</p> <p>Compare dispatch records to arrests to written reports, from agency records.</p> <p>Compare date of incident with date report becomes available.</p>	<p>Criteria for good investigation practices should be informed by local policies, regulations, and resources.</p> <p>Analyses should examine variations as well as average ratings to identify model practices and areas for improvement.</p>

Table 9.1: Impacts on Law Enforcement Agencies

Objectives	Specific Measures	Data Collection Procedures	Caveats
<p>Improve assistance to victims by officers, detectives, advocates, counselors, and other law enforcement personnel.</p>	<p>Percent of victims who are offered assistance such as notification of rights, service referrals, and assistance with immediate needs (such as transportation for emergency medical treatment or shelter, or referral for hotline crisis counseling). For victims offered each type of assistance, percent who received or did not receive it, and reasons why not (e.g., victim refused, assistance not immediately available, victim not eligible for assistance, etc.).</p> <p>Percent of victims who are offered follow-up assistance such as case updates, information on the legal system, and service referrals. For victims offered each type of assistance, percent who received or did not receive it, and reasons why not. Can also measure the number of offers over time and the length of time over which the victim receives offers of assistance.</p> <p>See also Chapter 7, proximate measures, for victim feedback on agency performance; Chapter 8 for characteristics of victim services.</p>	<p>Victim assistance can be assessed through victim surveys and reviews of law enforcement case files documenting assistance offered and whether victims accepted or declined the assistance. Victim surveys should also identify types of help that victims needed but were not offered.</p>	

Table 9.1: Impacts on Law Enforcement Agencies

Objectives	Specific Measures	Data Collection Procedures	Caveats
<p>Officers and detectives should enforce laws against VAW more vigorously. More cases should be substantiated, more warrants issued, and more arrests made.</p>	<p>Percent of substantiated cases = number of cases of DV or SA found to have probable cause divided by the total number of cases investigated.</p> <p>Percent of cases pursued = number of substantiated cases in which search warrants, arrest warrants, or other types of follow-up options were both needed and obtained, divided by the number of substantiated cases where follow-up actions were needed (whether they were accomplished or not).</p> <p>Percent of cases with arrests = number of substantiated cases with an arrest divided by the total number of substantiated cases.</p> <p>Percent of cases in which victim’s account of the incident (or 911 description) corresponds with the information that police included in their report, and with the resulting charge, if any.</p>	<p>Automated agency database should contain, for each case, whether the case was substantiated and what follow-up actions were taken, including warrants, arrests, and other relevant legal options. If there is no adequate automated database, this information should be available in case files.</p> <p>911 records, or independent interviews with victims, will be needed in addition to information from above sources.</p>	<p>There should be a trend away from victim-initiated actions and toward officer-initiated actions. If case files are reviewed, look for signatures indicating victims took the initiative to request the action (such as obtaining an arrest warrant or making an arrest).</p>

Table 9.1: Impacts on Law Enforcement Agencies

Objectives	Specific Measures	Data Collection Procedures	Caveats
<p>Agencies should have more and better resources to respond to cases of VAW.</p>	<p>Percent of agencies which have thorough, progressive written policies on VAW cases.</p> <p>Percent of agency personnel who receive training in a timely fashion when new policies are implemented.</p> <p>Percent of agencies that have adequate equipment and supplies to handle all VAW cases using up-to-date technology.</p> <p>Percent of agencies that regularly monitor dispatchers', officers', and other relevant staff's responses to calls and sanction non-compliance with agency policies.</p>	<p>Assess thoroughness and progressiveness of policies by content analysis. Policies should address a variety of important current topics.</p> <p>Assess training by review of training records for all agency personnel.</p> <p>Assess adequacy of equipment and supplies by making an inventory and interviewing managers and officers for their perceptions of whether they are sufficient given the number of cases.</p> <p>Interview managers, officers, dispatchers, and other staff to identify supervision practices and areas for improvement.</p>	<p>Important policy topic areas may be determined in part by current issues in the state or community. Some topic areas might include warrantless arrest, officer involvement in DV, providing services to victims, coordinating with other agencies, etc.</p>
<p>Agency staff should be culturally competent and have culturally appropriate resources available to them.</p>	<p>Percent of staff trained in cultural competence.</p> <p>Number and types of cultural competence resource materials available (e.g., information on cultural beliefs or norms about women and violence, for populations found in the jurisdiction).</p> <p>Staff demographic profile compared to community demographic profile.</p> <p>Percent of victims with language barriers for whom an appropriate interpreter is provided. Number and types of forms and brochures available in languages spoken by the population of the jurisdiction.</p>	<p>Examine agency training records and resource materials.</p> <p>Compare agency personnel records with census data on community population.</p> <p>Examine case records, and also records of use of interpreter service, if available. Use personnel records to assess staff language competencies. Examine forms and informational materials available to the public.</p>	

Table 9.1: Impacts on Law Enforcement Agencies

Objectives	Specific Measures	Data Collection Procedures	Caveats
<p>Special units or other ways of giving special attention to DV and SA cases should be more widespread, should handle many of the cases they were designed for, should be staffed with qualified personnel who should have reasonable caseloads, and should make valuable contributions to the agency's functioning.</p>	<p>Percent of agencies with special units or other ways to give special attention to DV, SA, or both. Percent of the total DV, SA, or DV and SA cases which are referred to the special unit (depending on what type or types of cases the unit is designed to handle). Number (in FTEs) and types of staff assigned to the unit (may include officers, detectives, counselors, or other personnel). Staff's average level of training and experience in DV, SA, or both. Number of cases per staff member or team (depending on unit's structure) as measure of caseload. Average rated value of the unit's contributions to the agency. Reduced number of interviews with victim, people involved with victim, duplicate information-gathering by prosecutors, allowing smoother transition to prosecution.</p>	<p>Assess special units or staff designated to give special attention; the number, types, and qualifications of personnel; and the improvements in case processing through interviews with management and with victims, and through reviews of agency policies, organizational charts, annual reports, personnel records, etc. Measure case referrals and caseloads through agency database (or case file reviews). Measure unit's contributions through interviews with managers and staff both inside and outside the unit, and with victims.</p>	
<p>Staff should be well-trained in VAW cases.</p>	<p>Percent of staff who meet minimum requirements for VAW training, as specified by state POST or other training authority, or by state or local law, policy, regulation, etc. Percent of specialized staff who have advanced training in DV, SA, or both, depending on type of unit. Percent of staff following correct procedures learned in training (e.g., training to identify the primary aggressor measured by reductions in number of dual-arrest cases).</p>	<p>Measure by review of personnel records or records of the training authority.</p>	

Table 9.1: Impacts on Law Enforcement Agencies

Objectives	Specific Measures	Data Collection Procedures	Caveats
<p>Agency's caseload should increase due to higher reporting rates and greater case activity (e.g., more investigation, more arrests, etc). Increased caseload will lead to the need for additional staff, overtime, reorganization, or redeployment of staff.</p>	<p>Reporting = number of VAW reports divided by the total number of crime reports, including 911 calls and walk-in reports at police stations and other places officers are stationed (such as hospital emergency rooms).</p> <p>Caseload = the average number of cases per staff and the average amount of time put into each case.</p> <p>Increased demand arises from increases in staff caseload and time spent per case, over time.</p>	<p>Dispatchers and others receiving reports must record classification of report (VAW vs. non-VAW).</p> <p>Caseload should be available in the agency's automated or manual management information system database.</p> <p>The agency's response to increased demand can be measured through management information on hiring, use of overtime, staffing deployment changes, and reorganization.</p>	<p>If improved classification procedures are implemented, the evaluator must be very careful that differences in reporting rates are truly due to changes in public reporting rather than changes in dispatcher classification procedures.</p>

Table 9.2: Impacts on Prosecution Offices

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Tougher case filing practices should be followed. This includes filing charges in more cases referred from law enforcement, and filing more charges, filing charges at higher severity levels, and filing charges which clearly indicate a crime of DV or SA (laws permitting). Prosecutors should seek out information on the perpetrator’s history and use it in charging decisions.</p>	<p>Percent of cases referred from law enforcement in which charges are filed vs. dropped vs. other disposition.</p> <p>For cases with charges filed, average number of charges and severity of the charges (i.e., felony vs. misdemeanor, class A vs. B, or other indicators of charge severity).</p> <p>Percent of cases referred from law enforcement as VAW which are charged with offenses not recognizable as DV or SA (e.g., disturbing the peace or drunk and disorderly rather than DV assault, assault or indecent liberties rather than sexual assault).</p> <p>Percent of cases in which prosecutors obtained information on the perpetrator’s history of protection orders, arrests, court cases, probation/parole status, firearms violations, and so on, and whether more cases with serious histories were charged more severely (compared with cases with comparable current allegations but less serious histories).</p>	<p>Automated office database should contain, for each case, whether and what charges were filed. If there is no automated database, this information should be available in case files. Whatever the source, charges should be specified as VAW-related or not VAW-related at each stage of the process, and this classification must be retained at each stage so changes in status can be noted. Prosecutors could be interviewed to determine their use of perpetrator’s history information in charging decisions.</p>	<p>Filing practices are more likely to be improved where law enforcement agencies have strong evidence collection and case preparation policies and practices.</p> <p>In assessing the severity of the charges filed, it is important to understand local conditions and policies since filing higher-level charges may not always be the best strategy for obtaining a conviction or a harsher penalty. For example, in some jurisdictions stronger penalties may be more likely by filing misdemeanor than felony charges, due to local court practices or other factors.</p> <p>For states which do not have special codes for DV or SA crimes, charge severity must be used as the measure of “charging down” at the filing stage.</p>

Table 9.2: Impacts on Prosecution Offices

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>More indictments should be sought and obtained, on more and higher charges. When guilty pleas are negotiated, less “down-charging” should occur. Requirements to attend treatment should still be accompanied by sentencing (charges should not be dropped).</p>	<p>Percent of cases brought forward for indictment vs. dropped vs. guilty plea reached pre-indictment.</p> <p>For cases brought to the grand jury, percent on which an indictment is obtained.</p> <p>Differences in the number, severity, and type (VAW vs. non-VAW) of charges brought for indictment vs. charges on which an indictment is handed down.</p> <p>Differences in the number, severity, and type of charges to which a guilty plea is entered vs. charges filed, when a guilty plea is negotiated.</p>	<p>Automated agency database should contain, for each case, whether an indictment was sought, a guilty plea entered, or the charges were dropped. With an indictment or plea, the database should include charges in the indictment or to which the defendant pled guilty. When an indictment was sought, information should also be included on the grand jury’s decision. If there is no adequate automated database, this information should be available in case files.</p>	
<p>Practices for setting the amount and conditions of bail should involve a systematic risk assessment, including risk of both elopement and further violence, and should provide more protections for victims.</p>	<p>Percent of cases in which a systematic risk assessment is conducted, and the nature of that assessment.</p> <p>For high risk cases, the number of special conditions imposed to protect victims, such as stay-away orders, surveillance conditions, participation in treatment for violence and substance abuse if appropriate, and the like.</p>	<p>Interviews with prosecutors and pretrial services staff can identify the type and use of risk assessment procedures. The automated agency database (or hard copy case files) can be reviewed to determine whether cases identified as high risk are more likely to receive special conditions (compared with lower-risk cases and/or high-risk cases from another jurisdiction or before risk assessment procedures were instituted).</p>	

Table 9.2: Impacts on Prosecution Offices

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Indicted cases are prosecuted more vigorously. Prosecutors make more use of alternative evidence presentation strategies (besides victim testimony); charges are not charged down as much in the plea negotiation process; and fewer cases are dropped.</p>	<p>Percent of prosecutors willing and able to pursue cases with evidence other than victim testimony when the victim is reluctant (such as admissible hearsay evidence, forensic evidence, witness testimony, expert testimony, etc.).</p> <p>For cases in which a guilty plea is reached through negotiation, differences in the number and severity of charges to which the defendant pleads guilty vs. indicted charges. Percent of indicted cases in which the charges are subsequently dropped.</p>	<p>Measure prosecutors' willingness and ability to use alternative evidence presentation strategies through interviews with prosecutors to determine when they use such strategies and what the barriers to using these strategies are.</p> <p>Measure "down-charging" and percent of indicted cases which are dropped through automated agency database or, if none available, reviews of case files.</p>	
<p>Stronger coordination mechanisms are used. Prosecutors use practices that ensure integration of cases across civil and criminal courts, and across misdemeanor and felony offenses. State and local prosecutors coordinate with U.S. Attorneys and federal agencies to ensure prosecution under the most appropriate jurisdiction. Mechanisms should also be used to facilitate cross-jurisdictional enforcement of protection orders.</p>	<p>Percent of cases with involvement in more than one court in which a single prosecutor handles all cases or prosecutors closely coordinate their work.</p> <p>Percent of cases in which prosecutor obtains and makes case decisions based on information about all relevant cases regardless of which court they appear in.</p> <p>Provisions of working agreements between prosecutorial offices and other relevant federal agencies (such as the Bureau of Alcohol, Tobacco, and Firearms, and the Federal Bureau of Investigation), and how well these links function.</p> <p>Usefulness of policies, legislation, databases, and other resources to enhance enforcement of civil or criminal protection orders from other states, counties, cities, territories, or tribal areas.</p>	<p>Searches of case documentation, either electronic or manual. Interviews with state, local, and federal prosecutors, and investigators in federal agencies, to assess working relationships and areas for improvements. Interviews with law enforcement, prosecutors, and court administrators and support staff to assess the infrastructure supporting cross-jurisdiction protection order enforcement, and areas for improvement.</p>	

Table 9.2: Impacts on Prosecution Offices

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Better case outcomes: more convictions are obtained, on more and higher charges. Method of case disposition may or may not change. Case processing should become more efficient, with expedited docketing, fewer continuances, and timely victim notification of all phases of the case.</p>	<p>Percent of indicted cases with each type of disposition: conviction by guilty plea, conviction by trial, acquittal, diversion or suspension, or charges dropped.</p> <p>For cases with guilty plea or verdict, number and severity of charges of conviction.</p> <p>Time from case filing to disposition, number of continuances, and number and timing of victim notification contacts.</p>	<p>Measure case outcomes through an automated agency database which should contain, for each case, date and type of disposition for each charge (also need date of initial case filing to compute case processing time). If there is no adequate automated database, the disposition information should be available in case files.</p>	<p>Local practices and conditions should be very carefully examined to determine what are desirable disposition types and favorable outcomes. For example, plea bargaining can be a very useful tool if cases are not bargained down to minor offenses with light penalties that do not protect victims. Faster case processing time may be desirable from the system's perspective but must not be too fast to deliver justice and protection to victims.</p>
<p>Improve assistance to victims by prosecutors, investigators, paralegals, victim/witness staff, and other personnel.</p>	<p>Percent of victims who are offered assistance such as notifications of rights, service referrals, case updates, information on how the court system works, and support for participation in the process.</p> <p>For victims offered each type of assistance, percent who received or did not receive it, and reasons why not (e.g., victim refused, assistance not immediately available, victim not eligible for assistance, etc.).</p> <p>Can also measure the number of assistance offers over time, the length of time over which the victim receives offers, and offers to victims who are not able to testify.</p> <p>See also Chapter 7, proximate measures, for victim feedback on agency performance; Chapter 8 for measures of victim services.</p>	<p>Victim assistance can be assessed through victim surveys and reviews of case files documenting assistance offered and whether victims accepted or declined. Victim surveys should also identify types of help that victims needed but were not offered.</p> <p>In addition, process analysis should use semi-structured interviews with advocates and other knowledgeable people to identify problems, barriers, and gaps in assistance as well as improvements and successes in serving victims.</p>	

Table 9.2: Impacts on Prosecution Offices

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Prosecutors' offices should have more and better resources to respond to cases of VAW.</p>	<p>Percent of offices with thorough, progressive written policies on VAW cases.</p> <p>Percent of office personnel who receive training in a timely fashion when new policies are implemented.</p> <p>Percent of offices with special units on DV, SA, or both. These units may consist of prosecutors, paralegals, victim/witness counselors, investigators, or other personnel.</p> <p>Special units' caseloads and the average rated value of the unit's contributions to the agency.</p> <p>Number of prosecution staff with whom victim must interact (fewer is better). Consistent responses and information about case from all prosecution staff.</p>	<p>Assess thoroughness and progressiveness of policy by content analysis. Policies should address a variety of important current topics.</p> <p>Assess training by review of training records for all office personnel.</p> <p>Assess special units and the number and types of personnel through interviews with managers, reviews of office policies, organizational charts, annual reports, personnel files, etc.</p> <p>Assess special unit's caseload through agency database (or case file review). Measure the unit's contributions to the agency by interviews with staff inside and outside the unit.</p> <p>Assess by reviews of case records, and by interviews with victims.</p>	<p>Important policy topic areas may be determined in part by current issues in the state or community, and may include aggressive prosecution practices, caseload management, special court structures, providing services to victims, coordinating with other agencies, etc.</p>

Table 9.2: Impacts on Prosecution Offices

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Prosecutors and other agency staff should be culturally competent and have culturally appropriate resources available to them.</p>	<p>Percent of staff trained in cultural competence. Number and types of cultural competence resource materials available (e.g., information on cultural beliefs or norms about women and violence, for populations found in the jurisdiction).</p> <p>Staff demographic profile compared to community demographic profile.</p> <p>Percent of victims with language barriers for whom an appropriate interpreter is provided.</p> <p>Number and types of forms and brochures available in languages spoken by the population of the jurisdiction.</p>	<p>Examine agency training records and resource materials.</p> <p>Compare agency personnel records with census data on community population.</p> <p>Examine case records, and also records of use of interpreter service, if available. Use personnel records to assess staff language competencies. Examine forms and informational materials available to the public.</p>	
<p>Staff should be well-trained in VAW cases.</p>	<p>Percent of staff who meet minimum requirements for VAW training, as specified by the state prosecutors' association, or by state or local law, policy, regulation, etc.</p> <p>Percent of staff in special units who have advanced training in DV, SA, or both, depending on type of unit.</p>	<p>Measure by review of personnel records or records of the training authority.</p>	
<p>Prosecution office's caseload should increase because of these improvements and/or improvements in law enforcement case preparation and referrals. Having more cases and prosecuting them more vigorously will lead to the need for additional staff, staff overtime, or redeployment of staff.</p>	<p>Caseload = the average number of cases per staff and the average amount of time put into each case.</p> <p>Increased demand arises from increases in staff caseload and time spent per case, over time.</p>	<p>Caseload should be available in the office's automated or manual management information system.</p> <p>The office's response to increased demand can be measured through management information on hiring, use of overtime, staffing deployment changes, and reorganization.</p>	

Table 9.3: Impact on Judges and Courts

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Court policies should make protections such as civil orders more available by streamlining and simplifying forms and procedures, providing more and better assistance, and removing barriers such as fees.</p>	<p>Number and types of changes in forms and procedures courts have made, and average ratings of usefulness of changes. Percent of victims who receive and court personnel who offer assistance in filling out court forms. Types of help given and average ratings of helpfulness, and types of help still needed. Number and types of barriers to court protection which have been removed.</p> <p>Percent of protection orders issued for which fees were waived for issuing the order; for serving the order.</p> <p>Existence of a protection/restraining/stay away order registry, and percent of all orders, including emergency orders, that become entries on the registry.</p>	<p>Changes in court forms and procedures, and removal of barriers, can be assessed by interviewing court personnel and reviewing forms and policies for changes. Interviews with staff who work in the court setting (whether court employees or private victim service agency staff) and surveys of victims can provide information on how helpful changes in forms and procedures have been, how many victims receive assistance when needed, who offers assistance, what type of assistance is available, how helpful the assistance is, and what needs are left unaddressed.</p>	
<p>Judges and other court personnel should have access to, and should use, information from other court cases involving the same victim and/or perpetrator in making sentencing decisions.</p>	<p>Number of cases where victim and/or perpetrator are involved in related cases in other courts for which judge has access to information about the other cases; number of cases in which the judge uses that information to make sentencing decisions.</p>	<p>Review of case records; assessment of electronic or other mechanisms for acquiring relevant information from other cases; interviews with judges and staff conducting pre-sentence investigations.</p>	

Table 9.3: Impact on Judges and Courts

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Case activity should increase as a result of greater case activity by law enforcement, prosecution, or corrections, and/or from new court policies or procedures making it easier to obtain court protection.</p> <p>Courts should issue more warrants (e.g., arrest warrants, search warrants), subpoenas, and orders (e.g., protection orders) to facilitate case investigation and prosecution, and to protect victims.</p> <p>Courts should hold more hearings and impose more penalties in response to reported violations of civil and criminal court orders and conditions of diversion, bail, probation, or parole.</p>	<p>Measures may include the number of such court actions, the percent of cases with such actions, or the percent of requests which are affirmed, as appropriate.</p> <p>Measures may include the number of hearings, the percent of cases with hearings, the percent of hearings which result in a finding of a violation, and the severity of penalties imposed on violations.</p>	<p>Automated court database should contain, for each case, information on warrants, subpoenas, and orders requested from the court, and result of the request (issued or denied). Violation reports, hearings, findings, and actions should also be available. If no adequate automated system exists, this information should be available in case files.</p>	
<p>Judges should make court orders more responsive to victims' situations and needs, and provide better protection and justice to victims. For example, there should be fewer mutual protection, restraining, or no-contact orders, and protection and other orders (such as separation, divorce, or custody decrees) should contain more provisions needed by victims. Adequate procedures for serving orders and documenting proof of service should be followed.</p>	<p>Percent of cases without mutual orders, the typical provisions of orders, service and documentation procedures, and the percent of victims who are satisfied with the provisions of their orders.</p>	<p>Surveys of victims and/or advocates can assess what orders are issued, what the provisions of the orders are, whether these provisions address the victims' circumstances and need for safety and justice, and how the service process worked. Reviews of automated or manual court files can also assess what orders are written and their provisions and service procedures.</p>	

Table 9.3: Impact on Judges and Courts

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Judges and other court personnel, including those doing pre-trial and pre-sentence investigations, should give victims' needs and input more consideration.</p> <p>Better protection should be provided for victims and witnesses in the courthouse.</p> <p>More assistance should be offered to facilitate victims' involvement with the court, such as court schools, transportation and food allotments, and child care.</p>	<p>Percent of judges, magistrates, and other decision-makers who give victims' input more consideration when deciding whether to issue warrants, impose conditions or sentences, and so on.</p> <p>Percent of courts that have metal detectors and guards at the doors, escort services, and private secure waiting rooms for victims and witnesses.</p> <p>Percent of court proceedings in DV or SA cases held in courtrooms with safeguards.</p> <p>Percent of courts that offer such services, and percent of victims who avail themselves of the services. Victims' assessments of the value of this assistance and areas for improvement.</p>	<p>Interviews with judges and other decision-makers can be done to assess how they consider the victims' input in decision-making, and what is needed to enhance the victim's role.</p> <p>Courthouse security can be assessed through inspection of court facilities; interviews with security personnel to assess safeguards in use and still needed; interviews with court personnel, victims, and witnesses to measure perceived safety; and review of reports of violent incidents maintained by court security personnel. Court database or case file reviews should indicate which proceedings were held in secure settings. Victim assistance activities can be assessed by reviews of court records on assistance activities, and interviews with victims to get their feedback.</p>	
<p>Judges should allow and support increased court use of appropriate evidence and testimony.</p>	<p>Percent of cases in which evidence is presented from DNA testing, polygraph tests for sex offenders, and/or use of expert witness testimony about rape trauma syndrome, psychological reactions to violence, recovery patterns, and belief structures (myths and attitudes).</p> <p>Also, the percent of cases in which rape shield laws are applied appropriately.</p>	<p>Automated court database or court files should document the evidence presented in proceedings.</p>	

