

## Chapter 6: Impact Evaluations

This chapter introduces the methods used to conduct impact evaluations. Even directors who are working with professional evaluators will find this background information useful as they progress through the steps outlined in the second half of this chapter.

### What Is an Impact Evaluation?

An impact evaluation answers such questions as “What is the effectiveness of the program?” or “What impact has the program had on participants?” Many Child Advocacy Center (CAC) directors are interested, for example, in knowing whether the cases processed through a CAC result in less system-induced trauma to children than traditional methods of processing cases. They also want to know the long-term outcomes of children served by a CAC. Although children’s well-being is of paramount concern, CAC directors also want to know whether cases processed through a CAC using the multidisciplinary approach are better investigated than cases processed through the traditional methods (e.g., law enforcement). Answering questions such as these requires an impact evaluation (Rossi and Freeman 1993, 116–117).

### Impact Evaluation Methodology

An impact evaluation compares program participants to nonparticipants with similar backgrounds on characteristics and experiences relevant to the evaluation (Rossi, Freeman, and Lipsey 1999). The comparisons are made so that causal statements can be made. For example, after comparing the levels of stress found in two groups of children alleging sexual abuse—those seen at a CAC and those not seen at a CAC—one can draw conclusions about the differences between the two groups and the reasons for the differences.

To better convey the complexities of comparing groups and how causal inferences can be made, the following section describes both the experimental and quasi-experimental methodology that form an impact evaluation.

### Experimental designs

Experimental designs have two basic and related characteristics: random assignment of participants and use of control groups.

#### **Random assignment of participants.**

“Random” is often considered synonymous with “arbitrary,” and to some extent this is the case. In experimental designs,

random assignment occurs when participants are assigned to one group or another based on chance alone. (A random numbers table may be used to make assignments; see, for example, the one found at <http://www.randomizer.org/form.htm>). Thus, participants are randomly assigned either to the group of individuals who will receive the intervention (i.e., their case is processed through the CAC) or to the group of individuals who will not receive the intervention (i.e., their case is not processed through the CAC). Those who receive the intervention are called the “treatment group”; those who do not are called the “control group.”

The underlying assumption of random assignment is that systematic differences between groups that might affect the outcome will be eliminated because each participant has an equal probability of being assigned to each group. Thus, if differences are found between the groups, the evaluator can be more confident that the differences are due to the CAC intervention rather than to some other cause.

**The use of control groups.** Control groups allow evaluators to make comparisons using such phrases as “better than” and “more than.” The control and treatment groups should be equivalent in all important and relevant respects. For example, members of both must be alleged victims of child sexual abuse (CSA). The only important difference between them is whether they received CAC services; because the participants are equivalent on all other relevant characteristics, causal inferences can be made (i.e., differences between the groups are due to the CAC intervention).

### Quasi-experimental designs

Experimental designs are the most rigorous methodologically. Many real-life situations simply do not lend themselves to this type of design, either for ethical or

practical reasons. In such cases, quasi-experimental methods may be used. However, quasi-experimental designs are less methodologically rigorous than experimental designs. Quasi-experimental designs have two primary characteristics: nonrandom assignment of participants and use of comparison groups.

**Nonrandom assignment of participants.** Nonrandom assignment of participants means that individuals are not randomly assigned to one group or another, as they are in experimental designs. Membership in a group has nothing to do with chance. Rather, there are naturally occurring groups that existed prior to the study and thus are not the result of the intervention. For example, one group may consist of children’s cases processed in a jurisdiction with an existing CAC and another group may consist of children’s cases processed in a nearby jurisdiction through the police department because there is no CAC.

**Comparison groups, rather than control groups.** The term “comparison group” is used in quasi-experimental designs and the term “control group” is used in experimental designs to distinguish the difference in methodology. The term “comparison group” denotes the inability to ensure there are no differences between the two groups because participants are not randomly assigned. Although comparison groups are not as “pure” as control groups, they are useful in making comparisons with the treatment group. The treatment and comparison groups should be as similar as possible in all important and relevant aspects.

Two types of comparison groups can be used in quasi-experimental designs.

**Simple comparison group design.** As mentioned, the comparison group should be as similar as possible in all relevant

characteristics, with the exception of exposure to the intervention. Therefore, for a fully operational CAC, an appropriate comparison group would be a group of children whose cases are processed through the conventional criminal justice system (for example, in a nearby jurisdiction that does not have a CAC).

