



*Serving the public and non-profit sectors through independent program evaluation, applied research, and technical assistance.*

## EVALUATION BRIEF

### Critical Issues in Evaluating Child Welfare Programs

September 2009

#### Introduction

With an annual budget of more than \$7 billion, the Children's Bureau within the Administration for Children and Families (ACF), U.S. Department of Health and Human Services, awards funds to both States and Tribes on a formula basis as well as to individual organizations that successfully apply for discretionary grants.<sup>1</sup> The Children's Bureau awards discretionary grants to promote research on and the development of innovative child welfare programs and practices that improve permanency, safety, and well-being outcomes for children and their families in or at risk of entering the child welfare system. Awards are made through a competitive peer review process to organizations that include State, Tribal and local government agencies, faith-based and community-based organizations, and other nonprofit and for-profit groups.

As a stipulation of their awards, child welfare discretionary grantees are expected to implement program evaluations to determine their success in implementing their funded programs and achieving the child and family outcomes outlined in their original grant applications. Over the past several years, discretionary grantees have confronted similar methodological and logistical challenges to implementing high-quality research. This evaluation brief examines several critical issues that have affected the evaluation efforts of many child welfare discretionary grantees and reviews a range of strategies that grantees may consider to address these issues effectively, including many approaches suggested by grantees based on their own experience, knowledge, and expertise. Common evaluation challenges that will be discussed in this brief include:

- Estimating the size and characteristics of the evaluation's target population;
- Obtaining Institutional Review Board (IRB) approval and/or client consent to conduct research;
- Identifying appropriate data collection instruments, tests, and protocols;
- Training staff in data collection and analysis methods; and
- Accessing and using child welfare data from State and local information management systems

---

<sup>1</sup> Administration for Children and Families (2009). *Programs and funding*. Retrieved July 27, 2009, from [http://www.acf.hhs.gov/programs/cb/programs\\_fund/index.htm](http://www.acf.hhs.gov/programs/cb/programs_fund/index.htm).

## Critical Issue #1: Estimating the Target Population

An accurate assessment of the size and characteristics of the population targeted for service is a crucial first step in the program planning process that also has significant implications for the success of proposed evaluation activities. During the initial planning stages of a new project, it is common for organizations to adopt a “build it and they will come” mentality in which they overestimate the likely number of program participants. The reasons underlying these often overly optimistic predictions of program enrollment include erroneous assumptions regarding:

- The depth and breadth of need in the community for the new program (i.e., existing services in the community may already address the perceived need);
- The true size and characteristics of the target population (e.g., the people for whom the new program was designed may have a different mix of presenting problems or fall within a different age range than expected, or the number of people eligible for service may be smaller than anticipated); and
- Prospective participants’ willingness to participate in the program. Even if the need for the new service has been established and a sizeable target population exists, project staff may encounter resistance from families to enroll in services.

Problems with the participant screening and assessment process may compound the issues noted above. For example, the eligibility criteria for program participation may be ambiguous or intake workers may have received inadequate training to identify the best candidates for the new program. All of these factors can have negative implications for an evaluation in that they may reduce sample sizes, and subsequently, the amount and quality of data available for analysis. Less high-quality data makes it more difficult to draw meaningful conclusions regarding the true impacts of the child welfare program in question.

**Strategies that evaluators can use to mitigate problems with identifying and screening the appropriate program participants are noted below:**

- *Conduct a community needs assessment:* Before the planning stages of a new program even begin, a systematic investigation of existing services and resources in the target community, as well as of the demographic characteristics and presenting problems of the prospective target population, will help to establish the need and appropriateness of the new program. Although their purpose is somewhat different, needs assessments often involve the same research tools as evaluations (e.g., interviews and focus groups with community leaders, analysis of data from existing health and human services information systems, directly canvassing residents to collect information on community needs). Although a needs assessment can be time consuming and costly, it may be a worthwhile investment if it establishes an actual need for the new program in the community or reduces future spending on unnecessary services and evaluation activities. Moreover, if implementation of the proposed program proceeds as planned, findings from the needs assessment can assist an evaluator in understanding

contextual factors in the target community that influence or explain observed participant outcomes.<sup>2</sup>

