EVALUATION BRIEF

Conducting Randomized Controlled Trials in Child Welfare Practice Settings: Challenges and Solutions[1]

November 2013

Introduction

The last two decades have witnessed a trend toward higher standards for accountability on the part of the recipients of grants awarded through Federal agencies. These expectations are also on the rise for child welfare organizations that receive funding through discretionary grant programs administered by the Children's Bureau (CB) within the Administration for Children and Families (ACF), U.S. Department of Health and Human Services (HHS). This push toward greater accountability and results in the child welfare field has fostered greater emphasis on the implementation of what are often referred to as evidence-based programs and practices (EBPs) (OMB, 2012). Although no consensus has emerged within the field regarding a standard definition or the key components of EBPs, in general they include services, practices, and interventions for which conclusive evidence of effectiveness exists based on findings from a rigorous and systematic evaluation.

As in other fields, the use of rigorous research designs is essential to building evidence-based programs and practices in child welfare (Gambrill, 2013). However, the child welfare field is generally regarded as behind other applied fields of social work and mental health in testing and building evidence regarding the effects of new interventions on child welfare populations and organizations. In this regard, one of the most debated issues involves determining the level of evidence necessary for a classification of "evidence-based," and as a corollary, which research designs are regarded as sufficiently rigorous to provide this evidence. Although no final consensus has emerged, some of the most prominent EBP classification systems, including the California Evidence-Based Clearinghouse for Child Welfare (CEBC)[2], the What

[1]This publication was developed by James Bell Associates on behalf of the Children's Bureau, ACF, HHS, under Contract Number GS10F0204K, Order Number 06Y010202. Its contents are the sole responsibility of the author and do not necessarily represent the official views of the Children's Bureau, ACF, or HHS. No portion of this document may be modified, reproduced, or disseminated without the written permission of James Bell Associates. Proper citation of this document is: James Bell Associates. (2013). *Evaluation Brief: Conducting Randomized Controlled Trials in Child Welfare Practice Settings.* Arlington, VA: Author.

[2]See www.cebc4cw.org/.

Works Clearinghouse developed by the U.S. Department of Education[3], and the Top Tier Evidence Initiative developed by the non-profit Coalition for Evidence-Based Policy[4], identify the randomized controlled trial (RCT) as the most methodologically rigorous design that offers the most definitive evidence of effectiveness. Despite the endorsement of these groups and by much of the research community in a variety of human service fields, RCTs remain comparatively rare in child welfare settings. For example, a search of CEBC's online database reveals that of the 103 programs documented by CEBC as having high relevance to the child welfare field, only 14 (11 percent) have been evaluated using one or more RCTs (CEBC, 2013). Familiar examples of child welfare-related interventions whose effectiveness has been established using RCTs include SafeCare (Chaffin, Hecht, Bard, Silovsky & Beasley, 2012), Homebuilders® (Frazer, Walton, Lewis, Pecora & Walton, 1996), and Trauma-Focused Cognitive Behavioral Therapy (TF-CBT) (Cohen, Berliner & Mannarino, 2010).

The purpose of this brief is to familiarize child welfare agencies, including current and potential Children's Bureau (CB) discretionary grantees, with RCTs and the value of implementing them in child welfare practice settings. Whereas many child welfare researchers are knowledgeable about and understand the importance of RCTs, it is imperative that agency administrators, front-line staff, and key service partners grasp the key elements of RCTs and their critical role in advancing the current state of knowledge regarding effective child welfare programs and practices. In addition, this brief highlights many of the common challenges faced by child-serving organizations in implementing RCTs in actual practice settings and suggests possible solutions that are informed by the experiences and lessons learned from current and former CB discretionary grantees.

Overview of Randomized Controlled Trials

The intent of an RCT is to test the effectiveness of a particular intervention in producing a desired outcome in a research participant. Another term for an RCT is an *efficacy trial*. An advantage of RCTs and experimental designs in general over qualitative studies, correlational studies, and quasi-experimental studies is their ability to rule out the impact of individual differences among participants through the random assignment of research participants to either an experimental condition or a control condition. With randomization, the researcher can be more certain that the results of a study are due to the intervention to which the participants were exposed rather than to unrelated personal or environmental variables that could affect observed outcomes.