Table 9.3: Impact on Judges and Courts

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Judges' use of sentencing practices and conditions of diversion, suspension, or other alternative methods of case processing should reflect greater efforts to hold offenders accountable and provide protection to victims. Diversion or suspension should be used only in ways that serve victims' interests (i.e., <i>not</i> for the convenience of the justice system when that conflicts with victim's interests).</p>	<p>For convictions, assess percent of cases with fine, jail time, probation time, community service requirements, and other conditions such as restitution or no-contact orders.</p> <p>Can also compute average fine and restitution amount and average probation, jail, or community service time.</p> <p>Assess percent of cases in which the victim is given the opportunity to adduce testimony on the <u>future risk</u> posed by the defendant, in addition to the impact of past behavior, and the use of this information in sentencing decisions.</p> <p>For diverted or suspended cases, assess conditions of diversion/suspension and how they meet victims' needs for justice and safety.</p> <p>Percent of offenders who are ordered to batterer intervention, substance abuse treatment, sex offender treatment, and other appropriate services.</p> <p>Percent of offenders ordered to specialized monitoring, such as electronic surveillance, intensive supervised probation, special victim safety check and/or notification requirements, and polygraph tests for sex offenders under supervision.</p> <p>Percent of cases for which offender has post-sentencing interaction with a judge, for monitoring. Frequency of such interactions.</p>	<p>Information on sentencing, victims' testimony, and conditions of probation, parole, and diversion/suspension (including orders for treatment and specialized monitoring) should be available through prosecutor's or court's automated database. Judges and victims can be interviewed for their perceptions of the impact of the victims' testimony on sentencing decisions.</p> <p>Measure whether sentencing and diversion/suspension practices respond to victims' needs through interviews with prosecutors and victims and/or victim advocates.</p>	<p>Exactly what "victim's needs" means might be different for each case, but in general court-ordered conditions should try to provide for victim safety from further assault or abuse and financial well-being (e.g., orders for restitution or separate maintenance) in so far as possible. Rather than using a set list of court-imposed conditions, it might be most useful to conduct more open-ended interviews and ask victims/advocates what they wanted, what they got, how it worked out, and what they still needed but didn't get.</p>

Table 9.3: Impact on Judges and Courts

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Courts offices should have more and better resources to respond to cases of VAW.</p>	<p>Percent of courts which have thorough, progressive written policies and/or bench guides on VAW cases.</p> <p>Percent of court personnel (judges, magistrates, clerks, bailiffs, and others) who receive training on new policies, and how soon after policy implementation the training begins.</p> <p>Percent of courts with an adequate number of research and writing staff to prepare legal memoranda on substantive, evidentiary, and procedural issues so that judges are fully apprised of legal precedent and public policy developments.</p> <p>Percent of jurisdictions with unified court systems or other ways of facilitating coordination of the different courts which may be involved in VAW cases (e.g., family, civil, and criminal).</p> <p>Percent of jurisdictions with specialized courts, specialized calendars, or specialized judicial assignments. Specialized court's caseload and the average rated value of the court's contributions to the jurisdiction, according to interviews with personnel on the specialized court and others in the jurisdiction.</p> <p>See also Chapter 7, proximate measures, for victim feedback on agency performance.</p>	<p>Assess thoroughness and progressiveness of policy by content analysis. Policies should address a variety of important current topics.</p> <p>Assess training by review of training records for all court personnel.</p> <p>Interview legal staff and judges to assess their the sufficiency of time available for qualified staff to prepare legal briefings, and the usefulness of their products.</p> <p>Assess cross-court coordination mechanisms through interviews with court personnel and review of policies, memoranda, shared information systems, and other documentation.</p> <p>Assess specialized courts and the number and types of personnel through interviews with managers and reviews of office policies, organizational charts, annual reports, personnel files, etc. Assess special unit's caseload through agency database (or case file review).</p>	<p>Important policy topic areas may be determined in part by current issues in the state or community, and may include such topics as cross-court coordination, working with other agencies to supervise compliance with court orders, etc.</p>

Table 9.3: Impact on Judges and Courts

<p>Objectives</p> <p>Judges and court staff should be culturally competent and have culturally appropriate resources available to them.</p>	<p>Specific Measures</p> <p>Percent of staff trained in cultural competence.</p> <p>Number and types of cultural competence resource materials available (e.g., information on cultural beliefs or norms about women and violence, for populations found in the jurisdiction).</p> <p>Staff demographic profile compared to community demographic profile.</p> <p>Percent of victims with language barriers for whom an appropriate interpreter is provided.</p> <p>Number and types of forms and brochures available in languages spoken by the population of the jurisdiction.</p>	<p>Data Collection Procedures</p> <p>Examine agency training records and resource materials.</p> <p>Compare agency personnel records with census data on community population.</p> <p>Examine case records, and also records of use of interpreter service, if available. Use personnel records to assess staff language competencies. Examine forms and informational materials available to the public.</p>	<p>Precautions</p>
<p>Court personnel should be well-trained in VAW cases.</p>	<p>Percent of staff who meet minimum requirements for VAW training, as specified by the state judges' association, or by state or local law, policy, regulation, etc.</p> <p>Percent of staff in specialized courts who have advanced training in DV, SA, or both, depending on type of court.</p>	<p>Measure by review of personnel records or records of the training authority.</p>	
<p>Court caseloads should increase because of court improvements and/or improvements in law enforcement and prosecution practices. Heavier caseloads will lead to the need for additional staff, staff overtime, court restructuring, or redeployment of staff.</p>	<p>Caseload can be measured by the number of cases, number of court appearances or actions per case, and case processing time.</p> <p>Increased demand arises from increases in number of cases and time spent per case, over time.</p>	<p>Caseload should be available in the court's automated or manual management information system database.</p> <p>The court's response to increased demand can be measured through management information on hiring, use of overtime, staffing deployment changes, and reorganization.</p>	

Table 9.4: Impact on Corrections (Probation and Parole)

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>More proactive assessment of supervision needs and efforts to ensure appropriate supervision is ordered.</p>	<p>Number of agencies that systematically assess supervision needs, and nature of the assessment procedures. When level or type of court-ordered supervision appears inappropriate, percent of cases in which the corrections agency either modifies conditions accordingly, or asks the court to modify its supervision orders (in jurisdictions where this is required).</p>	<p>Interview correctional offices to assess the nature and use of assessment procedures. An automated agency database or case file records should document modifications indicated by assessment results, and whether the supervising agency modified supervision accordingly or requested the court to do so.</p>	
<p>Offender monitoring should increase. Probation and parole officers should directly monitor offenders more closely and have more contact with other community agencies involved in the case to monitor offender compliance and victim safety. For example, it may be necessary to check with law enforcement to verify that a sex offender has registered as required and to be notified of rearrests; to communicate with batterer treatment providers to monitor compliance with court orders to treatment; and to communicate with victims and/or advocacy/service providers to detect repeat abuse and notify victims of offender status changes (e.g., release, furlough, escape, the opportunity to participate in reviews or hearings, and so on).</p>	<p>Direct offender monitoring = average number of contacts with offender per case, average length of time over which contacts occur in a case, and types of contacts most frequently made (e.g., visits to offender’s home or workplace, visits by offender to officer’s office, etc.).</p> <p>Compliance monitoring = average number of contacts with victims and other agencies involved in the case.</p>	<p>Interview probation and parole officers and staff of other relevant agencies, and review automated database of probation or parole agency, if available, or case files to document contacts with offenders and other agencies. Interview offenders to assess their perceptions of the monitoring process and what additional steps could be taken to fill gaps.</p>	

Table 9.4: Impact on Corrections (Probation and Parole)

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Closer communication with the courts on offender compliance or violations. More violation reports to the court, as a result of enhanced monitoring and/or stronger policies on reporting violations.</p>	<p>Communication with the court = average number of reports to the court on compliance or non-compliance with conditions of probation or parole. These may involve periodic court reviews or special reports issued in response to violations.</p> <p>Violation reports = average number of violation reports to the court.</p>	<p>Probation, parole, or court automated database if available; probation or parole officers' case files; and/or court case files.</p>	
<p>Supervision of offenders should improve: more referrals to treatment programs and other services needed or ordered by the court, more assistance with obtaining services (e.g., assistance with fees, transportation or other accessibility issues, waiting list priority, etc.), and greater service availability in prisons and jails.</p>	<p>Offender service referrals = average number of service referrals, types of referrals most frequently made (to batterer treatment programs, substance abuse treatment, etc.).</p> <p>Referral assistance = average number of assistance efforts given, types of assistance most frequently given.</p> <p>Service availability = number and types of service programs available in prisons and jails; Accessibility = number of inmates per program opening.</p>	<p>Interview probation and parole officers and offenders to assess factors which facilitate and hinder offenders from obtaining services, how the officers helped, and other interventions needed to make services more accessible. Review officers' records or agency database for information on referrals made and assistance provided.</p> <p>Interview prison and jail managers, counselors and other service program staff, and inmates to identify service programs and accessibility issues. Review management records on program activities and attendance.</p>	

Table 9.4: Impact on Corrections (Probation and Parole)

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Probation and parole should have more and better resources to respond to cases of VAW.</p>	<p>Percent of agencies with thorough, progressive written policies on VAW cases.</p> <p>Percent of agency personnel who receive training in a timely fashion when new policies are implemented.</p> <p>Percent of agencies with specialized staff (such as victim liaisons) or special units on DV, SA, or both.</p> <p>Special staff or unit caseloads and the average rated value of their contributions to the agency.</p> <p>Percent of cases for which corrections officers track and monitor offenders <i>after</i> probation and parole ends.</p>	<p>Assess thoroughness and progressiveness of policy by content analysis. Policies should address a variety of important current topics.</p> <p>Assess training by review of training records for all agency personnel.</p> <p>Assess special staff or units and the number and types of personnel through interviews with managers, reviews of office policies, organizational charts, annual reports, personnel files, etc.</p> <p>Assess special staff or unit's caseload through agency database (or case file review). Measure the staff or unit's contributions to the agency by interviews with staff inside and outside the unit.</p> <p>Assess by review of case records.</p>	<p>Important policy topic areas may be determined in part by current issues in the state or community, and may include models of intensive supervision, monitoring victim safety, etc.</p>

Table 9.4: Impact on Corrections (Probation and Parole)

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Staff should be well-trained in VAW cases.</p>	<p>Percent of staff who meet minimum requirements for VAW training, as specified by the state correctional officers' association, or by state or local law, policy, regulation, etc. Percent of staff in special units who have advanced training in DV, SA, or both, depending on type of unit.</p> <p>Percent of staff with explicit understandings gained in training (e.g., what constitutes a violation of a PO).</p> <p>Percent of offenders "violated" by probation or parole officer for different offenses, before and after training.</p>	<p>Measure by review of personnel records or records of the training authority.</p>	
<p>Correctional agencies' caseload should increase because of these improvements and/or improvements in prosecution and the courts. Having more cases and monitoring them more closely will lead to the need for additional staff, staff overtime, or redeployment of staff.</p>	<p>Caseload = the average number of cases per staff and the average amount of time put into each case.</p> <p>Increased demand arises from increases in staff caseload and time spent per case over time.</p>	<p>Caseload should be available in the agency's automated or manual management information system.</p> <p>The agency's response to increased demand can be measured through management information on hiring, use of overtime, staffing deployment changes, and reorganization.</p>	

Table 9.4: Impact on Corrections (Probation and Parole)

Objectives	Specific Measures	Data Collection Procedures	Precautions
<p>Corrections staff should be culturally competent and have culturally appropriate resources available to them.</p>	<p>Percent of staff trained in cultural competence. Number and types of cultural competence resource materials available (e.g., information on cultural beliefs or norms about women and violence, for populations found in the jurisdiction).</p> <p>Staff demographic profile compared to community demographic profile.</p> <p>Percent of offenders with language barriers for whom an appropriate interpreter is provided. Number and types of forms and brochures available in languages spoken by the population of the jurisdiction.</p>	<p>Examine agency training records and resource materials.</p> <p>Compare agency personnel records with census data on community population.</p> <p>Examine case records, and also records of use of interpreter service, if available. Use personnel records to assess staff language competencies. Examine forms and informational materials available to the public.</p>	

CHAPTER 10

MEASURES OF COMMUNITY COLLABORATION

Many individuals and organizations involved in addressing violence against women have heard about collaborative community-wide responses to ending such violence and in many communities across the country, efforts are under way to establish such systems. But what does coordinated community response really look like, and how do you know it when you see it? Furthermore, what is coordinated community response meant to achieve, and how does one measure this?

To most observers, the term “community-wide collaboration” conveys the fairly obvious, if somewhat vague, notion of different members of a community coming together to address a common problem. Clearly, there is more to community collaboration—or at least successful collaboration—than this. Christiansen et al. (1997)⁷ characterize collaboration as involving interdependence and ongoing give and take; solutions emerging through participants’ dealing constructively with differences; partners working beyond stereotypes to rethink their views about each other; and stakeholders assuming joint ownership of decisions and collective responsibility for future decisions. They believe that collaboration is an ongoing process, not something done once and finished. Partners in collaboration are always looking at how things are going, and talking through how to keep improving them.

It is important to keep in mind that there is no single ideal of community collaboration. What makes sense for one community may not work in another. There are, however, a variety of scenarios with common elements that appear to be effective. These elements or factors can provide us with models against which we can measure efforts in most communities.

The remaining sections of this chapter identify (1) the basic elements needed to start the process of community collaboration, (2) intermediate and system level outcomes associated with successfully establishing community collaboration, and (3) the ultimate outcomes of such collaborations. As noted above, collaboration is a never-ending process. Each success builds on the last one. This means that the process is circular—commitment (the first essential ingredient for a collaborative community-wide response) leads to system changes that lead to renewed commitment, etc. Similarly, each system change that is accomplished feeds a willingness to increase commitment for the next change. This complexity reflects the dynamic and fluid nature of establishing a community-wide collaborative response to violence against women.

It should also be noted that the greater the number (and longevity) of factors present in a given community, the more likely it is that the collaborative effort is truly entrenched within that community and that the effort will survive the withdrawal of a major source of financial support such as VAWA STOP grants. We have provided three tables (Tables 10.1, 10.2, and 10.3), each focused on one possible goal of a collaborative effort. Each table provides specific objectives

⁷ Helen Christiansen, Linda Goulet, Caroline Krentz, and Mhairi Maeers (Eds.). (1997). *Recreating Relationships: Collaboration and Educational Reform*, New York: State University of New York Press.

under each goal, measures of these objectives, and recommended data collection procedures. Please keep in mind that this is not an exhaustive list. Those who are closest to the communities and are actually assessing the collaborations may have other measures in mind and/or have a better idea of what is realistic in terms of data collection.

ELEMENTS OF COMMUNITY-LEVEL COLLABORATION

Commitment

The foundation of any type of community collaboration is commitment. This includes an understanding that (1) violence against women is a community problem that must be addressed by the community; (2) community-wide approaches are necessary to provide all victims with a full range of options and services free from unnecessary hassles and difficulties, to hold perpetrators accountable for their actions, and to create public confidence and trust in community institutions; and (3) changes are (almost certainly) necessary in all service units separately, and in the way they interact with each other, to promote excellent care and continuity for victims and to hold perpetrators accountable.

Organizing Structure

Coordinating bodies are a critical part of a community's efforts, providing a forum for identifying problems and developing solutions on an ongoing basis. When faced with the challenge of developing an integrated community system, it is important that the organizing body have some legitimized power or authority to make recommendations and/or decisions. This means that the committee or group is, at least, the recognized structure for interunit communication, for feedback about one's own unit's performances, for establishing priorities among the community-wide tasks to be done, and for public outreach and dissemination about its goals and activities.

Organizing structures can take many forms, but some of the most common are: a committee or task force; a nonprofit, nongovernmental body that assumes coordinating responsibility; or a small, informal network of two or more agencies that plan and implement activities together. As Table 10.1 illustrates, in addition to noting the structure of organizations, it is also useful to gather information such as the length of and reasons for their existence.

Committees. A coordinating committee or task force may employ a coordinator—full- or part-time—whose primary responsibilities are to organize all efforts of the committee, serve as the central contact for its members, and facilitate communication among members. If funding is not available for a full-time position, a suitable member from the community may serve as coordinator on a part-time or voluntary basis. In addition to aiding the work of the task force, the presence of a coordinator may help to ensure continuity in a politically changing environment. Committees (and their subcommittees) vary greatly in their frequency of meeting, generally ranging from once

a month to one or two times a year. Meetings provide a forum for updating members on progress toward set goals and the development of new issues.

Nonprofit Organizations. Another possible structure is a nonprofit organization that takes responsibility for some or all of the community's collaboration efforts. Possible scenarios include an organization that was created for the sole purpose of coordinating the community response or an already-existing organization that takes control of one or two specific projects.

Others. In communities where there is no formal coordinating structure, a number of agencies representing different areas (e.g., nongovernmental victim services and law enforcement agencies) may choose to work together to develop collaborative projects. In others, there may be more than one coordinating mechanism.

Formation of Additional Structures. In some cases, new VAW coordinating structures appear. They are formed for various reasons, which may include the realization that certain needs are not being met by the initial coordinating body, disagreement over objectives or means of achieving them, or simply the need to extend the reach of existing structures.

Composition of Committee/Group

Diverse Membership. Diversity of membership is important because it is either difficult or impossible for one person or agency to solve a problem as multi-faceted as violence against women. The challenge of collaboration is to blend the backgrounds of people from several agencies into workable relationships that are mutually beneficial. As noted in Table 10.1, the types of participants in a collaborative effort may vary depending on the specific goals of the community response, but should consist of representatives from different sectors, possibly including: battered women's and sexual assault services, law enforcement, prosecution, civil and criminal courts, health care agencies, child welfare agencies, batterer intervention programs, the corporate/business sector, clergy and religious institutions, and other opinion leaders.

It may also be important to have a combination of experienced/founding members and new members, since older members may be inclined to move on to other projects. The presence of new members will also indicate that there is continued community support and interest in the committee's efforts.

Continuity of Personnel. One of the keys to a successful collaboration is having some or all of the same individuals involved throughout the entire effort. Collaborations improve over time when participants become familiar with each other's strengths and weaknesses and learn to work as a team. Personnel changes occur for a variety of reasons, including work reassignment, termination, promotion, retirement, and, in some cases, a decision not to participate. Any disruption can have adverse effects. At a minimum, changes delay planned activities, and in the worst case the collaboration may dissolve.

Especially vulnerable are collaborations in which a dynamic leader is no longer available—what might be termed the “vanishing advocate” problem. Some collaborations are formed because an individual has taken a problem on and has forged a coalition to seek solutions. Without that leader, the effort folds.

Level of Participating Staff. Some successful groups comprise primarily high-level staff, while others achieve their goals with broad-based community and grassroots support. Whatever the level of participating staff, it is important to consider whether they are in a position within their agency or community that enables them to make changes and influence policy. They should also have the backing or commitment of their own agency to cooperate with the community system. Another factor that should be noted is whether or not those who are in power in the relevant agencies are supportive of ongoing collaborative efforts. (Suggestions for possible ways to collect these data are mentioned in Table 10.1.) For those found to be critical or ambivalent toward the group’s efforts, an important initial goal should be to get them on board.

Activity Level/Involvement of Participants. In some groups, the bulk of the work is done by one or two powerful members, while in others it is shared by many. Measures used to gauge the level of member activity might include the number of tasks initiated and completed by each member, the proportion of the member’s overall time that he/she dedicates to collaborative efforts, and how many meetings he/she attends.

Another issue is the degree to which participants are *allowed* to become involved. For example, one might ask who in the group is leading activities and making important decisions and whether or not all members have the opportunity to participate equally.

Levels/Sources of Support

Some collaborative VAW efforts enjoy support from all levels of government, while others are forced to work from the ground up, creating grassroots networks to inform elected and appointed officials and lobby for change. Some initiatives, for example, are formed when states pass laws affecting violence against women or governors or mayors make this issue a priority in their administrations. In other situations, the occurrence of a particularly tragic event or overwhelming frustration with the system prompts service agencies and/or victims to organize a community response from the grassroots level.

Both levels of support are equally important—albeit for different reasons—and should be measured. A high-level elected official such as a governor can help make a collaboration strong by directing state funds and legislative support toward the collaboration’s activities. This also happens at a local level when a mayor or other city official uses a portion of the city’s funds for this purpose. But due to the changeable political climate, collaborations should not become solely dependent on any one source of support. If they enjoy widespread public support, their activities will continue even after the loss of a strong political ally. For instance, even though grassroots

organizations generally are unable to provide the steady funding streams that the government can, they may offer other benefits to the collaboration, such as media attention and lobbying efforts.

COMMUNITY COLLABORATION: SYSTEM-LEVEL OUTCOMES

Communicating Effectively

Among Participants. Having *any* communication among agencies where there was none, more communication where some already existed, or improved or more effective communication that leads to greater understanding and trust are all indicators of an established community response. Communication may be measured based on frequency, type, level, basis, and nature of contact, as mentioned in Table 10.2.

While regular meetings are one way of keeping participants informed, it is also important to create and maintain connections among individuals. Informal dialogues, memoranda of understanding, reports, and e-mail transmissions can also contribute to effective communications. Ultimately, better communication among agencies and sectors may evolve into arrangements (both formal and informal) for different types of agencies or agencies in different geographic areas to share administrative or management information system databases.

Informal Networks. In addition to communication among formal committee/group members, participants should create their own informal communication networks. For example, even if the coordinating body is composed primarily of criminal justice members, there should be *connections* established with other community groups. Although a little more difficult to measure than formal networks, these connections are important because even if the coordinating body disbands or loses its coordinator, many of the connections will remain. One example would be to hold a monthly lunch with a brief program presented by a different agency each time.

Feedback from Community. Groups may or may not incorporate community feedback in their efforts. Two possible areas to measure noted in Table 10.2 are (1) whether meetings and materials are available to the general community and if so, how often, and (2) the frequency with which community feedback is incorporated into decisions and/or projects.

Developing a Shared Vision

The myriad agencies/organizations involved in addressing VAW may have different views on one or more of the following issues: fundamental principles of intervention, roles of each component, merits of collaboration, necessity of public accountability, and goals of reform. Despite differences of opinion, they should be able to develop a common vision and mission. It should be acceptable for parties to have different aims, as long as they are not conflicting and fit under an

overarching common purpose. In short, everyone needs to have the same understanding of the overall goals of the effort, and all collaborative activities should be working toward those goals.

Developing a common vision should result in fewer misperceptions about each member's roles/responsibilities. Each person should play a role in the collaborative effort and understand what that role is. Individual committees and task forces should have clearly defined goals and purposes. Other indicators of a shared vision include developing a strategic plan, a mission statement, prioritizing tasks, joint planning, and joint implementation (see Table 10.2 for ways to measure shared vision).

In addition, agencies should produce compatible protocols or guidelines for practice in each component. Services provided by different agencies should be as seamless as possible. For example, police and victim advocates may share a jointly developed protocol that specifies (1) that both will go on all calls; (2) once on the scene, the responsibilities of the law enforcement officer and responsibilities of the victim advocates; and (3) the responsibilities of each for follow-up and keeping the victim informed. Another example might be a protocol developed jointly by law enforcement and prosecution, describing the types of evidence police could collect that are of most value to prosecutors and the kind of feedback from prosecutors that would be most valuable to police.

Establishing Systems of Conflict Resolution and Evaluation

Beyond communication, there may be systems or mechanisms in place for the following: identifying problems and developing solutions, evaluating practice efficacy and incorporating community input (public accountability) and other forms of feedback, and monitoring adherence to adopted standards/practices. These systems/mechanisms should be known and be seen as legitimate, reliable, and timely. Table 10.2 suggests various ways to measure these systems.

Developing Trust and Mutual Respect

An important aspect of a successful collaboration is trust and mutual respect among participants. Some collaborations may “click” from the start, especially when the groups have worked together successfully in the past. But this is rare; most take time to build. Collaborative relationships usually move through several early steps: expression of interests, verification of each other, trust building, joint decision making, and small successes. One indication that mutual respect and trust are present is that the group has formed a flexible working arrangement, eliminating the hierarchy that normally characterizes relationships with participating agencies.

Gaining the trust and respect of other group members may itself be a reward for individual participants. When their expertise has been recognized, their ideas have been considered—perhaps even accepted—and their influence has led to solutions to a problem, they may gain a

sense of personal satisfaction. The collaboration itself may then be rewarded by improved performance by the individual.

Engaging in Joint Activities and Co-location

As mentioned in the *Shared Vision* section, a healthy sign of open coordination among different agencies is when they decide to embark on joint activities such as joint training of staff or designing new joint policies or protocols. In some cases, coordinating agencies may decide to co-locate services (e.g., victim service units within police departments and prosecuting attorney's offices). Alternatively, they may decide to respond together as when a victim advocate accompanies police officers on 911 calls. Similar arrangements can be made in other settings such as hospital emergency rooms and courts.

Reporting

Agencies involved in coordination should have a way to report performance, activities, and accomplishments. Ideally, this would be a formal report published on a regular basis. Public information would promote continued support for the initiative from the general public, the private sector, and the government, and may inspire other agencies to undertake similar efforts. As mentioned in Table 10.2, both direct and indirect results of publishing reports should be noted.