- *Develop well-defined eligibility criteria:* You can eliminate much of the uncertainty in identifying the most appropriate recipients of program services by establishing clear and concrete eligibility criteria for program participation. Well-defined criteria will reduce time, money, and personnel resources spent on outreach and recruitment of people who in the end are not good candidates for the new program. Make the eligibility criteria as discrete and specific as possible (e.g., individuals between the ages of *X* and *Y*, children in out-of-home placement for *X* number of months, people with a score of *X* or higher on a standardized assessment tool) to minimize ambiguity.
- *Formulate clear intake, screening, and assessment procedures:* Further guesswork can be eliminated from the participant recruitment and enrollment process by creating a systematic protocol for determining the eligibility of prospective participants. Using the explicit criteria noted above, create standardized intake and screening forms that contain questions and collect data that verify eligibility for the new program. When appropriate, standardized assessment instruments can be used to determine the presence or severity of problems that your new program is designed to address (e.g., using the *Rapid Alcohol Problems Screen* [RAPS] to assess problem drinking patterns as part of a substance abuse treatment and recovery project).
- *Provide adequate staff training:* Successful deployment of the eligibility criteria and assessment protocols described above depends in large part on the degree to which program staff can implement them effectively. As such, it is imperative to provide all staff involved in the intake, screening, and assessment process with sufficient training so they understand the purpose and importance of the eligibility criteria and assessment tools; know how to apply the criteria appropriately and fill out and/or score screening and assessment instruments; and make appropriate decisions regarding enrollment into the new program. For certain tools or protocols of a highly technical nature, special training may be required or the services of a clinician trained in the relevant screening and assessment processes may be necessary.

By beginning with a clear sense of the likely size of your target population and establishing clear guidelines for determining who belongs to this population, you will have much greater confidence in the amount and quality of participant data that will be available for study following program implementation.

---

<sup>2</sup> For more information on conducting needs assessments, see Soriano, F. I. (1995). *Conducting needs assessments: A multi-disciplinary approach*. Thousand Oaks, CA: Sage Publications; and Riviere, R., Berkowitz, S., Carter, C., & Ferguson, C. (Eds.). (1996). *Needs assessment: A creative and practical guide for social scientists*. Bristol, PA: Taylor and Francis.



## Critical Issue #2: Obtaining IRB Approval and Client Consent

Even when the need for service has been established and a target population has been clearly defined, problems with winning the trust of prospective participants and receiving approvals to conduct research (particularly among vulnerable populations) can delay evaluation activities. Most program evaluations, particularly those conducted under the auspices of colleges and universities, require IRB review and approval before research activities involving the collection of primary participant data may proceed. Designed to protect the rights and privacy of research subjects, IRB review can create challenges for the timely implementation of an evaluation. The IRB approval process can last many months (or longer), especially if modifications to the proposed research design, data collection protocols, or informed consent procedures are required. This process is slowed even further when IRB approvals from multiple organizations are necessary; for example, a program evaluation funded by a local child welfare agency and implemented by university faculty may require approval from the IRBs of both the child welfare agency and the university that employs the faculty. While IRB approval is pending, precious months to collect primary data slip away.

The establishment of clear procedures for obtaining informed consent from program participants to collect personal information is a key ethical requirement of all IRBs; however, even when IRB approval is received, some participants may be reluctant to give their informed consent to participate in research activities due to concerns about privacy, how their personal information will be used, or general mistrust of government and other large institutions. The more frequently participants refuse to cooperate with data collection activities, the less information will be available to measure program impacts.