In practice, prospective participants in an RCT are first recruited and assessed for their eligibility to participate in the study. They are then randomly assigned to either a treatment (or experimental) condition or to a control condition. Those participants assigned to the experimental condition participate in the intervention under investigation, while those assigned to the control condition generally receive "treatment as usual." The latter generally consists of any programs, services, or supports that would have been available to all participants in the absence of the new intervention. For example, in examining the effects of a therapeutic treatment for childhood trauma exposure, the control condition could receive an unrelated regimen of general psychotherapy, after which differences in outcomes between this control condition and the experimental group that receives the therapeutic treatment

[3] See www.ies.ed.gov/ncee/wwc/.

[4] See www.toptierevidence.org/.

would be compared; this difference in outcomes is referred to as the program effect. The hallmark of an RCT is that both groups of participants begin with an *equal chance* of being assigned to either the experimental or control condition.

The Importance of RCTs in Child Welfare

RCTs are often regarded as the "gold standard" for evaluating the efficacy of an intervention with the fewest qualifying assumptions because they are very good at isolating the impact of an intervention from other possible explanations for observed changes in participants' outcomes. Because the experimental and control groups look almost exactly the same—except that one group receives the new intervention and the other group does not—it is possible to conclude with more confidence that any differences in outcomes between the treatment and control groups are in fact due to the intervention and not something else. To appreciate the strengths of the RCT, and why it is generally preferable to answer questions about the impact of child welfare programs, it is useful to consider the limitations of other types of studies conducted in the social sciences.

One long-recognized drawback of other research designs is ambiguous causality. For example, a correlational study of an in-home maltreatment prevention program may identify an association between the introduction of the program and reduced maltreatment rates, but generally will not be able to pinpoint the specific variables (whether the prevention program itself or other factors) that are affecting the observed rates. The limitations of correlational studies are summed up in the aphorism that correlation does not imply causality. Similarly, a quasi-experimental study with non-randomized comparison groups of a program designed to expedite permanency among children in foster care may not adequately control for pre-existing differences between the groups that may be responsible for observed differences in outcomes; for instance, if children in the comparison group are older or have more developmental needs on average than children in the intervention group, it may prove more difficult to find permanent homes for them. In short, RCTs are the easiest and most effective way to assess what really works in improving outcomes for children and families. Building and sharing convincing evidence of effective programs and practices helps child welfare professionals avoid unhelpful or potentially harmful interventions and serves as a foundation for ethical child welfare practice (Gambrill, 1999).

RCTs in Child Welfare Settings – Challenges and Special Considerations

Although the benefits of RCTs have been clearly established, their implementation in child welfare practice settings poses a number of challenges that may require special consideration and accommodation. The tightly controlled clinical settings associated with RCTs are often difficult to achieve in real-world child welfare organizations, which can be hectic and unpredictable in the face of sudden and unanticipated crises, political considerations, and other factors that can undermine the internal validity of an experimental study. While not unique to the field of child welfare, some of the most common challenges to implementing RCTs in child welfare practice settings, and suggestions for addressing them, are highlighted below.

Organizational resistance. Resistance to RCTs among front-line staff, as well as among managers and administrators, is widespread in some child welfare organizations due to

misunderstandings regarding the ethics of random assignment. A common concern is that random assignment unfairly "denies" critical services to children and families. To address concerns about the ethics of random assignment, it can be helpful to remind child welfare professionals that the study is being conducted because the efficacy of the intervention in question is unknown. Researchers can appeal to the shared interest among child welfare professionals in improving the lives of children and families by stressing that RCTs offer the best means for determining what really works to improve child and caregiver safety and well-being, and that it is possible that a supposedly promising intervention is in fact ineffective or even has unintended negative effects. It may also be helpful to draw parallels between an RCT and a lottery. Most people would agree that a lottery is by definition fair because each participant has an equal chance of drawing a winning ticket; similarly, with random assignment each eligible person has an equal chance of "winning" as anyone else.