Funding

Often community-wide collaboration efforts can lead to changes in various funding mechanisms. In some states or localities, collaborative organizations have succeeded in establishing new sources of funding (e.g., through new marriage license taxes or line items in city budgets). In other cases, collaborative organizations may be given the authority and responsibility for distributing funds within the community, which increases their opportunities to support and foster cross-agency or cross-jurisdictional collaborative projects and other efforts.

COMMUNITY COLLABORATION: ULTIMATE OUTCOMES

Creating Permanent Policy and Practice Changes

One strong indication that a community's collaborative efforts have become entrenched is the existence of permanent policy changes. In some communities, permanent changes in the law and agency policies or procedures may occur as a result of collaboration efforts and will remain even if the original committee disbands or loses steam.

Treating Victims and Perpetrators Consistently

Coordinated community response should lead to more consistent treatment of both victims and perpetrators. The same services and options should be available to victims of all racial, linguistic, geographic, or other characteristics, irrespective of the point at which they enter the system (e.g., a 911 call to the police, an emergency room, child protection, etc.). Similarly, perpetrators should be held accountable in a consistent manner by everyone they come into contact with, such as law enforcement officers, prosecutors, judges, treatment providers, and probation and parole officers. Specific examples of how perpetrators might be treated consistently include: not reducing criminal charges; including participation in batterer and chemical dependency treatment programs and compliance with protection orders in sentences; rigorous enforcement of probation and parole conditions; having systems in place that allow one to track compliance with treatment programs and probation and parole terms; and coordinating terms and conditions with other legal cases (both civil and criminal) involving the perpetrator, victim, or other family members.

Creating More Options and Resources for Victims in the Justice and Human Service Systems

Effective coordinated community response should produce more options and less duplication of services for victims in the justice and human service systems. Because sexual assault and domestic violence victims often face a variety of difficulties and unique circumstances, there should be flexible solutions. For example, domestic violence victims may be offered a choice between staying at a temporary group shelter or receiving financial assistance in order to obtain a more permanent residence. A community may make various types of counseling, including group and individual therapy, available to victims of sexual assault and allow them up to two years to enroll in such services. Standards may be imposed to improve the quality and appropriateness of batterer intervention programs or sex offender treatment programs, and higher quality may lead to higher utilization and more resources. New options should also include more culturally sensitive responses to VAW on the basis of race, language, religion, culture, class, kinship networks, perspectives of the efficacy of the legal process, physical and/or mental disabilities, and urban/rural status.

Changing Public Knowledge of VAW and Reducing Unmet Need

Finally, effective community response should result in an increased level of awareness and understanding of violence against women by the public at large. A variety of public education strategies tailored to differing segments of the population (teens, immigrants, etc.) can be employed and the effects measured by individual surveys on levels of understanding and tolerance about violence against women. Chapter 11 discusses community attitudes toward violence against women further and offers specific tools for measuring these attitudes.

Chapter 10: Measures of Community Collaboration

The general public should also be aware of the services available to victims and their families. Increased knowledge and use of services by those that need them would indicate that the collaborative efforts had reached community members, who may in turn become interested in joining the groups' efforts.

Table 10.1
Community Collaboration, Goal 1:
Establish an Effective, Stable, and Continuing Community Response to VAW

Objectives	Specific Measures	Data Collection Procedures	Caveats
Commitment (Willing to collaborate)	Are agencies and their representatives dedicated to community collaboration? —Do they believe that violence against women is a community-wide problem that needs to be solved by the community? —Are they willing to make changes in their own agencies if necessary in order to further community-wide goals?	Interviews/surveys with key agency staff re. actual and planned participation in collaborative efforts and willingness/ability to change policies/procedures as needed.	
Presence of established formal/informal collaborative structures (Ability to collaborate)	What type(s) of organizational structures has existed in the community (formal task force, interagency lunch group)? How long has it been in place? How often does it meet? What is its focus? How does the structure support collaboration?	Interviews/surveys with actual (and potential) collaborating participants, as well as outside sources. Task force/committee/group agendas, minutes, memos, other documents.	
Presence of new committees or nonprofit organizations (Ability to collaborate)	Types, numbers, and purposes of secondary organizational structures. Reasons for the formation of new group (e.g., none other existed, focus of original group too narrow, membership of original group restricted, etc.). Activities of organizations and areas they cover (e.g., health services, law enforcement, etc.).	Interviews with original committee members. Interviews with new organizations.	
Achieve diverse membership in collaborative structure (engage diverse players in other collaborative efforts)	Appropriate types and numbers of members/players with respect to issue (DV and SA), systems (law enforcement, prosecution, corrections, advocates), and community (reflecting demographics of community). Each subcommittee or task force reflects the diversity of the whole collaborative structure. Length of time various members have been involved.	Committee membership records. Self-reported member/participant surveys.	

Table 10.1
Community Collaboration, Goal 1:
Establish an Effective, Stable, and Continuing Community Response to VAW

Objectives	Specific Measures	Data Collection Procedures	Caveats
Level/positions of participating collaborators.	<p>Are collaborators in positions within their respective agencies to: make changes and influence policy? to implement changes in service practice (i.e., line workers)? to understand barriers to success and how to overcome them?</p> <p>Are those who are in power in the relevant agencies supportive of VAW efforts? Have they: (a) taken an active role in; (b) been ambivalent toward; or (c) opposed VAW efforts?</p> <p>Why have/haven't certain people joined the efforts?</p>	<p>Track record of participants, e.g. have they delivered what they've promised?</p> <p>Observation: determine which agencies/political players have/have not shown an interest in or taken an active role in VAW efforts.</p> <p>Agency heads or public relations staff can be contacted to determine reasons for supporting/not supporting efforts.</p>	It may be difficult to get government officials to admit on the record why they don't support efforts. Informal contacts with agencies and political players may be useful to determine reasons for refusal of support.
Engage active participants	<p>Share of work time devoted to collaboration.</p> <p>Number of tasks each participant initiates, plans, and completes.</p> <p>Number of meetings attended by participant.</p>	<p>Self-reported participant surveys and interview with committee coordinator.</p> <p>Minutes of committee meetings.</p>	
Opportunities for involvement	<p>Nature and level of involvement of various members (set agenda, participate fully, observe only).</p> <p>Who makes major decisions? Who leads activities? If there is one leader, does he/she delegate authority to other members? Are all participants involved in the decision-making process?</p> <p>Number of policy changes that participants try to install in their own agencies. Success of such attempts.</p>	<p>Observation of meetings.</p> <p>Surveys of participants.</p> <p>Examination of activities: determine who was responsible for prioritizing and organizing the task and completing the work.</p>	
Level/source of support	<p>What is the highest level of support for VAW efforts? State? Region? Locality? Do the efforts also have support at the grassroots level?</p> <p>How has the collaboration benefited from this support (e.g., funding, media attention)</p>	<p>Observation; interviews with participants.</p> <p>Funding records of collaborative partners; press clippings.</p>	

Table 10.2
Community Collaboration, Goal 2:
Achieve System Level Outcomes

Objectives	Specific Measures	Data Collection Procedures	Caveats
<p>Achieve frequent, positive communication among members</p>	<p>Frequency of inter-agency contact: one time, as-needed, regularly scheduled.</p> <p>Type of contact: referrals only, telephone, meetings, conferences, training sessions.</p> <p>Level of contact: directors/senior managers, trainers, line workers.</p> <p>Basis of contact: individual initiative, contractual, formal agreements, memoranda of understanding.</p> <p>Nature of contact: hostile, reluctant, indifferent, cordial, amicable.</p> <p>Are participants informed about actions of the group/committee? Do participants provide updates on their activities for other members?</p>	<p>Labor intensive approach: Committee members and coordinator keep a general log of the people they contact, how often, the nature of the contact, etc. Members listed in logs may also be interviewed in order to double-check this information and account for possible differences in perception.</p> <p>Much less labor intensive approach: Periodic interviews (e.g., every quarter) with committee members and others involved in collaborative effort. Use questions with quantitative answer categories such as: "how often in contact with X...daily, weekly, monthly, etc."</p>	<p>The communication logs can provide information on both member and non-member contacts, but in this case only members would be interviewed.</p> <p>It is not likely that people will comply with the log approach, as it takes a lot of time and will contain a good deal of extraneous information. A quarterly interview is probably "good enough" to track changes in interaction patterns.</p>
<p>Create an informal communication networks (non-member contacts).</p>	<p>Number of non-member agencies or organizations contacted.</p> <p>Frequency of non-member contacts.</p> <p>Length of relationship.</p>	<p>This information can be obtained from the general communication logs. Non-members with whom contacts were made may be interviewed in order to compare information.</p>	
<p>Provide opportunities for feedback from general community.</p>	<p>Number of meetings open (and published materials made available) to the general public.</p> <p>Frequency with which community feedback is incorporated into decisions and/or projects.</p>	<p>View committee records.</p> <p>Survey of community members to determine if they are aware of committee's activities, have had opportunities to give feedback and have done so, and whether or not that feedback was incorporated into committee decisions and/or projects.</p>	

Table 10.2
Community Collaboration, Goal 2:
Achieve System Level Outcomes

Objectives	Specific Measures	Data Collection Procedures	Caveats
Develop a shared vision.	Do committee members have a clear and consistent perception of their roles and the roles of others? Have they been able to put differing opinions aside and agree upon a shared vision or primary goal?	Interviews with members and coordinator. Observation of meetings, records.	
Create a written mission statement and/or strategic plan.	Does the mission statement or strategic plan exist? Does it represent the opinions of the majority of the members?	Examination of the mission statement. Interviews/surveys of members and committee coordinator.	
Establish a mechanism for conflict resolution and feedback among members.	Number/share of committee conflicts that are resolved. Number of opportunities available for members to voice objections and/or give feedback. Frequency with which member feedback is incorporated in committee's activities.	Review of minutes; observation of meetings. Interviews/surveys of collaboration participants.	
Establish a system for evaluation of programs/ achievements.	Does a formal program evaluation system exist? If so, how many programs (and/or what percentage) are evaluated? How often are they evaluated? Who is evaluator? Does the group monitor general achievements?	Interviews/surveys of collaboration participants. Examination of committee records. Review evaluation results (if available).	
Develop trust and mutual respect.	How do group members perceive each other? Is there a general feeling of trust?	Surveys implemented initially and after the collaboration has been in existence for a substantial period.	

Table 10.2
Community Collaboration, Goal 2:
Achieve System Level Outcomes

Objectives	Specific Measures	Data Collection Procedures	Caveats
Engage in joint planning, prioritizing, and implementation of tasks; co-location.	Number/types of members and/or agencies involved in these activities. Number of these activities.	Observation of meetings. Examine joint activities; document who is involved. Interviews/surveys of participants.	
Create formal reporting procedures.	Is there a formal written report produced by the committee? If so, how often, what type of format, who receives the report, total number of reports distributed, specific and general effects of publishing the report.	View written document. Interview those receiving document and record their opinions regarding the usefulness and validity of report. Record direct and indirect results of publishing report (e.g., change in funding, attention from media, community interest).	It may be difficult to document results of publishing the report, but this may be achieved by interviewing recipients of the report (e.g., funders, media) and seeking their opinions.
Increase funding.	Are any new funding sources available as a result of community collaboration? Are you able to take better advantage of existing funding sources because of collaboration? Do collaborative structures have any responsibility for distributing/ allocating funds within community?	Review major sources of funds for VAW; identify roles and responsibilities of collaborative structures and collaborating agencies in terms of use and allocation of funds.	

Table 10.3
Community Collaboration, Goal 3:
Achieve Ultimate Outcomes

Objectives	Specific Measures	Data Collection Procedures	Caveats
Permanent policy changes.	Have the collaboration efforts resulted in permanent changes in agency procedures and/or laws (e.g., guidelines for arrest of abusers, new training curricula)?	Check legislative history of state/city/locality. Interview collaborating participants.	
Consistent treatment of victims and perpetrators.	Are the same options available to victims at all points of entry to the system? Are perpetrators consistently held accountable by all parts of the system?	Interview/survey personnel of various agencies (e.g., victim services, law enforcement, courts, child welfare). Survey victims using services. Interview law enforcement agencies and court personnel. Examine court records and/or newspapers to determine sentencing patterns for convicted rapists/abusers.	
More options and resources for victims.	More options: what types of services are available to victims and perpetrators in the justice and human service systems? Increase in number of referrals. Better options: standards written, standards meet criteria of model programs, standards enforced on programs throughout the community. Duplication: where does duplication of services occur?	Check state, city, and agency records. Review hotline records, comparing need stated/requests made to information given Review STOP subgrant award reports	
Increased knowledge in general community re VAW; increased use of services.	How many people are aware of: --the problem of violence against women; --services that are available; --laws pertaining to victims and abusers. How many people and/or what percentage of the population in need are actually using services and/or reporting violence?	Survey residents. Records of service agencies and law enforcement offices. Survey victims and their families/friends. See Chapter 11 for some survey resources.	

CHAPTER 11

MEASURING CHANGE IN COMMUNITY ATTITUDES, KNOWLEDGE OF SERVICES, AND LEVEL OF VIOLENCE

Many STOP grants have as either intermediate or ultimate outcomes the goal of making communities safer and more supportive of women victims of violence. Steps toward this goal can take many forms. Many projects try, through outreach, thoughtful location of services, education of professionals and the general public, and public service announcements to increase women's knowledge of the services that are available to help them deal with domestic violence and sexual assault. These projects want women to know that help is available if they need it, and where to find it. Other projects try to reduce the level of support for violence against women by reducing myths and attitudes that blame women for their own victimization, minimize the harm done by violence, discourage women from trying to improve their situation, and provide excuses for male perpetrators of violence. In efforts to change community behavior, projects may work to enlist support from community members other than the "usual suspects" from the justice systems and victim advocacy. These other community allies and partners can do many things to make your community safer for women. Finally, many projects want to know what the true level of violence against women is in their community, and most projects hope that the ultimate effect of all their efforts will be to reduce the overall level of this violence. This chapter provides a brief look at some ways that these goals might be measured. The chapter is far from exhaustive, but it does offer some resources for measuring the achievement of these goals.

KNOWLEDGE OF SERVICE AVAILABILITY

It would be hard to imagine a program offering services and supports for women victims of violence that was not interested in whether the community knew about its existence. Therefore any service might profit from doing some investigation into whether its name rings any bells among members of the larger community. The issue of community knowledge is especially important for new services, and for services that are trying to reach new populations. Since one of the major VAWA objectives is to extend services to previously underserved populations, many projects have been funded with this objective. All of these projects should be making an effort to learn whether members of their target community have heard of them.

What to Ask About

Basically, you want to know about *recall* and about *recognition*. "Recall" means that when you ask someone "Do you know anywhere you could go to get help if you were raped?" the person says, "Yes, I would go to the ABC Crisis Service." This answer indicates they *recall* the name and purpose of your agency, and do so without prompting or help from you. Even if people cannot recall the name of your agency, they may still *recognize* it when asked a question such as "Do you know what ABC Crisis Service does?" Or, you could ask "Which of these agencies would you go to if you had just been raped and wanted help?" and then give the person a list of

Evaluation Guidebook

agencies that includes ABC Crisis Service and see whether she pick out ABC Crisis Service as a place to go for this type of help.

With people who already know about your program, there are several questions you can ask to get important feedback. You can ask whether they know your location, your hours, whether you speak their language, and whether they could get to your agency using the transportation available to them. You could then ask them how they learned about your program, whether they have ever used your services, or whether they know anyone who has. Follow-up questions could ask for feedback about the program and its services, such as whether they have heard that you treat women with dignity and sympathy.

Sample questions include the following (answered “yes” or “no”):

- Do you know of a specific place or program in your community:
 - That can help if a family member is being violent to someone else in the family?
 - That offers information and assistance to someone who has just been raped?
 - Where someone could talk about how they’re feeling about a sexual assault that happened some time ago?
 - That will help you get a restraining order to keep someone who has been harassing you away from you?
 - And so on.
- For each “yes” answer, you could then ask:
 - Do you know the name of that agency? What is it?
 - How did you hear about that agency?
 - Do you know where it is?
 - Do you know what kinds of people use that agency?
 - And so on.
- Continue making up questions to get the answers you want. Be sure that the questions are worded very simply.

How, When, and Whom to Ask

There are many different places where you could ask these questions. Each place will let you learn about perceptions of your program from a different group of women. Different places may require you to use different formats, some of which will be very formal while others will be very informal, such as group discussions. Here are a few of the places and formats you might be able to use to get information about how well your program is known in the community:

- ❑ **Focus groups** with women who use a variety of generic service agencies (e.g., community health centers, welfare offices, employment centers, child care centers, cultural or recreational centers). These are very informal situations, in which you pick between 4 and 6 questions on which you want feedback, and keep the group focused on providing answers to those questions. Do not try to treat a focus group as a group interview; you will not get quantitative data from a focus group. But you will get a lot of feedback about your program. Also, you might want to be prepared to find out that few of them have ever heard of your program, and spend time exploring what it would take to spread the word.
- ❑ **Exit interviews or questionnaires** in generic service agencies such as health, social services, or welfare agencies, or in other VAW-related locations such as the police department. Interviews could be formal or informal. Questionnaires have to be written, of course, and should be short (1 page is best).
- ❑ **Surveys of community members.** This is the most formal approach. You would have to develop a formal, written survey containing the questions to which you want answers. Then you would have to do one of the following: find a firm to conduct the survey, recruit volunteers and organize several nights during which you call households at random, find an existing household telephone survey from which you can buy some time and attach your questions, or develop other methods yourself.

ATTITUDES TOWARD VIOLENCE AGAINST WOMEN

The causes of rape, other sexual assaults, and domestic violence are many and complex. One causal factor may be community support for or tolerance of violence against women. Attitudes such as acceptance of sex-role stereotypes, belief in rape myths, and beliefs that some circumstances justify battering help to create a climate that gives abusers and rapists a justification for their actions, and is hostile to victims of domestic violence and rape. A wide variety of Americans, including typical citizens, police officers, and judges, have been shown to hold beliefs that can be used to justify rape (Burt, 1980; Feild, 1978; Mahoney et al., 1986) and domestic violence (Broverman et al., 1970; Pagelow, 1981).

In addition to supporting violence against women, attitudes may also have the effect of creating an environment that is hostile to victims. Rape-supportive attitudes are partly responsible for low levels of rape reporting (Russell, 1982) and for a blaming-the-victim attitude that makes it difficult for victims to seek help and recover from their assaults (Ehrhart & Sandler, 1985). Belief in rape myths leads to a strict definition of rape and denies the reality of many actual rapes (Burt & Albin, 1981), which makes it difficult to prosecute rapists and support victims. Rape victims are often victimized twice—once from the actual assault and a second time when they encounter negative, judgmental attitudes from the police, courts, and family and friends (e.g., Deming & Eppy, 1981; Weis & Borges, 1975; Williams, 1984).

Myths and stereotypes about victims of domestic violence often make it difficult for these victims to seek help and to improve or sever relationships with their abusers. As mentioned earlier, police are less likely to arrest offenders if they believe that domestic violence is a “family issue” or that victims will refuse to press charges (Pagelow, 1981). And negative attitudes—such as apathy or hostility—that are often experienced by abused women when they reach out for help, may actually help perpetuate abuse by stigmatizing women and making it harder to leave abusive relationships (Stark, Flitcraft, & Frazier, 1979). Because they contribute to the perpetuation of violence against women, recognizing and measuring community attitudes can play an important part in evaluating the effectiveness of some STOP-funded projects.

Measuring Community Attitudes

A number of quantitative measures can be used to examine attitudes toward violence against women. Some, such as the *Attitudes Toward Women Scale* (Spence & Helmreich, 1972) measure general attitudes about the roles and rights of women, while others, including the *Rape Myth Acceptance Scale* (Burt, 1980) and the *Inventory of Beliefs about Wife Battering* (Saunders, et al., 1987) assess attitudes specifically related to rape and domestic violence. These measures can be used in a variety of ways, depending on the preference of the grantees and the intended target of their programs.

Chapter 11: Measuring Change in Community Attitudes

The most practical and least expensive way to implement the measures is not to survey the broad community, but to focus on those that could be considered a “captive audience,” such as students, potential jurors, or members of civic or religious organizations. For example, if a prevention program has given presentations to students in a particular high school, grantees may wish to survey students in that high school and in a school without the prevention program and then compare the results.

Some communities might have the capacity to conduct surveys with a broader range of community members. If you are trying to change attitudes in the whole community, it is much more convincing to conduct a community-wide survey than to use “captive audiences,” many of whom will surely have been exposed to your message. One way to do this, if your state has a periodic statewide poll, is to connect with the organization that conducts it and add a few questions about community attitudes toward violence against women. Another way is to find a sponsor who will support a special-focus survey of your particular community.

Rape Myths

Researchers in the field of rape have argued that these widely accepted myths—defined as prejudicial, stereotyped, or false beliefs about rape, rape victims, and rapists—support and promote rape. Rape myths are part of the general culture and people learn them in the same way they acquire other beliefs and attitudes—from parents, friends, newspapers, books, and television (Burt, 1991). Some of these myths include (1) women enjoy sexual violence; (2) sex is the primary motivation for rape; (3) women are responsible for rape prevention; (4) only bad women are raped; (5) women falsely report rape; and (6) rape may be justified (Lottes, 1988). Belief in rape myths is significant because it “provides men with a structural position from which to justify sexually aggressive behavior” (Marolla & Scully, 1982). For example, Scully and Marolla (1984, 1985a, 1985b) found that the myths that women both enjoy and are responsible for their rape were used by convicted rapists to excuse and justify their crimes.

RAPE MYTH ACCEPTANCE SCALE

Citation Burt, M. R. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology*, 38, 217-230. Copyright © 1980 by the American Psychological Association. Used with permission.

Description The Rape Myth Acceptance Scale is a 19-item measure of the acceptance or rejection of myths about rape. It was originally developed for a study of community attitudes, but has become a popular scale for use on a variety of issues, including the narrowness of people’s definition of a “real” rape, people’s use of different types of information in making rape judgments, jurors’ likelihood of convicting, college and other non-incarcerated men’s likelihood of raping, and in sex offender treatment. Some of its items may seem a bit dated, but it still works. It is intended for use with adults (18 and older), has been used extensively with college students, and has occasionally been used with younger adolescents.

Sample items:

Response categories:

- 1=disagree strongly
- 2= disagree somewhat
- 3=disagree slightly
- 4=neutral
- 5=agree slightly
- 6=agree somewhat
- 7=agree strongly

- 3. Any healthy woman can successfully resist a rapist if she really wants to.
- 5. When women go around braless or wearing short skirts or tight tops, they are just asking for trouble.
- 10. In the majority of rapes, the victim was promiscuous or had a bad reputation.

Chapter 11: Measuring Change in Community Attitudes

- References Burt, M. (1991). Rape myths and acquaintance rape. In A. Parrot & L. Bechhofer (Eds.), *Acquaintance Rape: The Hidden Crime*. New York: Wiley & Sons.
- Burt, M., & Albin, R. (1981). Rape myths, rape definitions, and probability of conviction. *Journal of Applied Social Psychology*, *11*, 212-230.

Another possible measure:

BUMBY COGNITIVE DISTORTIONS SCALE

- Citation Carich, M.S. & Adkerson, D.L. (1995). *Adult Sex Offender Assessment Packet*. Brandon, VT: Safer Society Press. Copyright © 1995 by Safer Society Press. Used with permission.
- Description The Bumby Cognitive Distortions Scales include both a Rape Scale (36 items) and a Molest Scale (38 items) These scales were developed by Dr. Kurt Bumby and take 5-10 minutes each to complete. They are intended for use by adolescents and adults. Dr. Bumby can be contacted for more information at Fulton State Hospital, Mail Stop 300, 600 East 5th Street, Fulton, Missouri 65251-1798. Many of the items on this rape scale are almost identical to those on Burt's Rape Myth Acceptance Scale, but there are almost twice as many items so other myths are also covered.

Sample items:

Response categories:

- 1=disagree strongly
- 2= disagree somewhat
- 3=agree somewhat
- 4=agree strongly

Rape items:

- 3. Women usually want sex no matter how they can get it.
- 6. Women often falsely accuse men of rape
- 13. If a man has had sex with a woman before, then he should be able to have sex with her any time he wants.