**Strategies that evaluators can consider to avoid delays resulting from the IRB review process and to win the trust of potential evaluation participants are highlighted below:**

- *Use secondary data whenever possible:* Large amounts of valuable information that does not entail primary data collection may already be available for your program evaluation. These secondary data sources include client data from child welfare information management systems, case notes, and other information on service delivery and outcomes contained in case managers' files and other hard-copy records. Because these data sources already exist for administrative and case management purposes, access to them does not usually require participant consent. In addition, because the use of secondary data generally poses fewer risks for prospective research subjects, IRB approval can often be obtained through an expedited review process.<sup>3</sup>
- *Begin the IRB process early:* Whether you will seek expedited IRB review or must undergo the full IRB review process, it is imperative to complete and submit an IRB application as soon as possible. For Children's Bureau discretionary grantees, the requirements for IRB approval from pertinent institutions can be determined and

---

<sup>3</sup> Expedited review consists of an abbreviated procedure most often conducted solely by the IRB chairperson. Research and evaluation projects that are deemed to present only minimal risk to human subjects are often eligible for expedited review. For more information regarding IRBs and the IRB application process, see James Bell Associates (2008, January). *Understanding the IRB*. Arlington, VA: Author.

preliminary work on the IRB application can begin before a grant application is approved. To begin the IRB application process you will need to have identified the individual(s) responsible for implementing the evaluation and have made significant progress in developing your evaluation plan. Assuming these components are in place, you can submit your IRB application as soon as your Federal grant is approved.

- *Obtain buy-in from community leaders and residents:* Ultimately, your success in obtaining informed consent from prospective program participants depends on your ability to earn their trust, specifically, their conviction that sharing personal information will result in no harm to them or their loved ones. Leaders as well as ordinary people who are part of the community to which program participants belong can serve as intermediaries to assuage concerns among prospective program participants about the research process. Strategies for garnering support from leaders and other members of the community include community forums to discuss the research process and brochures and other marketing materials that describe the evaluation and address frequently asked questions. Throughout this process, it is essential to communicate with community stakeholders early and frequently regarding the purpose of the evaluation, its importance, the benefits and potential risks of participation, and the merits of the proposed research design. Concerns and misinformation regarding the evaluation and research design can also be dispelled. In addition, community forums and similar venues can give members of the community an opportunity to share their ideas for improving the evaluation and for securing the cooperation of program participants.
  
- *Use incentives:* Monetary or in-kind incentives can serve as an effective device for increasing motivation to participate in surveys, interviews, and other primary data collection activities. Incentives should be big enough to be meaningful to your target audience (e.g., one dollar will probably not be sufficient and may be perceived as insulting) but not so large that it is perceived as a bribe or somehow coercive (e.g., a homeless person who is offered \$100 to complete a survey may feel that he cannot refuse to respond even if he does not really want to). Moreover, different types of incentives – whether cash or in-kind incentives such as gift cards – may be more appropriate or effective with certain target populations than with others. The process for securing buy-in from community stakeholders described earlier can serve as another source of information on the judicious and effective use of incentives. Ultimately, the nature and size of any incentives will also be affected by your budget for evaluation activities.

### Critical Issue #3: Identifying Data Collection Instruments

Surveys, questionnaires, and tests in various forms constitute some of the most common tools for collecting primary data from program participants. They are widely used because they (1) can capture a large amount of information in a relatively short period of time; (2) are able to measure changes across a wide range of outcomes (including changes in knowledge, skills, attitudes, and behavior); and (3) codify data in a form that can be readily processed, analyzed, and reported. However, the use of data collection instruments is fraught with challenges, including the issue of whether to use existing standardized tools or customized



“home grown” instruments. Drawbacks associated with both types of instruments are discussed in more detail below.

### Standardized Instruments

Standardized instruments generally include surveys, questionnaires, and tests that have been developed by third parties with special clinical training and/or academic experience; contain a uniform set of questions and require the use of standardized administration procedures; have been normed against one or more populations of individuals to determine the normal range of responses or scores; and have undergone statistical tests to assess their validity (i.e., the degree to which they measure the constructs they purport to measure) and reliability (i.e., the extent to which they produce consistent results across different testing environments, times, and test administrators). Standardized instruments enjoy many advantages, most notably the fact that they have already been developed and validated and are therefore ready for immediate use. They do, however, have notable disadvantages:

- *Identification takes time:* Scores of standardized instruments have been developed and tested over the last few decades in a range of medical and human service fields, including mental health, substance abuse, early childhood development, parenting skills, and child maltreatment risks. Choosing a tool that works best for measuring improvements in the specific areas your program is designed to address may require a substantial investment of time and a high level of professional knowledge and experience.
- *Program fit:* Despite the wide range of standardized tools available in many fields, an instrument may simply not exist that is adequately “tailored” to the unique features of your program and that addresses the constructs and changes in participants you wish to measure.
- *Cost:* Although some standardized tools are in the public domain, many are copyrighted and must be purchased, sometimes at substantial cost. Many project directors would be duly hesitant to make the investment in an instrument that later proves to be inappropriate for measuring change in their service populations.
- *Training:* Some standardized tools require specialized education and training to administer properly or to interpret the results; providing this training to designated staff or hiring professionals with the requisite skills further compounds evaluation costs.

### Customized instruments

An alternative to existing standardized tools is to develop a “home-grown” tool whose content is tailored to the unique features of your program and is sensitive to the characteristics of your program’s target population. In addition, the development of an instrument in house can avoid some of the copyright, cost, and training issues that characterize standardized tools and can also foster a greater sense of ownership and buy-in to the evaluation among project staff. However, customized instruments present their own unique challenges:

- *Time and effort:* Although home-grown instruments circumvent the up-front costs of purchasing a standardized tool, they are not truly free because they can require substantial staff time and effort to create and revise; these factors constitute real costs in terms of staff and evaluator productivity. Moreover, time spent developing a home-grown tool translates into delays in collecting valuable data from program participants; in contrast, an off-the-shelf standardized tool can often be deployed immediately after program implementation.
- *Validity and reliability:* Although they may appear to be good measures of change because they are tailored to the features of your program (i.e., they have face validity), home-grown instruments may in fact produce results that are neither valid nor reliable because they have not undergone the rigorous testing and validation process of standardized instruments. In these circumstances, a standardized instrument that appears to have less face value may actually do a better job of measuring the impact of your program. If use of a customized instrument is essential, one option is to undertake a validation study to vet the psychometric properties of the instrument and make modifications until acceptable levels of validity and reliability are achieved. However, validation studies require substantial amounts of time, money, and technical expertise – resources that many child welfare grantees cannot afford to expend.

A number of options exist for addressing the pitfalls of both standardized and home-grown tools:

#### For standardized instruments

- *Take advantage of on-line resources:* In recent years, a number of Web-based resources have become available that provide information regarding the use and psychometric properties of a wide range of standardized instruments in multiple fields. For program directors and evaluators unfamiliar with survey research in a particular profession or academic field, these on-line databases provide a fast, easy, and inexpensive way to review the features of large volumes of tools and narrow the focus on those instruments best suited for their evaluation needs. Examples of on-line resources for selecting appropriate standardized instruments include detailed instrument reviews at the website of Psychological Assessment Resources, Inc. (<http://www3.parinc.com/products/default.aspx>); and an annotated compendium of child and family measurement tools available through the National Resource Center for Community-Based Child Abuse Prevention (<http://www.friendsnrc.org/outcome/toolkit/annotalpha.htm>).
- *Adapt subscales from standardized instruments:* Selected subscales or items from standardized instruments are sometimes sufficient for measuring the constructs most relevant to your program; rather than attempting to administer an instrument in its entirety, these selected subscales can be adapted into a shorter instrument that result in more meaningful findings while reducing participant response burden. It is advisable, however, to contact the instrument's author(s) for permission to use

selected instrument items and to ensure that the adapted items will maintain adequate validity and reliability.

For both standardized and customized instruments

- *Pilot test:* Whether you decide to use a standardized or home-grown tool, pilot testing of any prospective instrument is strongly recommended. Pilot testing involves administering the tool to a small sample of program participants (or to people very similar to those who comprise your program’s target population) to identify and correct problems with the content, wording, or format of the instrument itself, or with the test administration procedures, before it is widely administered to participants as a whole. Although a pilot test is not a substitute for a full validation study, it will nonetheless detect the most serious problems with a prospective instrument and help eliminate those that are clearly inappropriate for your program or target population. For a little extra time and money, you can avoid implementing an unsuitable tool before you are too far along in the evaluation process and identify an alternative instrument that is a better fit.
  