A potential problem with using a nearby jurisdiction as a comparison group (aside from obtaining the cooperation of the agencies in that jurisdiction) is that there may be some systematic difference between the two jurisdictions. For example, the neighboring jurisdiction may have a significantly higher unemployment rate or lower average income levels. Thus, any differences found between the comparison group and the treatment group may be due to factors other than the CAC's intervention (e.g., economic resources).

**Pre-post design.** In a pre-post design, the comparison group would be children whose CSA cases were processed before the CAC opened. Once the center opens, the treatment group becomes the children whose CSA cases are processed through the center. Thus, for a center that is still in the planning stages, a pre-post design is appropriate.

*"We did ask kids about their feelings about being here for a medical exam. We had a before-and-after questionnaire. When they walked in, they were scared and didn't want to be here. When they left, 97 percent said they'd come here for the post-exam."*

The pre-post design reduces the potential systematic differences in comparison groups because all children come from the same jurisdiction. However, the majority of centers do not implement an evaluation prior to opening the center and thus most centers cannot use this design.

### **The case for quasi-experimental designs.**

Although quasi-experimental designs are less methodologically rigorous than experimental designs, they can yield credible estimates of the effects of ongoing programs. Quasi-experimental designs require strong theory and important assumptions about how people behave. Thus, evaluators who use quasi-experimental designs should think about the following issues:

- What will happen to the participants as a result of the intervention?
- What if-then statements is the evaluation using?
- Did the program have its intended effects? Was causality established?
- Were the measures focused on services provided?

## **Steps in Conducting an Impact Evaluation**

Impact evaluations involve nine steps:

1. State the impact evaluation's objective.
2. Develop the questions the evaluation should answer.
3. Predict the outcomes (i.e., state the hypothesis).
4. Select the impact evaluation's design.
5. Select the treatment and comparison/control groups.
6. Recruit participants.
7. Consider the long-term impact.
8. Identify influencing factors (i.e., moderating variables).
9. Select measurement instruments.

### **Step 1. State the impact evaluation's objective**

Developing the impact evaluation's objective is the first step. For example, an objective in developing a CAC might be "to reduce the amount of system-induced trauma that children would otherwise experience while in the criminal justice system."

### **Step 2. Develop the questions the evaluation should answer**

Next, restate the objective as a question. For example, "Do children whose CSA cases are processed through a CAC experience less system-induced trauma than children whose CSA cases are processed through the conventional criminal justice system? How do these children fare in the long run?"

### **Step 3. Predict the outcomes (i.e., state the hypothesis)**

After the objective is stated and the questions asked, the hypothesis needs to be clarified. For example: "Victims of CSA whose cases are processed through a CAC will experience significantly less stress (as measured by the Trauma Symptom Checklist) than children whose cases are processed through the conventional criminal justice system."

Making predictions that can be tested (i.e., forming hypotheses) is critical to research and evaluation because predictions force you and the evaluator to consider the relationships between variables, as well as the explanation for those relationships, before any data are collected.

### **Step 4. Select the impact evaluation's design**

Decide whether to use an experimental or a quasi-experimental design. Given the

ethical and practical considerations, most CACs will find that a quasi-experimental design is most appropriate.

### **Step 5. Select the treatment and comparison/control groups**

Next, determine eligibility criteria and decide who will be selected for the treatment and comparison/control groups. For the treatment and control groups, select children who meet the following criteria:

- Referred to the CAC for a CSA investigation.
- Under age 18.
- Reside within the CAC's jurisdiction.

For the comparison group, selection will depend in part on which type of design is used.

#### **Pre-post design using CAC children.**

If the CAC is in the planning stages, the center may be able to select a group of children whose cases are being processed through the current system or have been processed in the past, such as all children whose CSA cases were processed 1 year prior to the CAC opening. To do this, it is necessary to enlist the assistance of the multidisciplinary team's (MDT's) agencies to collect data on children whose cases were processed before the CAC opened. Although changes in procedures are probably already in progress (e.g., an MDT may already exist), comparisons may still be made. Some agencies may be concerned that a CAC is trying to make the existing systems look deficient by using children whose cases are processed through these various agencies. But given that these agencies have already agreed to develop a CAC, it may be easier to obtain their cooperation than if a CAC is not in the planning stages.

**Nearby jurisdiction without a CAC.** If the CAC is already operating, then a sample of children from a nearby jurisdiction without a CAC may be an appropriate comparison group. However, the various agencies may not be cooperative because the evaluation may be perceived as trying to imply their deficiency. Therefore, establishing a relationship with cooperating agencies will require the utmost sensitivity far in advance of the evaluation's start date.