Domestic Violence Myths and Stereotypes

Popular myths and stereotypes abound on the subject of domestic violence victims and batterers. The public and many professional groups often hold negative attitudes toward battered women (Dobash & Dobash, 1979; Gelles, 1976; Straus, 1976). Victims are often thought to be (1) masochistic; (2) weak; (3) “seeking out” the batterers; or (4) somehow at fault. Batterers are excused because they are “sick” or their force is perceived as justified because of the wife’s behavior (Greenblat, 1985; Pagelow, 1981). One way in which domestic violence myths are harmful is that they are often believed by the people responsible for aiding domestic violence victims: social workers, judges, health professionals, and law enforcement officers. For example, a common perception among police officers is that domestic violence victims are too weak-willed to press charges. This is particularly damaging in cases where officers base their decision to arrest an abuser on whether or not they believe the case will ever reach court (Bates & Oldenberg, 1980).

Several scales have been developed to measure attitudes toward battering, including the Attitudes to Wife Abuse Scale (Briere, 1987) and the Inventory of Beliefs about Wife Beating by Saunders and his colleagues (1987).

ATTITUDES TO WIFE ABUSE SCALE

Citation Briere, J. (1987). Predicting self-reported likelihood of battering: Attitudes and childhood experiences. *Journal of Research in Personality, 21*, 61-69. Copyright © 1987 by Academic Press. Used with permission.

Description The Attitude to Wife Abuse Scale (AWA) is an 8-item measure of attitudes towards women and the abuse of women. The AWA was developed by Dr. John Briere, and is intended for use with adolescents and adults.

Sample Items

Response categories:

- 1=disagree strongly
- 2= disagree somewhat
- 3=disagree slightly
- 4=neutral
- 5=agree slightly
- 6=agree somewhat
- 7=agree strongly

- 3. A husband should have the right to discipline his wife when it is necessary.
- 5. A man should be arrested if he hits his wife.
- 8. Some women seem to ask for beatings from their husbands.

INVENTORY OF BELIEFS ABOUT WIFE BEATING

- Citation** Saunders, D.; Lynch, A.; Grayson, M.; & Linz, D. (1987). The inventory of beliefs about wife beating: The construction and initial validation of a measure of beliefs and attitudes. *Violence and Victims*, 2, 39-57. Copyright © 1987 by Springer Publishing Company, Inc. Used with permission.
- Description** The Inventory of Beliefs about Wife Beating (IBWB) is a 31-item scale developed by Dr. Daniel Saunders to measure attitudes and beliefs about wife beating. It covers many more issues than the Briere scale, including attitudes toward appropriate intervention (none, arrest the husband, social agencies do more to help), attributions of responsibility for battering (e.g., the husband, because.....; the wife, because....), and the value of wife-beating (improves a marriage). Five reliable subscales emerge (1) WJ=wife beating is justified, (2) WG=wives gain from beatings, (3) HG=help should be given, (4) OR=offender is responsible, and (5) OP=offender should be punished. Scales proved to have high construct validity with other scales, and to differentiate well among known groups, especially differentiating abusers from college students and advocates, and male from female students.
- Sample Items**
- Response categories:
- 1=strongly agree
 - 2= agree
 - 3=slightly agree
 - 4=neutral
 - 5=slightly disagree
 - 6=disagree
 - 7=strongly disagree
3. Wives try to get beaten by their husbands in order to get sympathy from others.
7. Even when women lie to their husbands they do not deserve to get a beating.
18. If a wife is beaten by her husband she should divorce him immediately.
27. Occasional violence by a husband toward his wife can help maintain the marriage.

Adversarial Sexual Beliefs and Hostility Toward Women

Adversarial sexual beliefs and general hostility toward women are attitudes that contribute to violence against women. Adversarial sexual beliefs refer to the feeling that men and women are adversaries in their sexual relationships with one another—that relationships are exploitative and each party is manipulative and not to be trusted (Burt, 1980). Hostility toward women is related to adversarial sexual beliefs, but goes beyond sexual interactions to assess a more general feeling of animosity and distrust of women in all aspects of relationships and interactions. Check (1984) expanded Burt's Adversarial Sexual Beliefs Scale (Burt, 1980) to create a measure of general hostility toward women and found that men who scored high on the hostility scale tended to (1) have traditional sex-role beliefs; (2) believe in rape myths; (3) admit to using force in their attempts to get women to have sex with them, and say they would be likely to do so again; and (4) use high levels of physical punishment when rejected by women and given the opportunity to retaliate.

ADVERSARIAL SEXUAL BELIEFS SCALE

Citation Burt, M. R. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology*, 38, 217-230. Copyright © 1980 by the American Psychological Association. Used with permission.

Description This scale was developed as part of Burt's research on community attitudes supportive of rape, done in the late 1970s. It is strongly associated with belief in rape myths, but is not the same as rape myth acceptance. It is also related to sex-role stereotyping but, where sex-role stereotyping usually focuses on adult economic and family roles, Adversarial Sexual Beliefs focuses on interaction patterns related specifically to sexual behavior. It has been used frequently in research since first being published in 1980.

Sample items:

Response categories:

- 1=disagree strongly
- 2= disagree somewhat
- 3=disagree slightly
- 4=neutral
- 5=agree slightly
- 6=agree somewhat
- 7=agree strongly

- 3. A man's got to show the woman who's boss right from the start or he'll end up henpecked.
- 5. Women are usually sweet until they've caught a man, but then they let their true self show.
- 10. A lot of women seem to get pleasure in putting men down.

HOSTILITY TOWARD WOMEN SCALE

Citation Check, J. V. P. (1985). The Hostility Towards Women Scale (Doctoral dissertation, University of Manitoba, 1984). *Dissertation Abstracts International*, 45 (12). Used with permission of author.

Description The Hostility Toward Women scale (HTW) is a measure of anger and resentment toward women. Developed by Dr. James Check as part of his doctoral research, the HTW consists of 30 items. It began with the tone of Burt's Adversarial Sexual Beliefs Scale, and developed a set of more general items reflecting hostile attitudes toward women that covered a variety of different domains (not just sexual interactions). It is intended for use with adults.

Sample items:

Response categories (true/false):

1. I feel that many times women flirt with men just to tease them or hurt them.
11. I don't seem to get what's coming to me in my relationships with women.
13. Women irritate me a great deal more than they are aware of.
17. It is safer not to trust women.

Reference Check, J.; Malamuth, N.; Elias, B.; & Barton, S. (1985). On hostile ground. *Psychology Today*, 56-61.

Sex-Role Stereotypes

Traditional sex-role socialization or sex-role stereotyping, which views women as having a lower social status and lesser rights than men, appears to play an important role in violence against women. People who hold these beliefs—for example, feeling that a woman should get married and raise a family, be a virgin when she marries, and never contradict her husband in public—are more likely to be tolerant of and/or engage in domestic violence (e.g., Eisikovits, Edleson, Guttman & Sela-Amit, 1991; Kincaid, 1982; Koss et al., 1985, in Parrot). In one study by Finn (1986), traditional sex-role attitudes were found to be the most powerful predictor of attitudes supporting marital violence. Sex-role stereotypes also play a role in rape and sexual assault. In their analysis of sexually aggressive and non-aggressive men, Koss and Dinero (1988) found that sex-role stereotyping was causally related to sexual assault in that “the more sexually aggressive a man has been, the more likely he was...to accept sex-role stereotypes.”

ATTITUDES TOWARD WOMEN SCALE

Citation Spence, J.T., & Helmreich, R. L. (1972). The Attitudes Toward Women Scale: An objective instrument to measure attitudes toward the rights and roles of women in contemporary society. *Psychological Documents*, 2, 153. Used with permission of the author.

Description The Attitudes Toward Women Scale is a 55-item measure that was originally designed to assess opinions about the rights and roles of women. It focuses on family and economic roles mostly, and on norms and obligations of one gender to the other. It has been widely used in the 25+ years since it was first published, in its long form and also in a shorter 15-item version (see Spence & Hahn, 1997, below). This scale is intended for adults.

Sample items:

Response categories:

- 1=agree strongly
- 2=agree mildly
- 3=disagree mildly
- 4=disagree strongly

- a. Husbands and wives should be equal partners in planning the family budget.
- b. Intoxication among women is worse than intoxication among men.
- c. There should be a strict merit system in job appointment and promotion without regard to sex.
- d. Women should worry less about their rights and more about becoming good wives and mothers.

Reference Spence, J.T. & Hahn, E.D. (1997). The Attitudes Toward Women Scale and attitude change in college students. *Psychology of Women Quarterly*, 21, 17-34.

Masculinity Ideology

Also playing a role in support of violence against women is masculinity ideology. Masculinity ideology, which refers to beliefs about the importance of men adhering to culturally defined standards for male behavior, may be responsible for some problem behaviors—particularly in adolescent men. In their 1993 study, Pleck, Sonenstein, and Ku found that a higher score on the Masculinity Ideology Index was associated with, among other things, twice the odds of being sexually active and of ever forcing someone to have sex. And in a study of 175 college males, Mosher and Anderson (1984) found that a measure of macho personality with three components—callous sex attitudes towards women, a conception of violence as manly, and a view of danger as exciting—was significantly correlated with a history of self-reported sexual aggression against women.

MASCULINITY IDEOLOGY

Citation Thompson, E., H., & Pleck, J. H. (1986). The structure of male role norms. *American Behavioral Scientist*, 29, 531-543. Copyright © 1986 by Sage Publications. Used with permission.

Description The Masculinity Ideology Scale is an 8-item measure that assesses a male's endorsement and internalization of cultural belief systems about masculinity and male gender. It is adapted from Thompson and Pleck's Male Norms Scale (MRNS), a 26-item abbreviated version of the Brannon Masculinity Scale, Short Form. It is intended for use with adolescents and adults.

We do not have permission to reproduce actual scale items from this scale in electronic form.

Reference Brannon, R. (1985). A scale for measuring attitudes about masculinity. In A.G. Sargent (Ed.), *Beyond Sex Roles* (pp. 110-116). St. Paul, MN: West.

“BEYOND THE USUAL SUSPECTS”—COMMUNITY ALLIES AND PARTNERS

Most STOP projects will focus their efforts on the civil and criminal justice system agencies, and on providing services and advocacy to women victims of violence. These are “the usual suspects.” But some STOP projects are trying to bring more players into the effort to end violence against women. The previous section on public attitudes covered one approach to involving others, namely, trying to reduce beliefs and attitudes in the community that minimize perceptions of harm, blame women, and justify acts of violence. This section addresses efforts to involve particular local actors—preferably people with influence who are opinion leaders—as allies and partners in ending violence. Possibilities include the following (all of which have occurred in at least some communities):

- ❑ **Business and corporate leaders**—might (1) change the coverage of their employee health insurance policies, (2) establish special assistance opportunities through their employee assistance programs (EAPs), or (3) institute security procedures to protect women against harassment and stalking. In addition, they could (4) play a prominent role in a local anti-violence council, (5) organize a campaign to have several corporations “adopt-a-victim-services-program” (as has actually happened in at least one state, where corporations give ongoing financial and other support to “their” domestic violence shelter), (6) donate computer hardware and know-how to develop efficient databases, (7) fund a community telephone survey to establish levels of violence, and so on.
- ❑ **Clergy**—might (1) develop training sessions for other clergy that set anti-violence actions in a convincing doctrinal context and recruit more voices against violence, (2) preach against violence and for the rights of women to be violence-free; (3) get training to be supportive counselors when women approach them with a problem related to violence; (4) serve on a community-wide anti-violence council, and so on.
- ❑ **Doctors and therapists**—might (1) obtain training that would help them offer support and assistance to women victims of violence and, at a minimum, not make things worse, (2) develop into better screeners to identify more women victims of violence, (3) influence the practices of emergency rooms and other treatment settings that they control.
- ❑ **Substance abuse treatment professionals**—there is a tremendous overlap in treatment populations between batterers and men attending substance abuse treatment (especially those ordered to do so as a consequence of drunk driving and drunk and disorderly offenses). These treatment professionals could develop special sessions in which they deal with issues of battering and its relationship to substance abuse in the lives of treatment participants.

Chapter 11: Measuring Change in Community Attitudes

- ❑ **Journalists**—might (1) become familiar with the latest thinking about violence against women, (2) learn what is happening in their community to combat this violence, and (3) regularly publicize the issue with article on programs, community activities such as Take Back the Night marches, successful survivors, reformed batterers, and so on.

- ❑ **Politicians**—in addition to making public statements opposing violence against women, the main things politicians can do are (1) put money in city, county, and state budgets to fund services; (2) update laws and ordinances to conform to the latest legislative recommendations of anti-violence advocates; (3) mandate VAW-related training for all new recruits/professionals; and (4) mandate *regular in-service training* for all professionals dealing with women victims of violence, focusing on the latest in VAW-related legal and other developments (e.g., best practices).

For any project with objectives that include the development of community partners and allies, an evaluation will need to document the goals and the project's success in reaching them. Process analysis techniques will be the ones to use in this effort. Goals should be specified in terms of who (which allies are targeted), how many, how (what approaches will be used, and when), and for what (what do you want them to do). Goals can be documented by examining written project materials and by conducting interviews with key staff. Success in reaching the goals can be documented by interviewing key project staff, but that is only the beginning. *It is essential* that you also do the first bullet below, and that you consider the remaining bulleted suggestions:

- ❑ Interview the people who were targeted to see what they think of the campaign, what they believe their activities have been, and so on.

- ❑ To the extent that the community partners and allies have produced anything concrete, these products should be collected and described. Such products might include revised personnel policies, changing patterns of EAP use, newspaper articles, expanded budgets, expanded contributions to service agencies, new screening procedures in place, new policies in the emergency room, and so on.

- ❑ Other possible indicators to examine if they are relevant to your goals include the following:
 - The number of community leaders—such as business people, politicians, and clergy—speaking out against domestic violence and sexual assault.
 - The number of volunteers at rape crisis centers, domestic violence shelters, and hotlines.
 - Support from the school system for adolescent violence prevention programs.

- Media coverage that portrays rape and domestic violence in a victim-supportive manner.

MEASURING COMMUNITY LEVELS OF VIOLENCE AGAINST WOMEN

Ultimately, everyone involved in STOP wants to see reductions in levels of violence against women. However, documenting success on this goal usually falters on the shoals of inadequate measurement. Police statistics are known to underrepresent levels of violence, and are also biased (meaning that they miss particular types of women victims, as well as generally not getting all the incidents). The direction of bias may vary from community to community depending on local attitudes and practices, but some type of bias will virtually always be there. Records from victim service agencies and hotlines also undercount and have biases, although the biases are often quite different from the biases of police data. National surveys such as the National Crime Victim Survey undercount, are biased in various ways, and cannot be used to reveal what is happening in particular communities because their sample sizes are only adequate for national estimation.

There are no *easy* solutions to this dilemma, but that does not mean that there are *no* solutions. We recently came across a state-level effort that seems reasonably priced and reasonably successful in getting respondents. Further and most important, it obtained estimates of lifetime and recent victimization that seem in line with what one might expect (rather than being grossly low). We thought it would be worth sharing this resource with users of this *Guidebook*, because some of you may want to replicate this study in your own state or community.

In 1996, the Michigan Department of Community Health, using funding from the Centers for Disease Control and Prevention, developed and conducted a telephone survey of Michigan women to determine the lifetime and recent prevalence of violence in their lives.⁸ The violence examined included physical or sexual violence or threats of violence from strangers, boyfriends or dates, other acquaintances, current intimate partners including husbands and live-in partners, and ex-partners. Reports were limited to experiences since age 16. The structure and portions of the questionnaire content were based on an instrument that Statistics Canada used in its first national survey of violence against women in 1993. The telephone survey itself was conducted by the Gallup Organization, using random digit dialing.

Results represent the general population of women ages 18-69; 1848 women completed the interview. There is some bias in the survey respondents toward currently married women, employed women, and higher income women, as might be expected from a telephone survey

⁸ Contacts at the Michigan Department of Community Health include Patricia K. Smith, Violence Prevention Program Coordinator (513 335-9703), Ann Rafferty and John C. Thrush. The survey instrument is available from Michigan DCH and also from the STOP TA Project (800 256-5883 or 202 265-0967). Contact Patricia K. Smith for details regarding the conduct of the study, its costs, time and staff required, methodological lessons learned, and suggestions for revisions of the questionnaire.

Chapter 11: Measuring Change in Community Attitudes

methodology. Weighting was used to produce the final results, using age, race, and educational status as factors in the weighting to adjust for the sample biases and obtain representative data.

To give you a flavor of the results and their reasonableness, we quote (with permission) from information distributed at a presentation of this study in Indianapolis, at the 1997 Annual Meeting of the American Public Health Association. The study found:

- 58% reported experiencing some type of violence (physical, sexual, or threat) by a man since age 16.
- 12% reported at least one form of violence during the past year.
- 38% reported experiencing physical violence by a man, while 40% reported some form of sexual violence.
- Over one-third (38%) of women who have ever had a partner reported having experienced violence by a partner.
- Nearly half the respondents said they had experienced some type of violence by a non-partner since age 16.
- 27% reported that a non-partner either raped or attempted to rape them. 38% reported some type of unwanted sexual acts by a non-partner.

These results, and others available from the study team (see footnote 2) will suggest to anyone who has tried to do this type of survey that this team did something right. Anyone interested in conducting a similar survey would be wise to discuss methodology with this study team. In particular, important things to discover are (1) how they conducted telephone screening that identified women and established a safe and secure time for the interview, and (2) how they introduced the study as a whole and each section of the questions to elicit the level of self-disclosure that they achieved. At the time this *Guidebook* went to press, no final report was available for this study, but you should be able to get it soon from the study team, and it will undoubtedly make interesting reading.

ADDENDUM: ADDITIONAL READING

Bates, V., & Oldenberg, D. (1980). Domestic violence and the law. Paper presented at the annual meeting of the Pacific Sociological Association, San Francisco.

Broverman, I., Broverman, D., Clarkson, F., Rosenkrantz, P. & Vogel, S. (1970). Sex-role stereotypes and clinical judgements of mental health. *Journal of Consulting and Clinical Psychology*, 34, 1-7.

Evaluation Guidebook

- Burt, M. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology*, 38, 217-230.
- Burt, M. (1991). Rape myths and acquaintance rape. In A. Parrot & L. Bechhofer (Eds.), *Acquaintance Rape: The Hidden Crime*. New York: Wiley & Sons.
- Burt, M., & Albin, R. (1981). Rape myths, rape definitions, and probability of conviction. *Journal of Applied Social Psychology*, 11, 212-230.
- Check, J. (1984). *The Hostility Toward Women Scale*. Unpublished doctoral dissertation, University of Manitoba.
- Deming, M., & Eppy, A. (1981). The sociology of rape. *Sociology and Social Research*, 65, 357-380.
- Dobash, R. E., & Dobash, R. (1979). *Violence against wives: A case against patriarchy*. New York: Free Press.
- Ehrhart, J., & Sandler, B. (1985). *Myths and realities about rape*. Washington, DC: Project on the Status and Education of Women.
- Eisikovits, Z. C., Edleson, J. L., Guttman, E., & Sela-Amit, M. (1991). Cognitive styles and socialized attitudes of men who batter: Where should we intervene? *Family Relations*, 40, 72-77.
- Field, H. (1978). Attitudes toward rape: A comparative analysis of police, rapists, crisis counselors, and citizens. *Journal of Personality and Social Psychology*, 36, 156-179.
- Finn, J. (1986). The relationship between sex role attitudes and attitudes supporting marital violence. *Sex Roles*, 14 (5/6), 235-244.
- Gelles, R. (1976). Abused wives: Why do they stay? *Journal of Marriage and the Family*, 38, 659-668.
- Kincaid, P. J. (1982). *The omitted reality: Husband-wife violence in Ontario and policy implications for education*. Concord, Ontario: Belsten.
- Koss, M. & Dinero, T. (1988). Discriminant analysis of risk factors for sexual victimization among a national sample of college women. *Journal of Consulting and Clinical Psychology*, 57, 242-250.

Chapter 11: Measuring Change in Community Attitudes

- Koss, M., Leonard, K., Oros, C., & Beezley, D. (1985). Nonstranger sexual aggression: A discriminant analysis of the psychological characteristics of undetected offenders. *Sex Roles, 12*, 981-992.
- Lottes, I. (1988). Sexual socialization and attitudes toward rape. In A. Burgess (Ed.), *Rape and Sexual Assault II*. New York: Garland Publishing.
- Mahoney, E., Shively, M., & Traw, M. (1986). Sexual coercion and assault: Male socialization and female risk. *Sexual Coercion and Assault, 1*, 2-8.
- Marolla, J., & Scully, D. (1982). *Attitudes toward women, violence, and rape: A comparison of convicted rapists and other felons*. Rockville, Md.: National Institute of Mental Health.
- Mosher, D., & Anderson, R. (1984). Macho personality, sexual aggression, and reactions to realistic guided imagery of rape. Typescript.
- Pleck, J., Sonenstein, F., & Ku, L. (1993). Masculinity ideology and its correlates. In S. Osgamp & M. Costanzo (Eds.), *Gender Issues in Contemporary Society*. Newbury Park, CA: Sage.
- Russell, D. (1982). The prevalence and incidence of forcible rape and attempted rape of females. *Victimology, 7*, 81-93.
- Scully, D., & Marolla, J. (1984). Convicted rapists' vocabulary of motive: Excuses and justifications. *Social Problems, 31*, 530-544.
- Scully, D., & Marolla, J. (1985a). Rape and vocabularies of motive: Alternative perspectives. In A. W. Burgess (Ed.), *Rape and Sexual Assault: A Research Handbook*. New York: Garland Publishing.
- Scully, D., & Marolla, J. (1985b). "Riding the bull at Gilleys": Convicted rapists describe the rewards of rape. *Social Problems, 32*, 252-263.
- Stark, E., Flitcraft, A., & Frazier, W. (1979). Medicine and patriarchal violence: The social construction of a 'private' event. *International Journal of Health Services, 9*, 461-493.
- Straus, M. (1976). Sexual inequality, cultural norms, and wifebeating. *Victimology: An International Journal, 1*, 54-76.
- Weis, K., & Borges, S. (1975). Victimology and rape: The case of the legitimate victim. In L.G. Schultz (Ed.) *Rape Victimology*. Springfield, Ill: Charles C. Thomas.

Evaluation Guidebook

Williams, J. (1984). Secondary victimization: Confronting public attitudes about rape.
Victimology, 9, 66-81.

CHAPTER 12

MEASURING PERCEPTIONS OF JUSTICE

Many STOP projects are trying to change the ways that justice system agencies treat women victims of violence. These changes range from interpersonal behavior such as respectful listening through significant changes in skills and procedures applied to cases involving violence against women. Examples of the latter are better evidence collection, better communication with victims about the status of their case, better prosecution strategies, better monitoring and supervision of offenders, more attention to victims' wishes and circumstances in case disposition and sentencing, and so on.

Most, if not all, of the changes just described should affect how victims feel about participating in the justice system, and whether they feel that "justice has been done" in their case. Therefore, you may want to include measures of perceptions of justice in your evaluation, to let you know whether system changes have produced a better feeling about the system among victims. Your evaluation can measure perceptions of justice by the victims or parties in a legal action. You can also develop a rating of the extent to which the justice intervention is consistent with the rights of victims.

PERCEPTIONS OF JUSTICE

The first step in measuring perceptions about justice is to decide whether you want to measure (1) distributive justice, (2) procedural justice, or (3) both.

Distributive Justice

Distributive justice refers to perceptions of the outcome of police, court or other justice procedure (regardless of how those outcomes were arrived at). In rating the justice of an outcome, research indicates that people base their perceptions on how they answer two questions:

- Did the punishment fit the crime?** Measures would assess the extent to which the disposition of the case (or police call, hearing, or other intervention) resulted in a punishment that fit the offense.
- Is the outcome likely to reduce subsequent offenses?** Measures would assess the extent to which the disposition of a case, incident, or complaint reduces future harm to the victim or society.

Procedural Justice

Procedural justice refers to how the system acted, how it processed a case regardless of the outcome that resulted. It involves several components, including the consistency, fairness, and appropriateness of rules and procedures, and how those in authority treated the participants.

- ❑ **Consistency, fairness, and appropriateness of rules and procedures** used in handling the case or grievance. In general, research finds that people think something is fair or appropriate when they are able to tell their side of the story and they also perceive that all parties in a case or grievance have had a chance to tell their side of the story. This aspect of fairness has been measured by asking about the extent to which the parties have control of the process or input into the process (Shapiro & Brett, 1993; Tyler, 1994).
- ❑ **How participants are treated** by those in positions of authority during the process of handling the case or grievance. Two factors are particularly important in determining whether someone will feel that procedures are just (1) how/whether the intervention process/rules are explained to participants; and (2) how participants are treated during the intervention. Items used to measure this aspect of procedural justice usually ask people what they think of the third-party decision maker (e.g., the judge, police officer, prosecutor)—are they understanding, are they neutral/impartial, are they trustworthy, do they show consideration to those involved in the case (Tyler, 1989, 1994)?