- *Consider alternative data collection methods or formats:* A written instrument may not be the only or even the best method for collecting useful data about the effects of your program. For example, direct observation using checklists or videotaped interviews may provide more valid information regarding changes in participant behavior than self-report surveys. When a written survey remains the research method of choice, alternative administration formats should be considered if they improve data quality and quantity and/or reduce costs. For example, rather than investing the substantial time and resources necessary to design and administer a telephone survey among a random sample of participants, it might be more cost effective to survey an entire population using a paper mail-in or web-based survey without significantly compromising data quality. Similarly, a pre-post test targeted at a population with low literacy levels could be administered orally by trained program personnel, resulting in better response rates to test items.

**Critical Issue #4: Staff Training**

With limited budgets and few in-house evaluation resources (e.g., specialized research and evaluation departments), organizations sometimes turn to front-line staff and supervisors to collect data and conduct other evaluation activities. Although this may seem like a sensible solution at first, there are serious limitations to relying too heavily on program staff:

- Program staff frequently lack the appropriate education and experience to design, plan, and implement a full-scale evaluation effort. Even routine tasks such as collecting, analyzing, and managing evaluation data often fall outside of the experience of most workers and their supervisors.
  
- As is the case in many government and non-profit organizations, staff and their managers are often already overburdened with a range of job responsibilities. Adding an evaluation to the “to do” list on top of pressing programmatic duties may simply be

infeasible from a workload standpoint and may fuel resentment that undermines staff support for the evaluation.

- Even when staff with basic evaluation skills and time are available, they may not have ready access to tools and resources (e.g., statistical analysis software, budgets to purchase standardized assessment instruments) that are necessary to conduct an evaluation effectively.

When a substantial role for program personnel in collecting data or implementing other evaluation activities is unavoidable, **the following steps can assist in maximizing the contributions of staff while minimizing additional work burdens:**

- *Make clear the benefits of evaluation:* It is not enough to simply gain the compliance of front-line staff and supervisors to participate in research activities; their full support and cooperation can best be won by demonstrating the potential of evaluation to enhance their job performance, improve participant outcomes, and increase the overall quality of their work experience. Strategies for demonstrating the benefits of evaluation to program staff include involving them in the early planning stages of the evaluation (e.g., building the program's logic model, selecting appropriate data collection instruments); sharing and discussing evaluation findings with staff on an ongoing basis, particularly those that are relevant to their day-to-day work responsibilities; and inviting staff to share their knowledge and experience to address new or ongoing evaluation challenges (e.g., making suggestions for improving low response rates to surveys or increasing participation in other data collection activities). Program staff that perceive a personal stake in the success of the evaluation will be more motivated to play an active role in evaluation activities.
- *Clarify and simplify roles:* To the extent possible, keep the responsibilities of personnel involved in the evaluation limited to clear and simple roles. For example, it may be cost effective to hire a third-party consultant to develop the overarching evaluation design and assist in the selection or development of data collection tools, but then leave routine activities such as administering surveys, conducting observations, and data entry and cleaning to program staff. It is often advisable to delegate these routine activities to a small cadre of staff in order to minimize the impact on workloads throughout the organization and to maintain greater control over the work quality and consistency.
- *Maximize the use of local resources:* Many organizations may be able to make use of free or low-cost local human resources to assist with many routine evaluation activities. For instance, interns and graduate students from local post-secondary institutions may welcome the opportunity to hone their research skills by conducting interviews and focus groups or by performing statistical analyses of pre-post test data. Organizations with volunteer programs can also explore ways to involve volunteers in basic logistical activities (e.g., preparing paper surveys for mailing) or make use of special skills possessed by certain volunteers (e.g., website or database design).
- *Use simpler and shorter tools:* Selecting or developing surveys, questionnaires, and tests that are short and easy to administer is a straightforward way to reduce data

collection burdens on staff, both in terms of the time required to administer the tools and the effort to clean and enter the resulting data. The advantages of short and simple tools must be weighed against the need to choose instruments that are valid and relevant to the key features and target population of your program.