## Step 6. Recruit participants

Recruitment is discussed thoroughly in chapter 7 and reviewed briefly here.

**Determine the number of participants needed.** In a quasi-experimental design, a minimum number of participants are needed in both the treatment and comparison groups in order to conduct statistical tests of the difference between the groups. As a general rule, 20 participants are needed per group. A professional evaluator should conduct what is called a power analysis (Cohen 1992b). This is a method for determining how many participants are needed to detect differences between the groups.

**Recruit other agencies.** During the planning phase, obtain cooperation from the various agencies who will participate in the evaluation. It may foster cooperation to include agency representatives on your evaluation team. Depending on the center's relationship with the representatives of the partner agencies, the process might begin by enlisting the cooperation of supervisors and then explaining the evaluation to the line employees.

Begin planning the coordination effort early because it will take some time for the process to work smoothly. Although flexibility is desired, you should have a general coordination plan in mind prior to approaching the decisionmakers in each

agency. A fairly well developed coordination plan should be in place before the evaluation effort is explained to the line employees. Ask for feedback from line employees and take their suggestions into consideration.

**Coordinate with other agencies.** There are several ways to coordinate the process of recruiting families from other agencies. For example, a victim advocate from the police department or in the Child Protective Services (CPS) agency may contact a CAC evaluation member (such as the data collector) when a CSA case comes into the department or agency. The CAC team member can go to the police station or CPS office to make the initial contact with the family. The center may want either the police officer or the CPS worker to introduce the data collector to the family so the family is assured of the evaluation's legitimacy. Although this is an ideal scenario, it is not always possible, so it may be helpful to develop an alternative procedure for recruiting families that fits the needs of various agencies.

**Recruit families from other agencies.** After a plan is in place to coordinate evaluation activities with the various agencies involved in the evaluation, the center may begin recruiting families from those agencies. As with all recruiting efforts, the center is required to follow ethical and legal mandates (see "Confidentiality," chapter 7). For example, participation must be voluntary, but offering incentives to participate is appropriate (Boruch 1997). Remember that the consent of parents must be obtained to ask children questions.

To ensure that recruitment is consistent, the center may consider developing a recruitment script. A standard script ensures that the center will include all the information that potential participants need to know, while also ensuring that all potential participants receive the same

information. Consider adapting the recruitment script in exhibit 6.1.

### **Step 7. Consider the long-term impact**

According to Yates (1996), “What happens as a result of human service provision may be different from what happens after human service provision.” Although it is possible to simply compare groups at the same point in time, the question most directors ask is whether the CAC helps children in the long run. Answering this kind of question requires long-term followup. This requires collecting information from both the treatment and the comparison/control groups during their initial CAC visit and at specific points of time in the future, such as 1 and 2 years after they leave.

How often participants are asked to complete the questionnaire depends on the center’s adopted theory of change (see chapter 9). Chapter 7 describes methods to recruit families and stay in touch with them over time.

### **Step 8. Identify influencing factors (i.e., moderating variables)**

Directors need to consider—and measure—a number of possible factors that could influence the effect the center has on children. Factors that influence the outcome are referred to as moderating variables—the relationship between two or more items that are influenced by another factor (Mark, Hofmann, and Reichardt 1992). For example, the relationship between CAC activities and child stress may be moderated by the relationship between the parent and the child. That is, children may experience lower levels of stress during the investigation when they have a positive relationship with their parent(s), whereas children may experience higher levels of stress during the investigation when they have a poor relationship

with their parent(s), all other factors being equal. Thus, the CAC program may have less effect on children who have a positive parent-child relationship and a greater effect on children who have a poor parent-child relationship.

A number of influencing factors should be considered and measured, such as characteristics of the interviewer, characteristics of the child and family, and social support (Berliner and Elliott 1996). There are also socioeconomic and political processes beyond the control of the participants that affect children, such as social support, health status, and economic self-sufficiency. The following are some additional factors to consider collecting data on:

- Mother’s support of the child.
- Type of abuse.
- Child’s relationship with the alleged perpetrator.
- Mother’s relationship with the alleged perpetrator.
- Mother’s level of distress.
- The level of trust the child has with an adult.
- Child’s level of depression.
- Time of disclosure.
- Child’s coping style.
- Family’s level of conflict.
- Family’s level of cohesion.
- Degree of court preparation (stress inoculation).
- Demographic characteristics of participants, such as age, sex, ethnicity, educational level, household income, household composition (head of household, family structure), disability status,

prior work history, health status, criminal record, and employment status.

- Geographic location of participant's residence, such as neighborhood, political boundaries, ZIP Code, census tract, city, and county.