Somewhat surprisingly, ratings of procedural and distributive justice appear to be independent of each other. Participants may rate the outcome as just, but not the procedure—or the procedure fair, but not the outcome. When individuals are forced to choose between the two, some research indicates that the higher ranking is given to achieving a just outcome (Wagstaff & Kelhar, 1993).

The following items and scales draw on the work of Tyler (1989), but include some new wording and answer categories. The scales ask separately about police and courts and can be extended to other justice agencies such as prosecutors' offices.

Procedural fairness items use 5-point scales and include the following:

- ❑ Were the procedures used by the authorities (police, courts) very fair (1), somewhat fair (2), neither fair nor unfair (3), somewhat unfair (4), or very unfair (5)?
- ❑ Was your case treated very fairly (1), somewhat fairly (2), neither fairly nor unfairly (3), somewhat unfairly (4), or very unfairly (5)?

Perceptions of fair personal treatment (by police/courts/others) use “yes” and “no” as answer categories; items include the following:

- Were you angry?
- Were you frustrated?
- Were you pleased (with the authorities)?

General fairness of treatment by (police/courts/other) items, rated on 5-point scales, include the following:

- Do you think (police/courts/other) usually treat people and handle their problems very fairly (1), somewhat fairly (2), neither fairly nor unfairly (3), somewhat unfairly (4), or very unfairly (5)?
- Do the (police/courts/other) treat people fairly almost always (1), most of the time (2), some of the time (3), not very often (4), or never (5)?

Distributive justice (the fairness of the outcome) items, also on 5-point scales, include the following:

- Was the outcome of your case very fair (1), somewhat fair (2), neither fair nor unfair (3), somewhat unfair (4), or very unfair (5)?
- Did you receive what you deserved from the (police/courts/other)? Strongly agree (1), Agree (2), Neither agree nor disagree (3), Disagree (4), Strongly disagree (5).
- If you were to deal with the (police/courts/other) in the future, do you think you would be treated very fairly (1), somewhat fairly (2), neither fairly nor unfairly (3), somewhat unfairly (4), or very unfairly (5)?

PROTECTION OF VICTIM RIGHTS

You can develop a rating of justice by asking women victims of violence to tell you the extent to which they have been treated according to the principles contained in the Victims’ Bill of Rights that is the law in many states. This Bill of Rights lists certain legal remedies and protections that should be extended to victims. Not all states have this legislation, but your state may. You may also want to include rights accorded victims in other states. Rights mentioned in these bills include, among others:

- The right to restitution for financial losses;

- The right to have the person who caused the harm detained, tried, and sufficiently punished;
- The right to have relevant evidence presented in court;
- The right to be protected from further harm or reprisal;
- The right to be treated with respect, dignity, and fairness;
- The right to address the court at sentencing;
- The right to timely notification of proceedings;
- The right to notification that the perpetrator has been released or escaped from custody;
- The right to confer with the prosecution;
- The right to costs of medical or psychological treatment for injury inflicted during a crime;
- The right to confidentiality of records;
- The right to the prompt return of property seized as evidence; and
- The right to notification of all legal remedies.

This list can serve as a starting point for a scale of the justice victims receive when a crime (sexual assault or domestic violence) occurs by asking victims whether their rights were observed in these ways. The list can be tailored to the program you are evaluating. You can further clarify or define these items and/or add other rights to the list. Answer categories can be yes/no or ratings of satisfaction using five-point scales similar to those used above. However, items in a single scale should all use the same number of answer categories (two or five).

ADDENDUM: READINGS ON PERCEPTIONS OF JUSTICE

- Casper, J.D., Tyler, T.R., & Fisher, B. (1988). Procedural justice in felony cases. *Law & Society Review*, 22, 403-507.
- Feld, B.C. (1990). The punitive juvenile court and the quality of procedural justice: Disjunctions between rhetoric and reality. *Crime & Delinquency*, 36, 443-466.
- Folger, R., Cropanzano, R., Timmerman, T.A., Howes, J.C., & Mitchell, D. (1996). Elaborating procedural fairness: Justice becomes both simpler and more complex. *Personality and Social Psychology Bulletin*, 22, 435-441.
- Gilliland, S. W. (1994). Effects of procedural and distributive justice on reactions to a selection system. *Journal of Applied Psychology*, 79, 691-701.
- Lind, E.A., & Tyler, T.R. (1988). *The Social Psychology of Procedural Justice*. New York: Plenum Press.

- Miller, J.L., Rossi, P.H., & Simpson, J.E. (1991). Felony punishments: A factorial survey of perceived justice in criminal sentencing. *The Journal of Criminal Law & Criminology*, 82, 396-422.
- Shapiro, D.L., & Brett, J.M. (1993). Comparing three processes underlying judgments of procedural justice: A field study of mediation and arbitration. *Journal of Personality & Social Psychology*, 65, 1167-1177.
- Stroessner, S.J., & Heuer, L.B. (1996). Cognitive bias in procedural justice: Formation and implications of illusory correlations in perceived intergroup fairness. *Journal of Personality & Social Psychology*, 71, 717-728.
- Tyler, T.R. (1994). Psychological models of the justice motive: Antecedents of distributive and procedural justice. *Journal of Personality & Social Psychology*, 67, 850-863.
- Tyler, T.R. (1989). The psychology of procedural justice: A test of the group-value model. *Journal of Personality & Social Psychology*, 57, 830-838.
- Vingilis, E., & Blefgen, H. (1990). The adjudication of alcohol-related criminal driving cases in Ontario: A survey of crown attorneys. *Canadian Journal of Criminology*, 639-649.
- Wagstaff, G.F., & Kelhar, S. (1993). On the roles of control and outcomes in procedural justice. *Psychological Reports*, 73, 121-122.

CHAPTER 13

MEASURING THE IMPACT OF TRAINING

By Heike P. Gramckow, Ph.D., Jane Nady Sigmon, Ph.D.,
Mario T. Gaboury, J.D., Ph.D.⁸, and Martha R. Burt, Ph.D.

It is hard to imagine any training in the area of violence against women that does not seek, at a minimum, to increase knowledge, heighten sensitivity toward victims, and enhance skills that will help trainees contribute to efforts to protect victims. It does not matter whether those attending training are police officers, prosecutors, judges, victim advocates, hospital personnel, or probation officers—these basic goals will be the same even when the content of the knowledge and skills to be imparted will differ. Thus every STOP-funded training project will need to review this chapter, because every trainer needs feedback indicating whether the time and effort invested in training has any payoff in learning and in changed attitudes and behavior.

LOGIC MODEL OF A TRAINING PROJECT

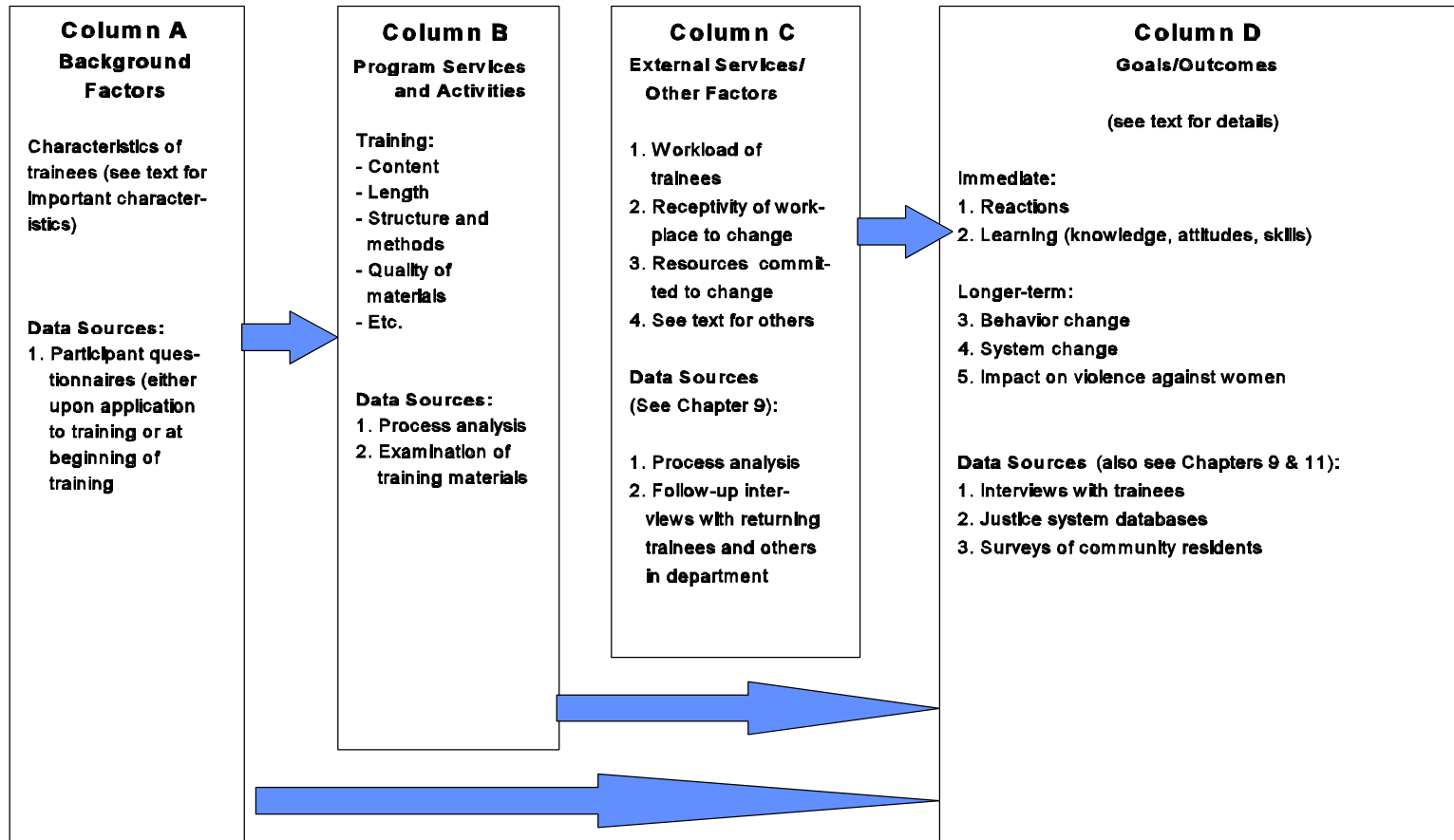
Underlying all impact evaluation is a common concept described in more detail in Chapter 6. Determining impact requires comparing the conditions of individuals who have experienced an intervention (e.g., a training program) with individuals who have experienced something else. We identify the impact of training by comparing data on participants and non-participants, or by measuring participants before and after an intervention (and also possibly while the intervention is in progress), or by other methods of comparison.

This chapter addresses issues in evaluating a particular type of project (rather than offering outcome measures relevant to a variety of projects), so it is a good place to practice developing a logic model. Exhibit 13.1 shows a schematic logic model, without a lot of detail. Column A contains background factors which, in this case, will be a variety of characteristics of the trainees (which characteristics will depend on the content of training). Column B shows characteristics of the training itself, which should be documented with the techniques of process evaluation. This is an essential step; process evaluation has inherent value in illuminating whether the training operated as intended, and is necessary to understand the results of an outcome evaluation.

⁸ Heike P. Gramckow, Director of Management and Program Development, and Jane Nady Sigmon, Director of Research, both from the American Prosecutors Research Institute, are working with others to complete an implementation manual for coordinating councils on violence against women (funded by the Bureau of Justice Assistance), and developing a curriculum to train prosecutors to handle domestic violence and sexual assault cases (funded by the Violence Against Women Grants Office). Mario T. Gaboury is the Director of the Center for the Study of Crime Victims' Rights, Remedies, and Resources at the University of New Haven. Jane Nady Sigmon serves as an Advisor to the Urban Institute for the National Evaluation of the STOP Formula Grants Program.

Exhibit 13.1

TYPICAL TRAINING PROGRAM



Column C identifies some common external factors that can enhance or undermine the impact of training. As with background factors, which ones are important in your evaluation will depend on the content of training, the origins of the trainees, and the circumstances that prevail in the trainees' home agencies and home community. Column D identifies the goals or outcomes of training, which are described in greater detail throughout this chapter. Note that the data source anticipated for Columns B and C is process analysis.

SELECTING OUTCOMES TO EVALUATE

The ultimate goal of VAWA-supported training is improved outcomes for women victims of violence. However, few training evaluations will be able to take this long-term view. The more immediate and intermediate outcomes such an evaluation *can* measure are those that facilitate or promote the ultimate goal. Immediate outcomes include new knowledge, changed attitudes, and new skills. Intermediate outcomes may be changes in the behaviors of those who received training, putting their new learning and attitudes into practice. More long-term but still not ultimate outcomes may include changes in police or court procedures, the adoption of new policies that affect the treatment of victims by criminal justice or allied professionals, or the development of a coordinating council that connects all community services for victims. The contents, and sometimes the structure, of training will be shaped by which of goals the training aims to achieve. Likewise, any evaluation must structure its measures to be able to assess the different goals and activities of the training program.

Avoiding Inappropriate Outcomes

Any evaluation of training should cover only outcomes that the training itself can be expected to accomplish. When the training's main goal is to increase the knowledge of its participants, then the evaluation should concentrate on assessing knowledge acquisition. While a number of other outcomes may result from the training, they cannot be the focus of an evaluation if the training itself did not posit that its goals included these other outcomes. It would be an odd training program that wanted only to convey knowledge and did not also hope or expect that participants would use the new knowledge in some way (i.e., that they would change some behaviors as a consequence of what they have learned). But suppose for the moment that this is the case. Then an evaluation would stick to assessing knowledge. However, when the training program's logic model includes the expectation that new knowledge will lead to different behaviors, and the program's training curriculum is set up to encourage changed behaviors and give people time to practice them, then the evaluator should ask whether the training participants use the new behaviors once they are back on the job.

Evaluators also need to be careful not to expect training programs to achieve ambitious goals that clearly require more than just training. When police officers participate in training that teaches them how to work closely with victim advocates in order to offer better protection to women who

have been assaulted, we would look for increased and improved coordination between police officers and appropriate victim advocates. But can we also expect to see that female victims of domestic assaults are actually better protected? The latter may be a long-term goal of the training effort that very likely requires the availability of additional resources (e.g., shelter, protection orders, electronic warning systems). While the training may actually teach police officers how to access these resources, it cannot make them available in each community. When training alone cannot, realistically, achieve a stated goal, then the hoped-for impact cannot be expected to occur and should not figure centrally in an evaluation of training.

Picking the Right Outcomes

Developing the right measures for goal achievement becomes more involved the more comprehensive the goals of training are. Usually the results of training that aims only to provide knowledge are relatively easy to measure. For example, when the training curriculum includes a session on new legislation that changes the situations in which protective orders may be issued or a session on new techniques for gathering evidence in sexual assault cases, participants should be assessed on their ability to describe the situations or techniques presented in training. When the training also has the goal of providing the skills to apply the new law or the new techniques and includes sessions for mastering these skills, then the evaluation should include an assessment of how much the skills are used in practice.

Specifying and collecting the right evaluation measures for assessing skill acquisition is more difficult than for knowledge acquisition, because one must collect evidence of changed behavior. This may come through self-report from a follow-up questionnaire, from the testimony of others (e.g., interviewing co-workers and supervisors about the trainee's use of the new skills), from actual observation, or from analysis of case records (assuming that the records contain adequate information). Obviously, it is more expensive and time-consuming to gather these types of information to document behavior change (or lack of it) than it is to give trainees a paper-and-pencil knowledge test on the last day of training. However, the importance of what you learn from your increased investment usually more than makes up for the additional trouble.

If training is to achieve its goals, the substantive information actually provided in the training session must be linked to those goals. Suppose that the goal of a training program is to increase the number of protection orders that officers obtain and the training provides law enforcement officers with knowledge of the legal requirements for obtaining protection orders, but does not teach them how and where to obtain the protection order. Then, considering the content of the training sessions, the desired result is not likely to occur, or if this result occurs, it is not likely to be a direct result of the training.

Also, the training must be implemented well enough that its critical elements have been delivered, and to the right people. If, for example, one section of a training program is not taught, you cannot expect to see the same results as if trainees had received the entire program. The missing

component has to be identified and accounted for since it is only logical to assume that only a completely implemented effort can have the expected impact. If a training component is missing this may actually be an explanation for why the assessment results indicate that the program did not have the predicted impact.

LEVELS OF EVALUATION FOR TRAINING PROGRAMS

You will need to gather several types of information before you can see whether a training is successful. First, you need information *from the trainers* about the goals and the content of training, to identify what was intended and what was delivered. Second, you need baseline information *from the trainees* about their level of knowledge, attitudes, and skills before the training. Finally, you need information *from the trainees* after the training to identify changes as a result of the training. In addition, training sessions related to violence against women will often include *victims, survivors, and advocates* as participants or presenters. These participants bring a unique perspective to a training session. They should be asked to give their perceptions of how well the other participants “got the message,” especially on the issues of attitudes and sensitivity toward victims.

A very widely accepted approach to evaluating training programs has been developed by Kirkpatrick (1996), who outlines four levels of measurement to address the range of potential training effects. Although many training evaluations focus on only one or two of these levels (reaction assessment and learning), he argues that to truly assess the outcome of training, one should gather data to measure all of the following:

- Reactions:** an on site measurement of immediate “customer satisfaction” with various aspects of the training program;
- Learning:** an assessment of changes in attitudes, increased knowledge, and/or improved skills;
- Behavior Changes:** an assessment of changes in participants’ performance or ability to apply learning; and
- Results, or Problem Impact:** an assessment of the impact or applied results of the training.

This model can be applied to any training regardless of the setting or topic. Each level of assessment can provide valuable information for developing, evaluating, and revising training curricula. However, assessment strategies such as the type of information to be gathered, timing of the assessment, and the method of gathering information are different at each level.

Table 13.1 shows Kirkpatrick’s four levels of evaluation, reaction, learning, behavior change, and problem impact (in our case, impact on VAW), and how they apply to the common training goals of knowledge and attitude change, skill acquisition, behavioral intentions and behavior change,

Table 13.1
Options for Evaluating Training Activities

Evaluation Level	Training Goals	Timing
Reaction	Perceived knowledge gained Perceived attitude change Perceived skills acquired New behavioral intentions	Immediate
Learning	Knowledge gained Attitude change Skills learned Plan in place for carrying out behavioral intentions	Before Immediate Short term Long term
Behavior Change	Skill use on the job Own behavior change beyond use of skills Changes in behaviors of others Changes in “behaviors” of organizations (see Chapters 8, 9, 10)	Short term Long term
Impact on VAW	Improved victim outcomes (see Chapter 7) Reduced levels of VAW in the community (see Chapter 11)	Long term

organizational change, and impact on victims and on violence. It also shows the most relevant timing for evaluation at each level, and the importance of taking direct and indirect external influences into account. *Immediate* assessments are those done during a training session or just before everyone goes home. *Short-term follow-up* assessments are done after trainees are back on the job for a while, but usually within the first two or three months, possibly up to six months, after training. *Long-term follow-up* assessments can be done months or even years after the training. Timing will depend on how long you think it will take for a particular impact of training to occur. Changes in trainees’ own behavior might be expected to occur relatively quickly (within the first few months), whereas changes in colleagues’ behavior might take up to a year to catch on, and institutional change could take several years (although there would be signs along the way, such as resource commitment, planning committees, protocol development committees, etc.). We discuss each level of evaluation shown in Table 13.1, including the measures that are most commonly used in relation to each training goal for each level.

Level 1: Reaction Assessments

The most basic level of impact evaluation focuses on capturing immediate, on-the-spot reactions to different aspects of the training, providing a measure of “customer satisfaction.” However basic, it is an important method of obtaining information that will help training organizers improve training programs for future presentations.

The information gathered usually asks for the trainee’s judgment about the content of each session, the materials used, the instructor’s ability to present the material, the relevance of the material to the trainee’s job, and the training facility. Examples of “reaction assessment” questions, to be answered on a scale from 1= strongly disagree through 5 = strongly agree follow:

1. The content of the session was relevant to my professional needs.
2. Overall, the information provided in the session was practical.
3. The time allocated to the subject was adequate.
4. The session enhanced my knowledge of the subject.
5. The instructor was very good at getting his/her points across.
6. The room where the session was held was adequate (big enough, warm/cool enough, appropriate arrangement of tables and chairs, easy to see the board/screen/flip-chart, etc.).
7. I learned a lot that I needed to know.
8. The session changed the way I will think about (_____) in the future.

All evaluation at the level of reaction assessments is immediate. One can assess reactions to the presentation of materials, regardless of what the materials were trying to do (that is, whether they were trying to convey knowledge, change attitudes, change behavior, etc.). One can assess trainee’s beliefs that they learned something new, their perceptions that efforts to change their attitudes or feelings about a topic succeeded or failed, their beliefs that they have acquired new skills, and their intentions to change their behavior once they get back on the job.

Even with reaction assessment, background and external factors (Columns A and C of your logic model) can play a role. To understand how these factors might affect reactions to training, you would also gather information about the ones you thought might make a difference and analyze your results against these factors. Some background and external factors will need to be collected through questionnaires completed by trainees (including their application for training), while others will be obvious to anyone who attended the training. For reaction assessments, the major background and external influences one would want to guard against are:

- Aspects of the trainees’ own background, prior exposure to the field and the material, and work situation; and
- Major disruptions in training (e.g., a power failure meant that two-thirds of the sessions were canceled; an outbreak of Legionnaire’s Disease meant that half of

the participants were sick or getting sick; the training materials were not delivered until the most important sessions were finished).

Since a reaction assessment gathers immediate perceptions and feelings, it is usually conducted during the training and immediately prior to the trainees' departure by distributing and collecting training evaluation questionnaires to be completed by all participants. These questionnaires can be done for each session, for each day, and for the training program as a whole, depending on the level of detailed feedback you want to receive. A *session* questionnaire should assess immediate reactions to content and presentation of each individual session, including the adequacy of opportunities to practice new learning if that is one of the goals of training. A combination of forced-choice (e.g., yes/no, agree/disagree) and open-ended questions that give people space to write in comments will facilitate participant feedback. A *daily* questionnaire should assess reactions to the balance of types of sessions (plenary, didactic, practicum, coalition-building, etc.), order of sessions, quality of presenters, and so on. When the training extends beyond one day, an *overall* training evaluation questionnaire should be used to gather participant assessments of the training program as a whole, the quality of written materials, format, schedule and organization of the course, trainers, and facilities. If the session is one day or less, this information should be requested on the *daily* questionnaire, which will serve also as the *overall* questionnaire. In an effort to encourage full and honest participation, such evaluations are typically submitted without names or other identifying information.

While one usually gets many completed questionnaires from a reaction assessment, it provides only limited information on the effect of the training. Its findings generally do not provide information about goal achievement, because a reaction assessment cannot capture even the most basic impacts such as the amount of learning or attitude change that occurred. Nor can it register how much of the newly acquired knowledge actually stays with the participant over time or how well it can be and has been applied in the field.

Reaction assessments are very much like the little card one is often requested to complete in a restaurant to indicate how much one liked the services and food. You can indicate being pleased with or displeased with the facility, the waiter, and the food, but there is no opportunity to add comments when the food did not agree with you several hours later. Some feedback of this immediate type is important but it cannot give the whole picture, nor does it capture the most important potential impacts of training. Immediately following training, participants may have the feeling that they learned much and may indicate this in a questionnaire administered on site. This result reflects the participants' perceptions of their learning, but the reaction assessment does not actually test that learning, nor does it assess whether the learning is applied on the job. Thus, a reaction assessment cannot adequately evaluate whether a training program accomplished any of its goals (beyond merely having gotten through the training day).

Level 2: Learning Assessments

An essential component of a training evaluation is an assessment of what participants learned. New knowledge and new attitudes are always primary training goals, and are also the basis on which trainers expect that trainees may change their behavior in the future.

Learning can be measured by assessing changes in participants' knowledge, attitudes, or skill levels from immediately before to some time after a training program. Such pre- and post-test measures of how well participants understood the material covered in the training program provide an objective indication of whether learning took place. Ideally, pre- and post-test assessments would be conducted with a group of training participants and with a comparison group who did not participate in training. By examining the differences between the two groups, evaluators can assess on the post-test whether participation in the training program improved participants' performance. However, few training programs have the funds and staff to use comparison groups. As a result, learning assessments are generally limited to those who participate in the training.