- *Maximize the use of secondary data sources:* Finally, the burdens imposed on staff by primary data collection can be mitigated by relying as much as possible on secondary data sources. As previously noted, large amounts of valuable information may be readily available in your or a partnering organization's child welfare information system or other relevant databases. Although information from these sources must still be extracted, cleaned, and analyzed, substantial time savings can be realized by avoiding the laborious process of planning and conducting primary data collection activities.

### Critical Issue #5: Accessing and Using Child Welfare Data

A frequent assumption made by many evaluators is that they will have access to critical information regarding program participants through information management systems maintained by local or state authorities. In reality, they often discover that they are barred from or severely delayed in accessing the necessary information; that certain data elements that were presumed to be in the database do not in fact exist; or that the quality of the available data are so poor as to be of limited value.

Evaluators can avoid these unexpected surprises by taking the following steps:

- *Secure buy-in:* During the initial planning stages of your evaluation, it is critical to obtain support for the evaluation effort from managers at the organization in which the data are housed and to assuage concerns they may have regarding data privacy laws or policies. In addition, the cooperation of database administrators and analysts is essential to ensuring that the data will in fact be produced in a timely manner. Strategies for securing buy-in from these key players include (1) cultivating contacts with program staff who can help you establish trusting relationships with personnel responsible for access to child welfare data; and (2) casting the issue of data access in as positive a light as possible, for example, by highlighting the use of the data for program development and continuous quality improvement. Formal presentations to key personnel and handouts that address frequently asked questions (FAQ sheets) regarding the ways in which you plan to use the data can help frame the issue of data access to your advantage.
- *Establish an MOU:* If you are a third-party evaluator representing an outside organization (e.g., a university or consulting firm), the creation of a Memorandum of Understanding (MOU) can serve as a formal means of codifying and enforcing your rights and responsibilities regarding data access. The MOU should delineate the specific data sets or data elements to which your organization will have access, key personnel responsible for extracting and transferring the data, and procedures for maintaining the privacy and security of sensitive client information. Of course,



informal buy-in from managers and other key personnel must first be in place before a more formal data sharing agreement can be finalized.

- *Run a pilot test:* Once access to the data is secured, request a sample extract to ensure that the requisite data (1) do in fact exist; (2) are valid (i.e., contain the information you actually need in a format you can use); and (3) are substantially complete. You can facilitate this process by crosschecking your pilot data with data from other information systems to assess the degree to which they coincide (e.g., demographic data such as age, race, and gender may exist in both a child welfare information system and in Medicaid or public assistance databases). As part of running the pilot test, verify whether the information management system maintains *historical data* on your target population, i.e., that old data are not overwritten by new information but rather are maintained in the system to facilitate longitudinal data analysis. By pilot testing prospective data sources, you can identify unsatisfactory results early on and invest time in cleaning and augmenting your data sets, or if necessary, select an alternative data collection strategy.

Although additional factors may affect the ultimate success of your research and data collection efforts, the resolution of the five critical issues discussed in this brief will go a long way toward improving the timeliness, quality, and usefulness of your evaluation.<sup>4</sup> For more information about strategies for addressing the child welfare evaluation topics explored in this brief, please contact a JBA team member at:

James Bell Associates  
1001 19<sup>th</sup> Street, North, Suite 1500  
Arlington, Virginia 22209  
703-528-3230 or 800-546-3230  
[www.jbassoc.com](http://www.jbassoc.com)

---

<sup>4</sup> A major issue not addressed in this brief involves the selection of an appropriate and practical research design. For a detailed treatment of this topic, see James Bell Associates (2009, September). *Selecting an evaluation approach*. Arlington, VA: Author.