### Step 9. Select measurement instruments

Appendix E contains several instruments for measuring child stress and trauma and influencing factors. Some are available only through a publisher and are described only briefly. Others are reproduced in their entirety. Select an instrument that is appropriate for the evaluation. Inclusion in this list does not imply endorsement. Please check each instrument for information on its validity and reliability.

#### Child stress and trauma impact evaluation questionnaires

- The How I Feel Questionnaire
- Child Anxiety Scale—Parent Form
- Family Stress Questionnaire
- Trauma Symptom Checklist for Children (TSC-C)
- Children's Depression Inventory (CDI)
- State-Trait Anxiety Inventory for Children
- Child Well-Being Scales (CWBS)
- Coping Responses Inventory—Youth Version
- Child Behavior Checklist (CBCL)
- Preschool Behavior Checklist (PBCL)
- Preschool and Kindergarten Behavior Scales (PKBS)
- Child Sexual Behavior Inventory (CSBI)

- Revised Children's Manifest Anxiety Scale (RCMAS)

#### Influencing factors impact evaluation questionnaires

- Children's Version of the Family Environment Scale (CVFES)
- Parenting Stress Index (PSI)—Third Edition
- Parent-Child Relationship Inventory (PCRI)
- Knowledge of Infant Development Inventory (KIDI) and Catalogue of Previous Experience With Infants (COPE)
- Conflict Tactics Scale—II
- Parent-Child Conflict Tactics Scale
- Exposure to Violence and Trauma Questionnaire
- Stressful Life Events Screening Questionnaire
- Family Adaptability and Cohesion Evaluation Scales (FACES III)—Family Version
- Family Environment Scale (FES)

### Additional Impact Evaluation Considerations

Several other issues should be considered when planning an impact evaluation.

#### Eliminating conflicting explanations

There are often multiple explanations for why changes occur in the target population. Therefore, it is important to eliminate as many competing explanations as possible to be confident that the program itself

is responsible for the evaluation results. There are two conflicting explanations unique to impact evaluations: history and maturation.

**History.** History may be relevant if the comparison group is different from the treatment group prior to the evaluation. For example, it would be problematic if children from the comparison group had higher levels of family conflict than children in the treatment group because differences between the two groups could be due to family levels of conflict (i.e., history) and not the program. Therefore, if random assignment of participants to groups cannot be made, take steps to ensure that both groups are equal on important variables. This can be done statistically if measures of influencing variables have been collected (e.g., family conflict).

**Maturation.** Maturation may be relevant when events outside the program cause the intervention group to change while children are in the program. For example, if an investigation is lengthy, a child may have a greater understanding of the investigation over time simply because of cognitive maturity. Accounting for history and maturation will help eliminate conflicting explanations for the findings.

### **Preexisting characteristics**

The concern here is that change in participants is due to the passage of time and not as a result of the CAC. One way to control for this type of error is to collect measures on characteristics that might change over time, such as age.

### **Timing issues**

A preferred design is one in which information is collected from participants both before (or as) they enter the program, and after they leave the program. This design

provides information about how participants were before they entered the program and after they completed the program.

### **Frequency issues**

A strong design is one in which information from participants is collected multiple times, including after they leave the center, to understand the long-term impact of the program on participants.

### **Societal influences**

Changes in existing laws, services, or public awareness may affect the evaluation's outcomes; therefore, more information on these factors may need to be gathered. For example, a new law may make it easier to convict perpetrators, allowing a more expedient prosecution of a child's case.

### **Selecting individuals to participate in the evaluation**

Selecting (i.e., sampling) participants for the evaluation is always a difficult challenge but critically important because who participates in an evaluation can make a tremendous difference in the results. Who participates in the CAC evaluation should be less of an issue because all individuals referred to the center should be eligible for participation in the evaluation. However, a significant challenge that will need to be addressed (and that must be explained) is refusal to participate in the evaluation. Without explanations for why clients refuse to participate, results will not be reflected accurately and will undermine the final report. Therefore, documenting refusals and collecting basic information on them for comparison with the final group of participants is critical. Individuals may refuse to participate for a number of reasons, and it may be beneficial to consider

tracking their reasons. For example, participants may refuse to participate because of lack of interest, inconvenience, busy schedules, objection to the approach, objectionable topic, poorly worded questions, distrust, or dislike of the recruiter.

### **The need for program monitoring**

Like an outcome evaluation, an impact evaluation typically requires a program monitoring component, because it is important to know how the existing system is being implemented, as well as how children respond to that system.