As mentioned earlier in this chapter, for training related to violence against women in particular, attitude change and increased sensitivity to victims must be an essential goal. There are some general points of sensitivity that should be included in any STOP-funded training evaluation, including the following:

- Behavior and language should reflect *a belief in what the woman is telling you*;
- Treat the woman *with respect*;
- Don't make disparaging remarks or jokes;
- Show by your behavior and remarks that you *take her situation seriously*;
- Protect the woman's privacy*, to the extent possible in the situation.

Questions assessing attitudes can be constructed to fit the specific work situations of the trainees. For example, in a police training focused on new laws specifying mandatory arrest of batterers, important attitudes to address in both training and evaluation are those which have been shown to reduce the likelihood that police will make an arrest (Ford, 1987). These include ideas about the seriousness and criminality of battering, whether some women are "unworthy" of police protection, whether women have acted on their own behalf, and whether battering is a "private matter." One can assess these attitudes on 5-point scales (e.g., 1=agree strongly, 2=agree somewhat, 3=neither agree or disagree, 4=disagree somewhat, 5=disagree strongly; or that a factor, if present in the situation, 1=strongly favors arrest, 2=somewhat favors arrest, 3=irrelevant to arrest, 4=somewhat against arrest, 5=strongly against arrest). The first group of items shown below uses the agree-disagree scale; the second group uses the favors-against arrest scale:

- Misdemeanor battery is too minor an offense to deserve arrest.
- A man does not deserve to be arrested for slapping his wife.
- If a woman does not leave a man who beats her she should not expect the police to protect her.
- It takes two to fight so one cannot blame the man alone when he hits his wife.
- It is best to arrest both parties in order to let the court decide who is at fault.

- The woman is highly agitated [as a reason *not* to arrest].
- The police had been to the residence before [as a reason *not* to arrest].
- The woman went out drinking against her husband's wishes.
- The woman does not want to prosecute.

The same types of attitudes are also relevant to prosecutors' decisions not to prosecute, judges' decisions to reduce sentences or waive punishment, and other actors' decisions about whether or not to support a battered woman. Parallel attitudes affect the decisions of police, prosecutors, judges, juries, and potential support people in cases of sexual assault. Attitude scales related to both sexual assault and battering are described in Chapter 11, where you can see other examples of attitudes that are important to include in training and to measure because they make a difference for people's behavior toward women victims of violence.

To assess learning, the content of a pre- and post-test evaluation instrument must be completely consistent with the training course content. The questions on the questionnaire should match the most important elements of the information that the training organizers were trying to convey. The most typical format is a combination of true/false and multiple choice questions that are based on material covered in the course. It is best to include both fact-based questions and questions that require the application of information to a problem situation. If attitude change was one of the training goals, questions should cover current beliefs related to the topics about which trainers hoped to change attitudes. If skill acquisition is a training goal, paper-and-pencil measures may not be enough to measure training success. In addition to participants' self-reports that they learned new skills, evaluators may wish to set up situations in which they can actually observe trainees to see whether they are using the skills. This can be done during training in practice sessions and also is good to do in the participants' regular job situation at some time after training. Finally, paper-and-pencil assessments toward the end of a training program can ask participants about their *behavioral intentions*—what do they intend to do when they get back to their jobs, how is it different from their usual behavior, and how do they plan to introduce these new behaviors into the workplace.

Timing is a critical dimension of learning assessments. You will want to collect data before training to assess "pre" levels of knowledge, attitudes, and skills. You will also certainly want to collect data at the end of training, to provide the easiest form of "post" evaluation data. However, bear in mind that "after" training covers a lot of time. It is very important for a training evaluation to try to include some assessment "down the road," as the effects of knowledge and attitude change are known to wear off quite rapidly if they are not quickly reinforced with

behavior change. So a good evaluation of training should try to assess *learning retention* two or three months after training, and should also track whether by that time participants have changed their behavior in any way that fulfills the behavioral intentions they stated at the end of training.

Since the measurement of learning is linked to the topic presented in a course, few measures can be applied to every training evaluation. We can, however, give an example in which judges were taught the contents and implications of a state's victim rights legislation. Measures of increasing levels of learning and how they can be established for this subject matter cover the following (Bloom, 1956):

- ❑ **Knowledge:** Training participants are able to quote the rule/legislation that victims have the right to speak at the sentencing hearing.
- ❑ **Comprehension:** Participants are able to summarize the rule/legislation that states that victims have the right to speak at sentencing hearings.
- ❑ **Application:** Participants see a video of a standard sentencing hearing and examine all methods that were followed and discover opportunities that were missed, that would have allowed the victim to speak.
- ❑ **Analysis:** Participants can illustrate how the victim may be allowed the opportunity to be heard in a sentencing hearing based on a scenario in which, for example, a victim who is severely brain damaged and, as a result, unable to speak.
- ❑ **Synthesis:** Participants devise a new rule for victim participation when victims are unable to speak for themselves.
- ❑ **Evaluation:** Participants are able to assess the adequacy of methods used to allow victims the right to be heard at sentencing hearings.

If the purpose of a training program is to increase the skills of participants in a particular area, a performance-based pre- and post-test evaluation may be utilized on site to collect immediate information about the effectiveness of the training and to provide immediate feedback to the trainees. Evaluators using this approach can assess learning, attitude change, and behavior change as a result of training. For example, if crisis intervention hotline volunteers are being trained to answer hotline calls, one method of evaluating their skills acquisition and proficiency in using them would be to have each trainee at the beginning of training handle a number of mock phone calls that present a variety of crises, and do this again at the completion of training. A trainer could observe and rate the responses of the trainee. In this instance, the actual performance of handling a representative range of phone call scenarios would provide trainers with valuable information about the individual trainee's learning and level of skill development, as well as strengths and weaknesses in the training program.

Pre- and post-test learning assessments generally are not anonymous, because they often are used to provide specific feedback to individual training participants. If you do not want or need to give

this type of feedback, anonymity can be preserved by assigning an identification number to each trainee and attaching the number to pre- and post-test learning assessments and any other follow-up assessments.

Background and external factors (Columns A and C of your logic model) may affect learning from a training program. Background factors—factors the trainees bring with them—are the only ones likely to affect immediate assessments. You may want to collect data from trainees through registration materials or a background sheet completed at the beginning of training to help you assess whether background, pre-training knowledge, or existing skills influenced the impact of training. Both background factors and external factors can influence learning assessments taken at some time after the training ends. The persistence of a supervisor's or work group's mistaken knowledge or negative attitudes may overwhelm any attempts of trainees to put new learning into place, just as lack of resources in the community may hinder trainees from establishing new cooperative or referral behaviors. An evaluation should try to learn about as many relevant external influences as possible. You can often use trainees' perceptions to identify some of these, especially when it seems likely that there will be little long-lasting effect of training and you are asking them "why not?" In addition, you could try to get "outside" views of "why not?" by asking other informants in the workplace, department, or community.

Level 3: Behavior Change Assessments

Assessments of behavior change try to identify whether the knowledge gained during the training is actually applied later, so of course these changes must be measured some time after the training takes place. It is also important to point out that evaluations of behavior change should build on previous assessment levels. If the learning assessment, for example, indicates that trainees did not learn very much, it is unlikely that behavior will change significantly. If such a change occurs nevertheless, it is likely to be the result of some external factor. Also, if the learning assessment indicates that a high degree of learning occurred but the behavior assessment show little impact, the evaluator might look for external factors that prevent trainees from applying what they learned in training on their jobs.

Changes in behavior can be captured through pre- and post inquiries to the training participant, or to his or her peers, supervisor, or clients (e.g., victims). Table 13.2 gives an example of questions posed to police officers who received training about how to respond more effectively to calls for domestic disputes. The officers answered on a scale of 1 = never to 5 = almost always, and the results for the pre- and the post-training questions were later marked on the same questionnaire to create a visual display of areas of improvement.

Such an instrument is also helpful if behavior changes are to be observed, either in a mock situation or in the field. An observation guide with similar indicators for trainee behavior is provided to the observer (who could be a supervisor or researcher), who notes the rating before the training and after. The results are then compared to indicate changes.

Table 13.2
Comparing Responses Before and After Training

Indicator question	Indicator response pre-training					Indicator response post-training				
How often do you:	Never	Rarely	Sometimes	Often	Almost Always	Never	Rarely	Sometimes	Often	Almost Always
...encourage the victim to seek medical attention?	1	2	3 ✘	4	5	1	2	3	4 ✘	5
...explain to the victim and the assailant that the arrest decision is required by law?	1	2 ✘	3	4	5	1	2	3	4 ✘	5
... discuss a safety plan with the victim?	1 ✘	2	3	4	5	1	2	3 ✘	4	5

Another avenue for identifying the pervasiveness of behavior change following training is to look at institutional documents. Some of these may reveal whether the lessons of training have penetrated through to the daily life of trainees. For example, many police departments have formal field training programs for new officers (that is, after they finish with the academy, they receive further training in their own department in conjunction with assuming their role as police officers. This field training often is guided by manuals that specify certain skills that police must be able to demonstrate. An evaluator of a training program who wanted to see whether behavior had changed in the field could examine these field training manuals to see whether new training being given at police academies has permeated through to field training.

Level 4: Problem Impact Assessments

Indicators of a training program’s impact on the ultimate issues of violence against women, such as increased victim safety and reduced overall levels of violence against women, are the same as those described in other chapters of this *Guidebook*, especially in Chapters 7, 8, 9, 10 and 11, and are not repeated here. However, since the chain of reasoning that links training activities to these ultimate outcomes is very long and complex (as the logic model you develop for your training program should make clear), your main challenge as an evaluator will lie in establishing the connection between training provided and changes in a problem situation as revealed by the indicators you select. Your logic model will help you decide where you have to look for impact, and what steps and external influences you will have to document along the way. In general, it is not likely that you will be able to attribute any changes in these ultimate impacts to training, or to training alone. The resources it would take even to attempt this would probably be better spent on improving your ability to measure outcomes that are a bit more connected to the training itself.

Assessing External Influences

If, despite what we just said, you still wish to include measurement of ultimate impacts in your training evaluation design, you will have to do an excellent job of measuring external influences. Therefore, we have saved the greatest part of our discussion of these influences for presentation after considering Level 4 evaluations, although they also apply to the long-term assessments of behavior change in Level 3, especially if you want to assess changes in the behaviors of others or of institutions.

External influences might include *direct* factors such as the characteristics of the people who received the training. Adult learning theory recognizes the many types of learners and great diversity in education levels, life experiences, goals, abilities and disabilities among learners entering adult training (Thermer, 1997). The different levels of knowledge that participants bring to training have an impact on how well they will be able to apply what is taught when they are back on the job, and as a result, influence to what extent a training program has the desired effect.

Other external influences include *indirect* factors such as other training efforts, resources to implement what is learned, resistance or cooperation from others to implementing what is learned, financial and legal barriers to implementing what is learned, and so on. Our earlier example of a training program whose long-term goal was to increase victim safety through better cooperation between police and victim advocates is a case in point. Even if cooperation actually improves dramatically, the ultimate goal of increased victim safety may not be reached because other resources to accomplish this goal were not available or were reduced.

It is also important to document external influences that operate simultaneously with training to produce the ultimate outcome. Again using our example, the training to increase cooperation may not have been very effective, but shortly after the training took place the police department and victim service agency issued guidelines requiring officers and victim advocates to work closely together, and supervisors in both agencies made every effort to implement this new guideline. Evaluation would most likely show that the goal of training was achieved, but it would be wrong to attribute to training more than a facilitative role in providing the skills and attitudes that helped individual officers and advocates make this transition more smoothly.

The following list, while not exhaustive, shows many frequently observed external influences that would be considered to be indirect effects in evaluations of training impact:

- Change in policies or procedures,
- Change in legislation,
- Change in organization or management,
- Change in staffing,
- Change in resources,
- Demographic changes in the community,
- Other training provided, and

- ❑ Information campaign conducted.

OTHER ISSUES

Timing of Training Impact Evaluations

Depending on the type of training program, its impact can be evaluated continually or periodically. An important consideration in determining when a training evaluation needs to occur is the time in which the impact can realistically be expected. You need to consider that some goals can only be achieved over time. If we again look at the example of a training program that promotes learning designed to increase law enforcement's cooperation with other agencies so that women victims will receive better protection, it is obvious that the first goal, learning, is the one you can assess closest to the time of training. Increased and improved cooperation is a goal that is more likely to be achieved a bit later, but still in a shorter time than the goal of greater protection for female victims.

To identify whether impact lasts a long time, you have to do your assessment a while after the training occurred. Usually such post-tests are scheduled within 6 months to one year after the training occurred. The later the assessment is conducted, the higher the likelihood that the impact of the training is "contaminated" by too many other influences (e.g., on-the-job experience) and the results can no longer be directly related to the training. It also becomes more difficult to track the participants, especially in law enforcement settings where assignments are rotated frequently or in settings with high turnover. In one study to assess the impact of a training program to assist law enforcement and prosecution in developing an asset forfeiture program, almost 40 percent of the law enforcement staff participating in the training effort had moved into other assignments at the time of the one-year follow-up training assessment.

When you should conduct a post-training assessment to assess training impact also depends on the type of training provided and the goals to be achieved. Long-term goals can only be assessed on a long-term basis. Among the training programs funded under the VAWA STOP grants this caveat about the timing of assessment applies especially to training efforts to increase cooperation among different agencies, which is a longer-term goal requiring sufficient time to materialize.

Methods to Evaluate Training

We have already made a number of suggestions about methods for assessing training impact at each level of evaluation. Selecting the appropriate method of evaluating VAWA-related training involves technical issues such as how best to measure changes in attitudes, knowledge, behaviors, and actual impact on a specific problem. As with any program evaluation, training impact can be measured through surveys, observations, focus groups, and analysis of agency records. Method selection will also depend on practical issues such as budget constraints and timelines.

Traditionally, how training participants react to training, what they learned, how they changed attitudes or behavior, and to what extent a specific problem was affected by a training session is measured through a written or oral survey. This survey can elicit information from the individual being trained, and also from others with whom the training participant interacts such as a supervisor, a co-worker, or a woman victim of violence (a client).

Another way to evaluate training impact is to conduct process evaluation. This would include observing the participant in real-life situations, where you can see how well tasks that were addressed in the training are actually handled. This observation can be made by an independent researcher or expert, or by a peer or supervisor. Another mechanism is a focus group with supervisors, co-workers, or clients to identify whether the training had the desired impact. These mechanisms provide qualitative information.

Other assessment approaches can focus on collecting quantitative data from agencies involved in violence against women issues (law enforcement, prosecution, victim services, courts, and others). Such data can be meeting records (if more coordination was a goal of the training), agency case records (if increased numbers of protection orders was a goal of the training), and other similar “hard” data. Both qualitative and quantitative data collection have their benefits and drawbacks. Generally, a combination of both provides the most complete picture.

Using Training Evaluation Results to Improve Services

Chapters 4 and 5 discussed in general some of the ways that programs can use evaluation data to improve their own agency functioning. In this chapter we want to focus on specific applications of training evaluations to improve programs and services for women victims of violence, including monitoring training progress, adjusting the subject matter or presentation style to the target audience, and justifying further use of resources for training. To ensure that a training program is up-to-date and continues to deliver the information required to the right audiences a number of steps should be implemented, including (1) revising training goals and evaluation measures, (2) developing a feedback loop, and (3) implementing a systemic change mechanism (Levin, 1974).

For a training program to deliver the needed services, training providers continually need to assess whether the goals, techniques, and materials of the training program are still adequate to the needs of the target population. If the goals, techniques, or materials change, the measures to identify impact need to be revised accordingly. To ensure that the training continues to be relevant, it is helpful to establish an ongoing mechanism to receive feedback not only from those who are currently participating in training, but from those who have participated in the past, those who are potential training participants in the future. It might also be informative to include those who benefit from the training indirectly, such as co-workers and supervisors, and the clients who actually receive the services that the training is trying to improve.

All of the information collected to identify training impact, to revise the training program, and to receive continuous feedback on training needs and the adequacy of the training provided is helpful to ensure that the training delivered is the training needed. There is, however, a larger issue that needs to be addressed to ensure that training resources are spent efficiently and effectively—the issue of ensuring that the knowledge and skills learned in a training course can actually be applied. Training providers need to identify factors that prevent trainees from using what they learned and possibly incorporate into training ways that the trainees may try to overcome these obstacles. Evaluation can help you identify these external influences and may possibly point the way to how they can be handled.

ADDENDUM: READINGS ON TRAINING EVALUATIONS

Other fields have developed an extensive body of literature regarding the evaluation of training programs that may be helpful to criminal justice professionals. These include business and industry (Burkhart, 1996; Ezzeddine & Holand, 1996; Holcomb, 1993; Jackson, 1989; Rothwell, 1994); adult education (Gailbraith, 1997; Lea & Leibowitz, 1992; Moran, 1997); vocational programs (Strecher & Hasner, 1992, 1993, 1995); and literacy (Leef & Riddle, 1996). The following references provide some resources for those interested in further reading on training evaluations (see also the addendum to Chapter 6 for readings on general evaluation issues).

Basarab, D. Sr., & Root, D. (1992). *The Training Evaluation Process*. Norwell, MA: Kluwer.

Bloom, B.S. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I: Cognitive Domains*. New York: McKay.

Brinkerhoff, R. (1968). *Achieving Results from Training*. San Francisco: Jossey-Bass.

Burkhart, J. (1996). *Evaluating Workplace Effectiveness*. Denver: Colorado State Department of Education, State Library and Adult Education Office.

Dugan, M. (1996). Participatory and empowerment evaluation: Lessons learned from training and technical assistance. In Fetterman, D.M., Kaftarian, S.J. and Wandersman, A. (Eds.), *Empowerment Evaluation: Knowledge and Tools for Self-assessment and Accountability*. Thousand Oaks, CA: Sage.

Ezzeddine, A., & Holand, K. (1996). How we developed our professional development programs: Two models and why they work. *Small Business Forum*, 14(2), 40-53.

Ford, D. (1987). The impact of police officers' attitudes toward victims on the disinclination to arrest wife batterers. Paper presented at the Third National Conference for Family Violence

Evaluation Guidebook

- Researchers, University of New Hampshire, July 1987. Indianapolis: Indiana University at Indianapolis.
- Gailbraith, M. (1997). *Administering Successful Programs for Adults: Promoting Excellence in Adult, Community, and Continuing Education*. Professional Practices in Adult Education and Human Resources Development Series. Malabar, FL: Krieger Publishing Co.
- Holcomb, J. (1993). *Making Training Worth Every Penny*. Playa del Rey, CA: On Target Training.
- Jackson, T. (1989). *Evaluation: Relating Training to Business Performance*. San Diego, CA: University Associates.
- Kirkpatrick, D.L. (1986). *More Evaluating Training Programs*. Alexandria, VA: American Society of Training and Development.
- Kirkpatrick, D.L. (1993). *How to Train and Develop Supervisors*. New York: AMACOM.
- Kirkpatrick, D.L. (1996). *Evaluating Training Programs: the Four Levels*. San Francisco: Berrett-Koehler Publishers, Inc.
- Lea, H.D., & Leibowitz, Z.B. (1992). *Adult Career Development: Concepts, Issues, and Practices*. Alexandria, VA: National Career Development Association.
- Levin, H.M. (1974). A conceptual framework for accountability in education. *School Review*, 82(3), 63-391.
- Leef, C. & Riddle T. (1996). *Work Together, Learn Together: a Model That Incorporates Community Partnerships*. Peterborough, Ontario, Canada: Literacy Ontario Central South.
- Moran, J. (1997). *Assessing Adult Learning: A Guide for Practitioners*. Professional Practices in Adult Education and Human Resources Development Series. Malabar, FL: Kreiger Publishing Co.
- Phillips, J. (1991). *Training Evaluation and Measurements Methods*, 2nd Edition. Houston: Gulf.
- Rothwell, W. (1994). *Beyond Training and Development: State-of-the-Art Strategies for Enhancing Human Performance*. New York: AMACOM.
- Strecher, B.M., & Hasner, L.M. (1992). *Local Accountability in Vocational Education: A Theoretical Model and its Limitations in Practice*. Santa Monica, CA: RAND.

Chapter 13: Measuring the Impact of Training

Strecher, B.M., & Hasner, L.M. (1993). *Beyond Vocational Education Standards and Measures: Strengthening Local Accountability Systems for Program Improvement*. Santa Monica, CA: RAND.

Strecher, B.M., & Hasner, L.M. (1995). *Accountability in Workforce Training*. Santa Monica, CA: RAND.

Thermer, C.E. (1997). Authentic assessment for performance-based police training. *Police Forum*, 7(3),.1-5.

CHAPTER 14

DATA SYSTEM DEVELOPMENT

by Susan Keilitz and Neal B. Kauder⁸

This chapter provides guidance on how to evaluate data systems developed or revised to improve system responses to domestic violence and sexual assault. The information presented in this chapter is drawn primarily from the experience of designing, implementing, and evaluating criminal justice system databases. Lessons learned in this context are closely related and readily applicable to the evaluation of data systems developed under VAWA STOP grants. However, these lessons are somewhat generic and therefore will not address every issue that is likely to arise in the evaluation of data systems funded by STOP grants. VAWA STOP grantees therefore should keep in mind the following considerations as they adapt the information presented here for use in their evaluations:

- ❑ STOP grants are supporting the development of a number of different types of data systems, e.g., protection order registries, court case management systems, law enforcement incident reporting systems, criminal history records, victim notification systems, victim hotline and service delivery databases, and locally integrated data systems.
- ❑ These various data systems have different purposes and uses, e.g., protection order enforcement, law enforcement tracking of past incidents where no arrest was made, bail and charging decisions, crafting court orders, sentencing, enhancing victim services, research and planning, policy development, report production, and needs analysis.
- ❑ The data systems have different sets of users, e.g., law enforcement, pretrial services, prosecutors, courts, probation, victim services, social services, and researchers.
- ❑ Many systems are seeking to link or integrate data from different systems, both cross-agency and cross-jurisdiction, and may encounter technical, financial, philosophical, legal, and political obstacles.

All VAWA STOP evaluations of data systems, regardless of the system type and purpose, should document information on five factors that might affect the development or revision of a data system. These factors follow:

- ❑ Whether the system was developed with the full support of high level policy makers and decision making authorities.

⁸ Susan Keilitz of the National Center for State Courts in Williamsburg, Virginia and Neal B. Kauder, a Principal at VisualResearch in Richmond, Virginia are evaluating STOP projects in Purpose Area 4, Data and Communications Systems, under a National Institute of Justice grant.

- ❑ Whether the intended users and beneficiaries of the system participated in system development.
- ❑ The existence or degree of communication between the administrators of various data system components.
- ❑ Whether the system is new or an enhancement to an existing system.
- ❑ Whether the data system project is funded exclusively by STOP grants or is cofunded by other sources.

EVALUATION METHODOLOGY

Chapter 6 of this *Guidebook* addresses methodological issues related to conducting evaluations of STOP grant projects. One of these issues is the value of having information from a time period before the project began against which to compare information gathered in the evaluation. Evaluations of data systems may be more likely than others to suffer from the absence of pre-project information, because in many instances no data system existed to which the new system can be compared. New automated systems can be compared to the manual systems they replaced, however. Likewise, the effects of a new system can be assessed by comparing whether and how particular functions were performed before and after the system was implemented.

One way to gather information to use in an evaluation of a data system without incurring the expense of collecting pre-project data is to conduct a **users' needs assessment** before a system is designed or restructured. The results of a needs assessment can serve as an objective baseline for measuring how a system is meeting the needs of its users. Among other things, needs assessments typically identify who intends to use the data system, what pieces of information are required, and how the information will be used. Whether or not specific groups of people have received the information they need is a critical measure for assessing the utility of the system.

The lack of pre-project comparative information should not discourage evaluation of data system projects. This chapter suggests ways that VAWA STOP grantees can conduct an evaluation that provides valid and useful information about how well a data system is functioning and how it might be improved. To accomplish this goal, evaluations of data systems should involve at least three basic components, discussed below:

- ❑ **Documenting and describing what the data system is intended to do and how it operates.** This information should include the purpose of the system, what data it contains, who operates or manages the system, who uses the system, and what they use its information for.

- ❑ **Determining if the information in the data system is accurate, reliable, accessible, and useful.** Document the extent to which the information is inaccurate or missing, whether intended users have ready access to the system, whether the number and types of people who can access the information increased, and whether the information has the specificity needed to allow line staff and researchers to do their jobs.

- ❑ **Assessing whether the system makes a difference.** Determine whether and how the existence or improvement of the system has changed the behaviors or performance of system users and to what extent research has been undertaken or new policies developed using data from the system.

Potential sources of evaluation information and methods for gathering the information are discussed in Chapter 6. Many of the sources described in that chapter should be tapped in an evaluation of a data system. These include reviews of existing system documentation and reports generated about or by the system; interviews with and surveys of system developers, participants, and users; audits of the system to compare data from its original source with the form and content of the data in the system for accuracy and completeness.

LOGIC MODEL FOR A DATA SYSTEM PROJECT

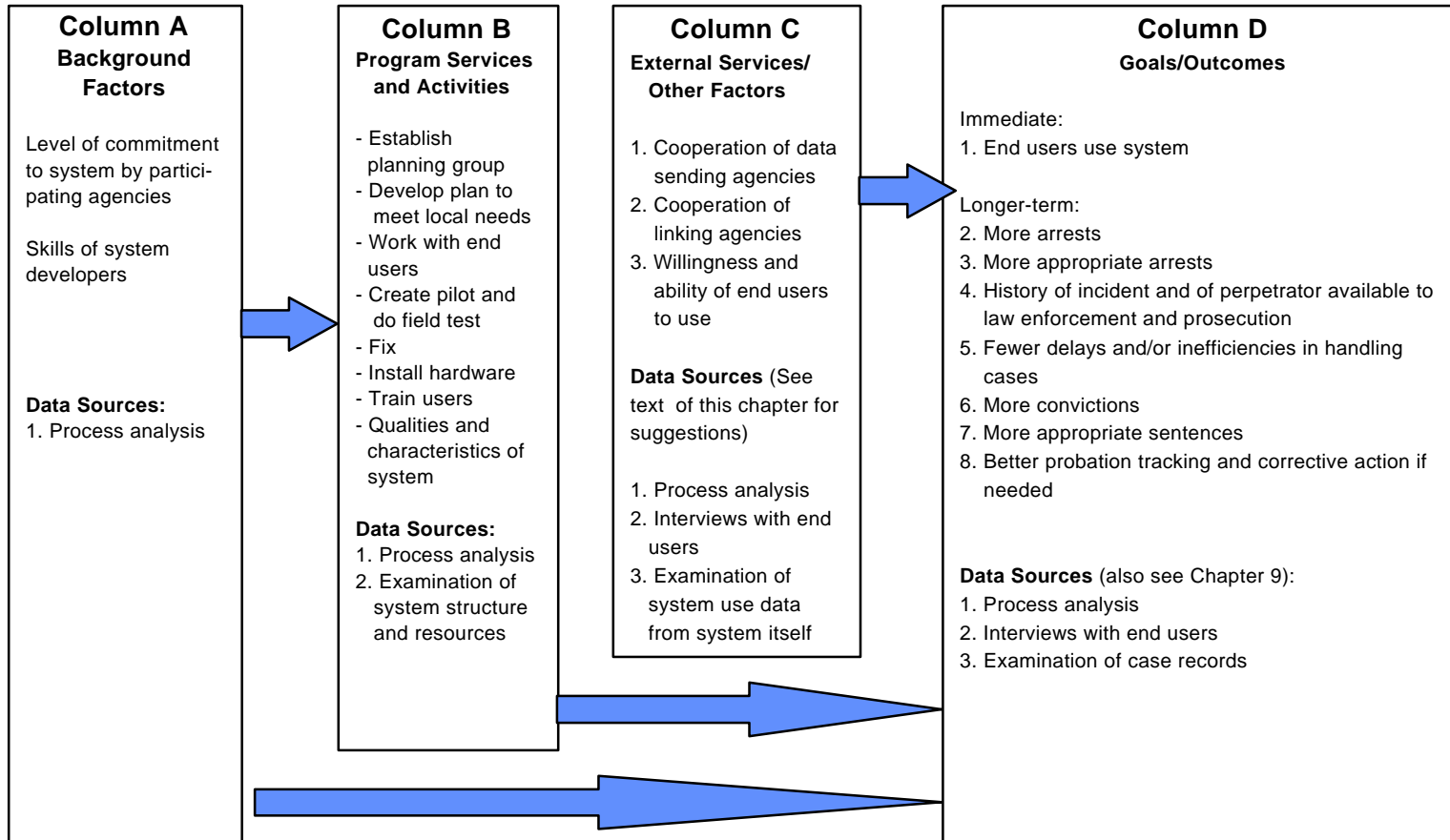
As with Chapter 13 (training), this chapter addresses issues in evaluating a particular type of project, and also lends itself to developing a logic model. Exhibit 14.1 shows a schematic logic model for a data system project, without a lot of detail. Column A contains background factors which, in this case, are mostly characteristics of systems, although some individual characteristics may also be relevant. Column B is also a bit different from other logic models in this *Guidebook*, as it shows the *steps in developing* the data system, as well as the operating characteristics of the system itself. As with the training evaluation, the steps and operating characteristics should be documented with the techniques of process evaluation.

Column C identifies some common external factors that can enhance or undermine the development of a data system and its ultimate usefulness. Column D identifies the obvious primary goal of having the intended users actually use the system. It also identifies a variety of system-level impacts that might, in the long run, be influenced by the availability and use of a good data system.

The remainder of this chapter discusses the wide variety of variables that might be used for different parts of this logic model, and gives examples of the topics one might cover and the questions one might ask.

Exhibit 14.1

DATA SYSTEM DEVELOPMENT PROJECT



DOCUMENTING THE PURPOSES AND OPERATIONS OF THE DATA SYSTEM

This component of an evaluation lays the foundation for the other components of the evaluation. It should provide you with a thorough understanding of the technical and operational features of the data system. This understanding should guide the evaluator in identifying issues to explore further and what information sources will be needed. For example, if a purpose of the system is to produce a monthly report for the prosecutor and the court, you will know to inquire of prosecutors and court staff whether these reports are useful and produced in a timely manner. If a users' needs assessment was completed before the system was implemented, you should be able to conduct this component of the evaluation in less time and with fewer resources.

The first task you need to do for this evaluation component is to gain a basic understanding of the system by reviewing documentation of the system's technical requirements, data dictionaries, and any background reports. You should be able to get this information from the individuals responsible for the design of the data system and from those responsible for managing it. The ease or difficulty of obtaining system documentation is itself a measure of the quality and utility of the data system.

The next step is to gather information from the people who participated in developing and operating the system and those who use it. You will need to interview the system designers and developers, staff who enter and delete information from the system, system managers, and system users. If several individuals or groups use information in the system, mail or telephone surveys may be an efficient way to gather information from the universe of users. It is important that you include all users of the system in this part of the evaluation.

Finally, functional diagrams of data system should be constructed. These system or functional diagrams serve as a visual guide to the system and to the evaluation. They can help you understand the flow of information through the system as well as focus on and record how the system functions at each juncture. These diagrams also will facilitate the documentation of system changes in the second and third components of the evaluation (see below).

The system review and interviews should address the following issues: What are the intended purposes of system? Who was involved in the design and implementation of the system? What resources supported system development and implementation? What are the technical requirements of the system (e.g., hardware and software used and compatibility with other systems, field lengths, data transmission capabilities)? What information is contained in system? Can the information be accessed as individual case records, or can only aggregated summaries be produced? Can individual variables within be accessed and analyzed? Who enters information into the system? When and how is information entered? How is information updated? How and when is information audited? Who uses the information in the system? How is the information used (e.g., to produce periodic or occasional reports; to track activity daily, monthly, yearly; to exchange with others; for evaluation)? What individuals or groups can access the information and

how is this accomplished? What measures are in place to ensure confidentiality of the information in the system? Which, if any, systems are linked and how?

MEASURING THE ACCURACY, RELIABILITY, AND UTILITY OF THE DATA SYSTEM

The second component of an evaluation of a data system builds on the information gathered while documenting the purposes and operation of the data system. As in the first component, this component entails interviews or surveys of system users. It also involves reviews of original sources of information contained in the system to determine the degree to which this information matches the information in the system. You may be able to obtain the information on user perceptions of the system while gathering information in the first evaluation component.

Assessing User Perceptions of Accuracy, Reliability, and Utility

User perceptions of data system quality and utility typically are measured through surveys and interviews to ensure broad representation of users at various agency or occupational levels. Comparing the results of an initial needs assessment with the results of follow-up satisfaction interviews allows you to gauge the degree of change related to the actual data system implementation. Regardless of whether an initial needs assessment was completed, questions can be posed about former and current use of the data system to identify the benefits and problems associated with system changes. The following items typically are addressed in an assessment of user satisfaction with the data system:

- The actions or functions that are the responsibility of the user.
- How the user obtains information from/accesses the data system and how this compares with the user's experience before the system was created/revised.
- How easy or difficult it is to gain access to the system and how this access has been improved or hampered.
- What pieces of information are used in the data system, how the information has been improved, and what needs further improvement.
- What pieces of information have been added (or should be added) to the data system and how that information is now used.
- Whether or not the information obtained from the system appears to be accurate.
- How often information is missing.
- The noticeable impact concerning the user's job performance or duties that can be directly or indirectly attributed to the data system changes (see below).

Assessing Data Quality

Developing a process for tracking the accuracy and reliability of a data system will vary according to the level of specificity and detail needed or desired from the outcome measures. For example, it may be useful to know how often agencies fail to record dispositions for domestic assault arrests into the data system, but it may be even more helpful to determine which types of domestic assault cases are most likely to have missing data of this type. A jurisdiction may find that information fields are filled out more often in felony cases that are handled in the court of general jurisdiction than they are when the cases are misdemeanors heard in limited jurisdiction courts. It may be as important to describe the nature or accuracy of an arrest or disposition entry as it is to determine whether it has been entered.

The need to evaluate the quality of the data, as opposed to its mere existence, is becoming more important as domestic violence protocols and risk assessment instruments are relying more heavily on data to guide their development. When evaluating a specific data system it is also important to consider data quality across other systems, including law enforcement (arrest), court (disposition), and corrections (treatment/incarceration) data systems. The data system under evaluation may be functioning well but other data systems to which it relates or upon which people also rely may be flawed. This consideration becomes more critical as government officials and communities attempt to efficiently and reliably share or merge data across various databases and organizations (see below).

Considerations for Assessing Data System Quality. Many jurisdictions track measures of data quality as part of an audit or data quality assessment program. For example, managers of criminal history repositories report rates of missing or inaccurate data as part of their ongoing effort to improve criminal history records. Lessons from these experiences provide a few considerations for establishing measures for assessing the quality of other data systems:

- ❑ **Determining the accuracy of information in the data system.** Any data system evaluation should assess the degree to which information stored in the system is correct. Accuracy in data system reporting helps ensure that entries are valid—that they are measuring what they intend to measure. After the degree of accuracy is established, you can then attempt to measure data system reliability—whether or not the data system is consistently measuring information in a stable or uniform fashion. It is important to note that before being concerned with data system consistency (reliability), you should first establish levels of accuracy (validity). A consistent, inaccurate measure of offender behavior, or for that matter, any other justice system phenomenon, is one that is predictably wrong and therefore not very useful.

- ❑ **Defining the term “complete entry” when describing events within a data system.** With input from all relevant stakeholders, you (and system designers) should determine a realistic standard for how an entry should appear to users,

managers, or researchers when viewing the information. Once the standard is developed, you can determine the extent to which entries meet the “complete entry” standard. Arriving at a consensus definition for a “complete entry” may be one of the most difficult tasks in the evaluation process, however. You will need to consider the definitions applied in the past, as well as any new definitions established as part of the intended system changes or revisions. In this sense, the degree to which data quality is improved or that the data system becomes more useful can be measured over time by the extent to which the standard for completeness can be raised. Monitoring outcome measures over the long term is important, because data system improvements may evolve slowly, depending on the extent or pace of procedural and form updates, software and hardware upgrades, user training and experience, and staff turnover.

- ❑ **Tracking/matching statute offenses.** Improvements in data systems often involve using new offense codes or descriptors. These schemes are designed as an alternative to “free format” or text fields that are often too subjectively or inconsistently applied. The power of coding offenses or factors (such as weapon use) rests in the ability to describe the elements of a crime more precisely. For instance, as a result of modifying offense reporting, prosecutors may be able to “flag” offenders based on their recidivism status or potential for future violence. Many jurisdictions also have specific statutes or policies that increase penalties or call for mandatory arrests for specified subsequent offenses. Objectively defined offense codes can be used to immediately alert police and court officials that a specific case falls within these criteria. This concept can be applied to other data system contexts, such as standardizing codes for protection order terms entered in protection order registries.

- ❑ **Conducting a data system audit.** Comparing original source records or documents (e.g., police arrest reports, civil protection orders) with information found in automated records is a straightforward and easily interpretable method for determining data system quality. An audit designed and executed correctly will yield more direct and objective data quality measures than you will get from interviews or surveys. You can expect these measures to include the timeliness of data availability in the system and rates of complete records, accurate records, and successful linkages among multiple systems.

Audits are time and labor intensive, however, because they require a manual examination of the source documents. This process can be facilitated by accessing transaction logs (maintained by some jurisdictions) to provide an audit trail of all inquiries, responses, and record updates or modifications. It may be necessary to perform random sample audits, depending upon a number of factors, including the level of resources available to conduct an audit and the extent to which system access crosses agency or jurisdictional boundaries. Pulling a sample properly is important to ensure that the entire universe of cases (all records within a data system) have an equal chance

of being selected for audit. In some cases, you may want to select more cases (oversample) from a subgroup of cases or cases with particular characteristics that are more important or that occur with less frequency, such as violent domestic assaults and homicides. (See Chapter 6 for discussion of a number of other considerations concerning sample selection.)

MEASURING EFFECTS OF THE DATA SYSTEM ON BEHAVIOR, PERFORMANCE, AND POLICY CHANGE

In most scenarios, a new or redesigned data system can be considered effective if there are positive changes in the behavior or performance improvements of those who directly or indirectly access the system. Benefits may be quite obvious, as is the case with quicker responses for information requests or queries, or more subtle, as in the way information is formatted or presented to users.

One method for determining whether there have been changes or improvements in the behavior or performance of users involves constructing detailed diagrams of the various access points of a data system. These diagrams can depict how improvements at one stage of a data system can aid others at lateral or later points in both the data system and in the responses of the justice system and service providers. Identifying whether information gathered or improved at one segment of the system has benefited those involved in other segments is critical. The degree to which information reporting at the front end of a data system affects those at the back end should be communicated clearly to all system users, because those at earlier stages can rarely view the connection between their job performance and overall justice system improvements.

The diagrams can guide the selection of individuals across the system(s) to interview or survey. These system users may include data entry personnel, law enforcement officers, prosecutors, judges, court staff, probation officers, treatment providers, shelter staff, legal services providers, and other victim services providers. Information from these interviews or surveys can be combined with official justice system statistics (e.g. arrests, charges, dispositions, protection orders issued, enforcement actions, sentences) to more accurately describe the impact of the data system on behavior and policy change.

There are a number of indicators that data system changes have affected user behavior, system responses, and policy improvements. A few examples are described here

Changes in User Behavior, Job Performance, and System Responses

Those who access an information system for purposes of doing their job are acutely aware of changes in data accessibility, interface formats, and the quality of information stored. For example, as a result of improvements in a criminal history repository, law enforcement may conduct background checks on suspects at the earliest stages of the arrest or investigation

process. They may also indicate a greater degree of confidence in the information they obtain from these data requests. As a result, officers may be more aware of outstanding warrants or civil protection orders, which in turn leads to a more accurate and complete charges. Improvements in protection order registries may result in more arrests for protection order violations, because the specific terms of orders are identifiable in the system. Police may report to you that their job performance has improved, while data obtained from police statistics may show a concomitant increase in arrests or an increase in the frequency of multiple charges.

Likewise, prosecutors may find that improved access and data specificity reduces the time they have to spend calling or waiting for authorities from other agencies to respond to their inquiries. These prosecutors may find that their behavior on the job has changed, because they have more time to investigate the facts of a case or they may be able to spend more time working with the victims. An examination of conviction rates before and after implementation of the system may indicate that more cases are successfully prosecuted.

Services Affecting Victims, Offenders, and Other Clients

The range of benefits derived from improving a data system may not be noticed entirely for several years. For example, using STOP grant funds to integrate data across agencies or organizations may have the most impact after a database on offenders (or victims) has been solidly established. In this instance managers and researchers can begin to more accurately profile those who come into contact with the justice system. Information obtained from this type of analysis can be used for a number of purposes, including enhancing victim services, determining long-term treatment effectiveness, describing the related factors and tracking changing patterns in domestic violence and sexual assault incidents, objectively assessing offender risks to public safety at the various criminal justice stages, and assessing the effectiveness of new laws or policies aimed at deterring or preventing domestic violence

Use of System Data by Government Officials and Others in Policy-Making Positions

Government officials and policy making bodies are well-established consumers of data analysis and research. The extent to which they access and utilize data from the new or revised system can be assessed in a number of ways, including the following:

- ❑ **Examining management reports that are generated.** One clear sign that data are being used more frequently is an increase in the frequency and types of management reports or statistical summaries as compared to those generated from past data systems. For example, specific definitions of victim-offender relationship may be reported for the first time as a direct result of a data coding change. Similarly, the ability to report rates of recidivism or how often offenders

successfully complete treatment programs may indicate the successful integration of a law enforcement, court, and correctional data system.

- ❑ **Conducting individual level (incident-based) case studies.** Research and policy analysis is most effective for informing long-term decision making when data are available in a raw, individual record format, rather than in a format where individual records (e.g., cases, events, or persons) have been aggregated to produce a summary report. The summary format makes it impossible to analyze the circumstances and effects of justice or service system actions on an individual case level. From an evaluation perspective, the emergence of individual or incident-based studies indicates that a data system is sufficiently flexible to switch from an aggregate or operational case tracking system, to an incident-based, management and research database.

Expanding the Scope or Policy Application of Studies

Another indication that a data system is becoming more effective in aiding program development relates to the scope of the studies that are being conducted. For example, the ability to develop risk assessment instruments or interview protocols based on empirical information can only be accomplished with high-quality and detailed data that describes all aspects of a domestic violence event.

Increased Variety of Organizations Requesting Information and Data

Knowledge of the availability of high-quality data from a justice system often spreads across a variety of organizations. This will be more likely to occur if system data are not only well maintained and analyzed, but also are presented in formats that are readily understood and policy relevant. Legislative staff, judges, and executive branch advisors are the types of officials most likely to be aware of the benefits of an accessible and reliable data system. You may find it useful to survey these groups to test their knowledge of a system and to determine in what ways they use the data. Representatives of these groups can also recommend ways for improving the types of analyses that are conducted and the formats used for presenting results.

Evaluating the relationship between data system modifications and changes in justice and service system responses can be a complex undertaking for a number of reasons. Many data system improvements come at a time when organizations are redefining their roles or changing their structures completely. For this reason, isolating the effects of a data system on behavior and policy change may not be easily achieved. Although statistics and survey information gathered before and after the implementation of a data system may indicate a positive change, the data system alone may not be the cause of the change. For this reason, it is important to control for other intervening factors, and at a minimum, to search for the existence of alternative

explanations. Policy makers will benefit most from evaluations that explain both how a data system improved job performance and system responses and how observed successes and failures occurred within the context of broader system changes.

NEW APPROACHES TO DATA ACCESS AND INTEGRATION

As states and localities move toward improving the quality and utility of their data systems, the need is arising to break down walls among the various departments and agencies that store common pieces of information. To this end, system designers are beginning to approach the problem of data access not from a data “sharing” viewpoint, but from a more seamless data “integration” perspective. In this sense, data from different systems are merged continuously through a network or Intranet environment.

The data integration approach allows individual agencies to maintain basic data structures and formats that have evolved for decades with significant commitments of resources. The recognition is growing that creating a single unified database, centrally housed and requiring each agency to report data in exactly the same format, is an insurmountable task. Instead, the technology exists (through “middleware” software products) to take data from a variety of data systems and merge or store the information in a common format without having to rebuild entire information systems. This new environment still requires data integrity and completeness at the various junctures of the system but offers a technological solution to the existing problems in the transfer of data across systems or agencies.

Because the data integration approach appears to be the direction in which data systems are heading, current STOP grant project evaluations may benefit from an assessment of the possibilities of becoming part of an integrated data system environment. One of the first tasks will be to ensure that those involved in violence reduction initiatives have a voice during deliberations relating to the development of comprehensive integrated information systems. These individuals and the groups they represent are not only in the best position to assess information needs related to victims and offenders, but they are also well situated to know how linking to alternative data sources can help them to perform more effectively. They may also recognize how the information they hold can in turn benefit others.

A thorough discussion of the issues related to integrating data systems is beyond the scope of this guidebook, but VAWA STOP grantees should be aware that the environment is changing. A few questions that might be considered initially are presented here: Should access to our data be granted to other officials in the judicial process, and will this in turn facilitate access to data held by those other entities? Are there ways we can design or structure our system to better prepare for future data integration across systems? Has our system attempted to integrate information across multiple law enforcement, judicial, and correctional platforms? Do any local or wide area networks or Intranets exist to allow seamless access and sharing of information? Can we download information for further research and policy analysis?

CHAPTER 15

SPECIAL ISSUES FOR EVALUATING PROJECTS ON INDIAN TRIBAL LANDS

by Eileen Luna, J.D., M.P.A.⁸

The VAWA mandates a set-aside of 4 percent of annual funding to support STOP projects on Indian tribal lands. In addition, services to women in Indian tribes is one of the seven purpose areas for which states may use their STOP allocation. Projects run specifically by and/or for women on tribal lands are relatively new activities in the area of violence against women. Therefore it is important that we use evaluation to learn as much as possible about the success of these projects. All of the specific suggestions for conducting evaluations that we have made earlier in this *Guidebook* are relevant to evaluations of projects on Indian tribal lands. In addition, Indian tribal lands may be subject to some legal and jurisdictional complications that do not apply elsewhere. That, combined with the fact that tribes are sovereign nations with a wide variety of traditions and experiences, means that evaluators would profit from some background before jumping into an evaluation.

This chapter was written to provide some relevant background. Evaluators who wish to work with Tribal governments need to understand something of the history, culture, traditions, and protocol surrounding Indian tribes. They need to know that jurisdictions can be uncertain, varying from place to place, tribe to tribe, and over time. And they need to be prepared to work with Tribal governments who have little or no experience collecting systematic data on crimes, services, or outcomes, or complying with grant conditions that are routine for state and local governments in the United States.

THE INDIAN TRIBAL GRANTS PROGRAM

Grants to Indian tribes are subject to the same requirements as other STOP grants. The VAWA mandates that the STOP grantees focus 25 percent of their grant monies in each of three priority areas: law enforcement, prosecution, and victim services. The remaining 25 percent can be spent at the grantees' discretion. A number of Tribal STOP grantees have elected to expend this discretionary percentage on the development and implementation of elements of tribal court systems.

Within each priority area, the grantees have undertaken to develop training programs as well as appropriate policies, protocols, and procedures; to undertake the collection of data; and to facilitate communications within and among programs and governmental agencies.

⁸ Eileen Luna, of the University of Arizona American Indian Studies Program (AISP), is evaluating STOP Indian Tribal Grants under an NIJ grant. She was assisted in the development of this chapter by her colleagues in AISP, particularly Jay Stauss and Robert Hershey.

Given the recent development and implementation of most tribal police departments (amounting to a 49 percent increase in two years),⁹ the development of appropriate training programs and policies is of critical importance. However, it is unwise, particularly in Indian Country, to take rules and policies directly from other jurisdictions or to train law enforcement and other program personnel without regard to compatibility with the particular tribal culture.

The evaluation of Tribal programs should take into consideration the legal and jurisdictional issues unique to Indian Country. Evaluation should also consider the fact that many Tribal programs, particularly in the fields of administration of justice and victims services, have been created very recently. There may be little history and few models to assist with programmatic development. It is essential that the evaluation team commit to helping the Tribes develop competent and effective programs, while conducting the necessary critique of that which has been developed and implemented.

THE ADMINISTRATION OF JUSTICE IN INDIAN COUNTRY

The difficulties inherent in working with, and within, Indian Country, particularly in the field of program evaluation, are often difficult to foresee. Mistakes can be made that can hinder the development of relationships and trust—factors that are essential components of thorough evaluation. These mistakes can be avoided if one considers the legal framework of Tribal sovereignty, the jurisdictional issues that criss-cross Indian Country, and advisable protocols. The following should help to clarify the legal and jurisdictional issues faced by Tribal programs, explain the protocols that facilitate working in Indian Country, and thus make the evaluation of these programs more comprehensive.

Tribal Sovereignty

Tribes are sovereign and self-governing. Except for express limitations imposed by treaties, by statutes passed by Congress acting under its constitutionally delegated authority in Indian affairs, and by restraints implicit in the protectorate relationship itself, Indian tribes remain “independent and self-governing political communities.” A recognition and understanding of the “government-to-government” relationship, which has existed historically and has recently been emphasized by

⁹ A survey being conducted by the University of Arizona AISP and the Kennedy School of Government, funded by the Police Executive Research Forum, indicates that tribal funding of law enforcement has grown considerably in recent years. In Fall 1995, 114 tribes had tribally funded police departments, compared to 170 as of June 1997. The remaining 340 tribes have their law enforcement needs met by the BIA’s law enforcement system, or by the states through Public Law 83-280.

Chapter 15: Special Issues for Evaluating Projects on Indian Tribal Lands

President Clinton in his 1994 Executive Order¹⁰ and subsequently by Attorney General Janet Reno,¹¹ is necessary if an evaluation is to be fairly and impartially done.

Tribes have the power to pass and enforce laws to prevent violence against Indian women on their reservations. The programs should reflect the cultural circumstances of the particular tribe. The programs should be historically consistent. They must also address the needs of Tribal members, while supporting the development of Tribal governments and the concept of sovereignty.

In most situations tribes retain a broad degree of civil jurisdictional authority over members and non-members on the reservation, particularly where the conduct threatens or has some direct effect on the political integrity, economic security, or the health or welfare of the tribe. But the right of Tribes to exercise criminal jurisdiction over non-Indians is not so clear. The Supreme Court ruled in *Oliphant v. Suquamish Indian Tribe*¹² that Tribal sovereignty does not extend to the exercise of criminal jurisdiction over non-Indians. However when the Supreme Court supported this decision again in *Duro v. Reina*¹³ by expanding *Oliphant* to preclude criminal jurisdiction over non-member Indians, the case was specifically overruled by Congressional action. This action by Congress, called the “Duro-fix” gives Tribal governments rights of criminal action against non-member Indians.¹⁴

Tribal Jurisdiction

Jurisdiction is the power of a government to make and enforce its own laws. Tribal jurisdiction presents the most complex set of issues in the field of federal Indian law and policy.

Today, when Indian people and their leaders speak of tribal sovereignty, what they are usually talking about centers on questions of tribal jurisdiction—questions of whether or not a tribal government has the political power and legal authority to act, legislate, and enforce its laws with respect to certain persons and subject matter. A great deal depends on the answers to such questions, because when Congress or the courts are called on to resolve the jurisdictional disputes

¹⁰ Executive Order, 4/29/94, *Government-to-Government Relations with Native American Tribal Governments*.

¹¹ Reno, Janet. (1995). U.S. Department of Justice Commitment to American Indian Tribal Justice Systems. *Judicature*, Nov.-Dec.

¹² 435 U.S. 191 (1978).

¹³ 495 U.S. 676 (1990).

¹⁴ P. L. 102-137. This one-sentence amendment to the Indian Civil Rights Act simply maintains Tribal criminal jurisdiction over both member and non-member Indians for crimes committed in significant part on Tribal lands.

that can arise among tribes, states, and the federal government, what is ultimately being determined is who governs the land, the resources, and the people in Indian Country.

In the evaluation of tribal programs and the problem of preventing violence against women in Indian communities, crucial issues to be analyzed include:

- The source and scope of tribal jurisdictional authority.
- The geographic area encompassed by tribal jurisdiction.
- The extent to which United States law has altered the original exclusive jurisdiction of the tribes and has transferred all or part of that jurisdiction to the federal government or state governments.
- The problems created for tribes in enforcing law and order on their reservations by the limitations imposed on tribal jurisdiction under United States law.¹⁵

Source and Scope of Tribal Jurisdictional Authority

Indian tribes existed as independent, self-governing societies long before the arrival of European colonial powers on the North American continent. This status was recognized by the Supreme Court in one of its earliest and most important decisions, wherein Indian tribes were held to be “distinct, independent political communities.”¹⁶ This history and status as sovereign, self-governing societies distinguishes Indian tribes from other ethnic groups under the Constitution and laws of the United States and must be taken into consideration when evaluating tribal programs.

Within their reservations Indian tribes have inherent sovereign powers of civil and criminal jurisdiction over their own members. Tribes control their own membership, maintain their own judicial and criminal justice systems, and regulate the reservation environment and its resources. Tribes also have absolute jurisdiction over internal matters such as tribal membership and domestic relations. By virtue of their sovereign, self-governing status, tribes have the power to pass and enforce laws. Although they are not empowered to exercise criminal jurisdiction over non-Indians, they do retain a broad degree of civil jurisdictional authority over non-members on the reservation.

¹⁵ See David Getches, Charles F. Wilkinson, Robert A. Williams, Jr. (1993). *Federal Indian Law: Cases and Materials*, 3rd edition.

¹⁶ *Worcester v. Georgia*, 31 U.S. (6 Pet.) 515, 559 (1832).

Indian Country

The term “Indian Country” is the starting point for analysis of jurisdictional questions involving Indian tribal governments and whether or not P.L. 280 applies. Federal law defines the geographic area in which tribal laws (and applicable federal laws) normally apply and state laws do not as follows:

...[T]he term “Indian country,” as used in this chapter, means (a) all land within the limits of any Indian reservation under the jurisdiction of the United States government, notwithstanding the issuance of any patent, and including rights-of-way running through the reservation, (b) all dependent Indian communities within the borders of the United States whether within the original or subsequently acquired territory thereof, and whether within or without the limits of a state, and (c) all Indian allotments, the Indian titles to which have not been extinguished, including rights-of-way running through the same.

This definition, which appears in 18 U.S.C., Sec. 1151, a criminal statute, is also used for purposes of determining the geographic scope of tribal civil jurisdiction. Individual allotted parcels of land not located within the reservation may still constitute Indian Country if the Indian title to the land has not been extinguished.¹⁷

Criminal Law in Indian Country

The imposition of federal and state criminal laws in Indian country has caused great confusion. The Major Crimes Act¹⁸ gave federal courts jurisdiction over thirteen violent felonies.¹⁹ The Assimilative Crimes Act²⁰ and the Organized Crime Control Act have specifically been held to apply to Indian Country.

In states other than those where P. L. 280 applies (discussed below), subject matter jurisdiction of federal, tribal, or state courts is usually determined on the basis of two issues (1) whether the parties involved in the incident are Indians and (2) whether the incidents giving rise to the

¹⁷ See Robert N. Clinton. (1976). *Criminal Jurisdiction Over Indian Lands: A Journey Through a Jurisdictional Maze*, 18 Ariz. L. Rev. 503, 515.

¹⁸ 18 U.S.C. Sec. 1153 (1988).

¹⁹ The federal courts have jurisdiction over murder, manslaughter, kidnapping, maiming, a felony under chapter 109 A, incest, assault with intent to commit murder, assault with a dangerous weapon, assault resulting in serious bodily injury, arson, burglary, robbery, and a felony under section 661 of this title within Indian Country.

²⁰ 18 U.S.C. Sec. 13 (1988).

complaint took place in Indian Country. For the purpose of this particular analysis, an Indian is defined as a person of Indian blood who is recognized as a member of a federally recognized or terminated tribe. Indian Country includes (1) all land within the limits of any Federal Indian reservation, (2) all dependent Indian communities, and (3) all Indian allotments.

This erosion of tribal authority has caused tribes to become uncertain about the extent and scope of their civil and criminal jurisdiction. Many crimes go unprosecuted for two reasons: jurisdictional vacuums are created by the uncertainties of tribal jurisdiction, and the realities of federal and state prosecutorial priorities normally do not give a high priority to prosecuting reservation crimes. The sense of helplessness and frustration engendered by the jurisdictional confusion that exists in present-day Indian country results in many crimes going unreported. Tribal police and courts are underfunded and training is inadequate to deal with the complexities created by limitations on tribal jurisdiction. Tribes often find it difficult to secure the cooperation of neighboring state and local law enforcement authorities. The same situation often also applies to federal law enforcement agencies in addressing reservation crime.

State Versus Tribal Jurisdiction

Implications of P. L. 280. Under general principles of federal Indian law, states do not have direct jurisdiction over reservation Indians. However, Congress has the power to vest federal authority with the states, which it did with the 1953 passage of P.L. 83-280. Further, since 1968 when Congress amended the act, states with P.L. 280 jurisdiction are allowed to retrocede jurisdiction to individual tribes within their state, upon petition by the tribe and pursuant to the approval of the federal government. Should retrocession be granted, tribal jurisdiction over reservation Indians is determined according to general principles of federal Indian law.

Six states were delegated criminal jurisdiction over reservation Indians and civil jurisdiction over cases arising against Indians in Indian Country under P.L. 280. Other states were permitted to assume such jurisdiction, pursuant to the passage of appropriate state jurisdiction and/or state constitutional amendments. Tribal consent was not a requirement for the assertion of state authority for either the mandatory or option states under the original legislation.

After this legislation passed, ten states accepted such jurisdiction. Then, in 1968, Congress amended P.L. 280 to include a tribal consent requirement, which required a tribal referendum before states could assume jurisdiction. Since that date, no tribe has so consented.

In those states that have assumed jurisdiction (see Table 15.1), P.L. 280 established state jurisdiction without abolishing tribal jurisdiction. Thus the powers are concurrent. The problem, however, is that after passage of the act in 1953, many tribes in P.L. 280 states did not develop their own court systems, relying instead on the misunderstanding that the legislation had deprived tribes of adjudicatory powers. However, the existence of a tribal court system is necessary before a tribe can assert concurrent jurisdiction.

**Table 15.1
State by State Overview of Public Law 280**

Mandatory States	Indian Country Affected
Alaska	All Indian Country in the state except the Annette Islands with regard to the Metlakatla Indians
California	All Indian Country within the state
Minnesota	All Indian Country within the state except Red Lake Reservation (Retrocession accepted for Nett Lake Reservation)
Nebraska	All Indian Country within the state (Retrocession accepted for Omaha Reservation)
Oregon	All Indian country within the state, except the Warm Springs Reservation (Retrocession accepted for Umatilla Reservation)
Wisconsin	All Indian Country within the state (Retrocession accepted for Menominee Reservation and Winnebago Indian Reservation)
Option States	Indian Country Affected
Arizona	Air and water pollution
Florida	All Indian Country within the state
Idaho	Seven areas of subject matter jurisdiction; full state jurisdiction if tribes consent: compulsory school attendance, juvenile delinquency and youth rehabilitation, dependent, neglected and abused children, insanities and mental illness, public assistance, domestic relations, motor vehicle operation.
Iowa	Civil jurisdiction over Sac and Fox Reservation
Montana	Criminal jurisdiction over Flathead Reservation: full state jurisdiction where tribes request, counties consent, and governor proclaims (Retrocession accepted for Salish and Kootenai Tribes)
Nevada	Full state jurisdiction, but counties may opt out; later amendment required tribal consent (Retrocession accepted for all covered reservations)
North Dakota	Civil state jurisdiction only, subject to tribal consent
South Dakota	Criminal and civil matters arising on highways: full state jurisdiction if United States reimburses costs of enforcement
Utah	Full state jurisdiction if tribes consent
Washington	Eight subject areas of jurisdiction on Indian trust land; full state jurisdiction as to non-Indians and Indians on non-trust land, although the state has allowed full retrocession fairly liberally (Retrocession accepted for Confederated Tribes of the Chehalis Reservation, Quileute Reservation, Swinomish Tribal Community, Colville Indian Reservation, Port Madison Reservation, and Quinault Reservation)

With regard to criminal jurisdiction, the legal situation in P.L. 280 states is as follows:

- Tribal courts (where they exist) have criminal jurisdiction on the reservation over Indians, subject to the provisions of the Indian Civil Rights Act.

- ❑ Federal criminal jurisdiction on the reservation extends to offenses committed by Indians or against Indians if a non-Indian is either the perpetrator or the victim, and over major crimes committed by Indians.
- ❑ State criminal jurisdiction on the Reservation is exercised, with exceptions, over both Indians and non-Indians.

With regard to civil jurisdiction, the legal situation is as follows:

- ❑ Tribal courts have civil jurisdiction on the reservation over Indians and non-Indians, which is limited by principles of federal Indian law.
- ❑ Federal civil jurisdiction on the Reservation is determined using the same standards that are used to determine federal civil jurisdiction off the Reservation.
- ❑ State courts have civil jurisdiction, with some exceptions, over suits involving Indians or non-Indians (e.g., custody proceedings involving the Indian Civil Rights Act and the Indian Gaming Regulatory Act of 1988).

State regulatory jurisdiction. State regulatory jurisdiction was not granted under P.L. 280. The states were specifically not allowed to tax the reservations for services, such as law enforcement and access to state courts, rendered pursuant to such jurisdiction, nor were they allowed to infringe on water rights nor interfere with, control, or regulate any rights or privileges related to hunting, fishing, or trapping afforded under federal treaty, agreement, or statute.

In all states tribes may regulate, through taxation, licensing, or other means the activities of both tribal members and non-members who enter consensual relationships with the tribe or its members, through commercial dealing, contracts, leases, or other arrangements. A tribe may also retain inherent power to exercise its civil authority over the conduct of non-members within its reservation when that conduct threatens or has some direct effect on the political integrity, the economic security, or the health or welfare of the tribe.²¹ Further, a tribe may banish from tribal land any non-member against whom charges have been brought.

Law Enforcement Issues

Approximately 170 of the 230 reservations that are federally recognized at present have law enforcement departments. These departments are of five types. The types are not mutually exclusive, so more than one type may operate simultaneously within the boundaries of a given reservation.

²¹ *Montana v. United States*, 450 U.S. 544 (1981).

Chapter 15: Special Issues for Evaluating Projects on Indian Tribal Lands

The Bureau of Indian Affairs is involved with two of the types of law enforcement: BIA-LES and 638. Where P.L. 280 operates, the states are responsible for law enforcement on the reservation. Throughout the United States, even where P.L. 280 exists, many Indian Nations have their own tribal police that they fund and control. These tribal police departments often operate on reservations covered by other forms of law enforcement, including BIA, 638, and/or self-governance-funded law enforcement programs. All this, of course, results in problems of overlapping jurisdictions and conflicts of law.

Given the attention paid in recent years to the expansion of law enforcement services, this area is one where confusion regarding jurisdiction and conflict of laws may easily arise. For example, a problem arose recently regarding criminal statistics obtained from the Tribal governments that operate 638 and self-governance police departments. Most Tribal governments have not traditionally provided criminal incident statistics to the U.S. Department of Justice as is required of all other law enforcement departments in the United States. Now, however, Tribal governments have begun to receive funding from the Omnibus Crime Bill and through the Community Oriented Policing Services (COPS) program for the expansion of tribal law enforcement programs. This funding since 1995 has included over \$22.4 million for new police services to 128 federally recognized Indian nations, over \$6 million for programs aimed at reducing violence against women and almost another \$5 million to fund the development of juvenile justice and other community-based programs that emphasize crime reduction.²² When Tribal governments accept funding and contract with the U.S. Department of Justice to provide specific law enforcement services, there are requirements attached which include statistical reporting, something which may never have been done before by the Tribe. Also, the likelihood increases that the Tribe will be impacted by federal legislation aimed at changing or increasing the regulations under which law enforcement operates.

The burgeoning growth of Tribal police departments has increased the potential for conflict with federal, state and local law enforcement agencies. The issues of who has jurisdiction, whether that jurisdiction is exclusive or mutual, and which is the lead agency during a given incident can have sweeping repercussions. The failure of an agency to recognize or credit the personnel of another can seriously jeopardize the success of a law enforcement action. Further, the failure of one agency or institution to recognize an order by another can considerably hamper the administration of justice and impair the protection of a victim or witness.

Tribal Court Systems

More than 140 tribes now have their own court systems. Of these, approximately twenty-five have BIA-appointed judges and have retained the Code of Federal Regulations. The rest have

²² *Nation to Nation*, newsletter of the U.S. Department of Justice, Office of Tribal Justice, August, 1996.

developed their own tribal courts and have their own laws and codes, many of which may not be in writing.

Some tribes give members a choice of appearing before a more Euro-American-oriented tribal court or the more traditional tribal court. Others have established intertribal courts that are shared among tribes and are also often used for appellate hearings.

Many tribal judges are not attorneys. They may have received training from the National American Indian Court Judges Association or the National Indian Justice Center. Some have become active in state judges' associations.

The role and status of lawyers also varies by tribe. Often tribes do not require a lawyer to be a member of the state bar association. A few tribes have their own bar examinations. Many tribal courts do not provide legal representation for a defendant, but all allow defendants to have representation at their own expense, pursuant to the Indian Civil Rights Act.

The disparity between tribal court systems, and between tribal court systems and the Euro-American system, can engender distrust and even fear among non-Indians. State and local officials can resist granting recognition to tribal court orders, and issues of Full Faith and Credit abound.

Tribal Codes

Prior to colonization, the governmental structures and legal systems of tribes were largely unwritten and informal. However, they were no less effective than other systems. To a great extent they vested authority in Tribal councils, granting them the power to create laws, to negotiate with other tribes and governments, and to deal with issues of tribal land and resources.

Often these laws and powers are not written down, or if they are, they are not explicit. Much is left to the interpretation and discretion of those in authority. Particularly in the realm of domestic abuse, discretion is often left to clans, families, and others with traditional responsibilities.

The definition of violence against women may vary from tribe to tribe, with some codes having specific lists enumerating criminal offenses and others including acts that would be of a more civil nature in a Euro-American code. Other codes include as criminal and/or civil wrongs those acts that have a damaging effect on the spirit of a person²³ or which threaten the health or welfare of their children.

²³ See identification of instances of *disharmony* in Rules for Domestic Violence Proceedings, Courts of the Navajo Nation Rule 1.5.

Chapter 15: Special Issues for Evaluating Projects on Indian Tribal Lands

The remedies for incidents of violence against women may also vary from tribe to tribe. Generally, however, victims of domestic violence are accorded greater or equal protection under tribal codes than they would have under state laws. Some tribes handle all charges of violence against women as criminal, while others use civil statutes. Still others combine criminal and civil sanctions.

Unfortunately, tribal governments are generally at an early stage in their code development. Often tribal leaders do not fully comprehend the significance of this task and may be working in the dark, without the assistance of professionals who have written these types of laws before. Others may not wish to tie themselves down to specific delegations of power and thus may prefer to keep laws informal and fluid. If codes are promulgated, they may be modeled on Euro-American legal systems or on those of other tribes. These may not be culturally compatible and may require resources for implementation that do not exist in the particular Tribal community.

Protocol

Issues of proper protocol when dealing with Tribal government leaders can make or break the evaluation of Indian programs. Setting up and following a proper protocol when dealing with indigenous peoples is of such concern that some organizations have written specific guidelines for researchers to follow (discussed below). However, a simple guideline for researchers to follow when dealing with Tribal leaders is to consider them as one would the President of the United States and/or members of Congress. Tribal Chairs and members of Tribal Councils are the chosen representatives of sovereign peoples. Many are elected. In some instances they accede to their positions through inheritance or birth, or are chosen by the spiritual or clan leaders of their communities. They carry a heavy mantle of responsibility and should be accorded great respect.

It is also essential that researchers or evaluators realize that they bring to any encounter with indigenous leaders their own preconceptions and biases. These biases should not be denied but instead realized and compensated for. It is natural and normal to have opinions about how things should be done. It is critical, however, to put these preconceptions aside and to realize that just because Indian Tribal leaders may speak English, dress in Euro-American style, and be highly educated does not mean that they necessarily think and perceive issues in the same manner as might members of the majority community. The lives of Indian people may proceed pursuant to a rhythm different from those in the majority community. It is important for evaluators to realize that ceremonies and rituals often take precedence, even over previously scheduled interviews with evaluators. This rhythm should be appreciated by the evaluator as an indication of a healthy community and not treated impatiently or as avoidance.

The first step prior to any contact with the tribe should be a literature review on the tribal people who are to be evaluated, as would be done if working in a foreign country. The evaluator should be introduced to representatives of the Tribal government, as well as to any others who have approved the grant proposal or have any responsibilities toward the project funded by the grant.

Evaluation Guidebook

It is essential that the evaluator involve tribal people from the inception of the evaluation. The evaluator should be prepared to explain who they are, why they are there, what they are doing, and how they are proceeding, and to explain this again and again, often to small groups that include persons to whom the explanations were given before. Opportunities to attend social events or activities should be valued because they can result in the development of a relationship that can assist an effective evaluation.

It is also essential to value conversation without insisting on directing it, lecturing, or filling in silences. There are natural breaks in conversations among tribal people, when thoughts are formed that are deemed pertinent. It is important that this process take place within its normal rhythm and that it not be hurried. Evaluators who conduct themselves as learners, rather than as teachers, will hear and learn more.

Issues of Tribal sovereignty and self-determination underscore many contacts and decisions made in Indian Country. If tribal leaders and people believe that the concept of Tribal sovereignty is understood and honored by researchers, they will be more cooperative and forthcoming. This cooperation can go a long way in facilitating the evaluation of programs.

The Global Coalition for Bio-Cultural Diversity of the International Society of Ethnobiology has developed a model covenant to help guide researchers in proper protocol with indigenous communities. Although this covenant is focused on the protection of intellectual property rights for indigenous peoples, it contains language that may be helpful for effective and impartial evaluations of Indian tribal programs. Pertinent sections read as follows:

Covenant Between a Responsible Corporation, Scientist or Scientific Institution and an Indigenous Group

Prologue

Indigenous peoples are unanimous in identifying their primary concern as being self-determination, which subsumes such basic rights as recognition of and respect for their cultures, societies, and languages, as well as ownership over their own lands and territories, and control over the resources that are associated with those lands and territories...This COVENANT should not be viewed as a finished product defining equitable partnerships, but rather a process of consultation, debate, discussion, and creative thinking from the many peoples and groups concerned...

Spirit of the Covenant

--This Covenant is celebrated in order to:

Support indigenous and traditional peoples in their fight against genocide and for their land, territory, and control over their own resources, while strengthening the culture and local community through recognition and support of the groups' own goals, values, and objectives...Thereby establishing a long term relationship built through joint decision-making based upon the principles of equality of relationships and protection of

Chapter 15: Special Issues for Evaluating Projects on Indian Tribal Lands

traditional values, knowledge and culture; if these basic elements are not respected, then the Covenant is endangered, and along with it, the spirit of trust and partnership between responsible businesses/scientists/institutions and local communities that is essential for the future well-being of the planet...

Basic Principles to Be Exercised by All Partners

- 1. Equity of partners, including..joint planning and goal setting, informed consent, and full disclosure in all aspects of the project, including results...*
- 4. Dedication to the promotion of harmony and stability within a group, between groups, and within the region, meaning that activities creating tensions (between indigenous and non-indigenous alike) are contrary to the Spirit of the Covenant.*
- 5. Confidentiality of information and resources, meaning that information imparted by the indigenous group to the Partner cannot be passed on to others without consent of the Giver.*
- 6. Continual dialogue and mutual review, supported by independent monitoring...*
- 8. Development, strengthening, and support of local (indigenous and non-indigenous) educational, health, research, and non-governmental institutions...*
- 10. Establishment of local autonomy and control over all aspects of the projects as early as possible.*

CONCLUSION

The fair, thorough, and unbiased evaluation of programs in Indian Country is essential if tribal governments are to be able to successfully compete for much-needed funds. Although the creation and carry-through of appropriate programs is the responsibility of the tribes, any assistance that they can receive through an impartial and even sympathetic evaluation can go a long way in helping tribes to develop needed competency and to reinforce their worthy efforts.

The creation of an equitable partnership between evaluator and program should be welcomed. The processes of joint planning and goal setting, as well as the continual discussion and mutual review of the evaluative process, will enrich both partners and make present and future efforts more comprehensive as well as more enjoyable.

The evaluation of a program does not have to be fearsome for those studied. With a close working relationship between evaluators and program staff it can be a growth experience, where the evaluator is viewed as an ally in the formation and articulation of the vision. It is to this challenge that evaluators should dedicate their work.