Despite certain misunderstandings regarding the ethics of random assignment, valid ethical concerns may arise in certain situations. For example, a project evaluation involving an RCT may not be feasible in a small or isolated community in which many potential service recipients know one another and will inevitably find out who did or did not receive the new service, thereby generating resentment and ill will toward the project. Legitimate objections to random assignment may also come up when the effectiveness of a given intervention has already been established through earlier studies, or when the intervention is of clear and immediate value to potential participants. In these circumstances, one of the rigorous alternatives to an RCT discussed later in this brief may be considered.

Interference with RCT protocols. Even when child welfare organizations and personnel accept RCTs in principle, they may undermine their execution in actual practice out of well-meaning motives to help their clients. For example, program staff may attempt to circumvent random assignment procedures in order to place clients who are perceived to be in greater need into the treatment group. If the proportion of participants with more severe needs assigned to the treatment group exceeds what would be expected through pure randomization, an otherwise effective intervention can appear to be ineffective because it must serve a larger number of participants with more serious problems than would be found in the target population as a whole.

Legal and policy mandates. Legal and policy mandates can also create impediments to the smooth implementation of RCTs. Examples of these mandates include state laws that require the placement of sibling groups together whenever possible, court orders to provide services to children and families (when these services are what are being tested as part of an RCT), and contractual obligations between child welfare agencies and service providers to serve a certain number of clients. All of these situations can violate the principle of true random assignment that defines an RCT. Waitlist and cluster randomization designs, which are described in more detail below, offer possible alternatives when legal, policy, judicial, or contractual realities make pure random assignment unfeasible.

Client attrition. Child welfare populations can be characterized by high levels of geographic mobility and instability; a sudden move or family crisis can make it more difficult for children or caregivers to participate fully in an intervention, which both undermines the intervention's potential efficacy and interrupts the collection of critical data on the intervention's outcomes. The junctures at which participant attrition is most likely to occur include initial study recruitment and selection, during exposure to the intervention itself, and during the post-intervention phase of follow-up data collection. Throughout all of these phases it is

imperative to maintain frequent and ongoing contact with study participants and to minimize "down time" when participants are not actively engaged in either the intervention or data collection activities. Specific strategies for minimizing attrition are presented below in the discussion of recommended steps for implementing an RCT.

Small samples sizes. Related to the challenge of attrition is the broader issue of small sample sizes, which can be especially problematic in evaluating interventions implemented in areas with small target populations (e.g., rural communities) or those in which available resources limit the number of cases that can be served. Samples that are too small may not generate enough statistical "power" to determine whether observed differences between a treatment and control group are in fact statistically significant. To avoid issues with small samples, efforts must be made to minimize attrition using the recommendations discussed later in this brief and to include as much of the potentially eligible target population in a study as possible. In addition, random assignment ratios can be adjusted to increase the number of cases assigned to the treatment group (e.g., a two-to-one instead of a one-to-one ratio), as long as the control group sample can maintain adequate statistical power. Modified assignment ratios can also increase buy-in to RCTs among project stakeholders by allowing more eligible cases to receive the intervention under study.

Challenges with implementation fidelity. Implementation fidelity, also referred to as adherence, integrity, and quality of implementation, refers to the extent to which the delivery of an intervention adheres to the protocol or program model as intended by the developers of the intervention (Dane & Schneider, 1998; Domitrovich & Greenberg, 2000; Mowbray, Holter, Teague, & Bybee, 2003). However, variations in staff and participant motivation, skills, and characteristics may weaken adherence to all of the critical components of an intervention; this in turn can undermine an RCT (or any research design) because the intervention of interest is now something different from what the study was originally intended to investigate, or may differ little from traditional services or case practices available to the control group. "Dosage", i.e., the number and/or duration of discrete intervention components that a provider offers, is one of the most critical aspects of fidelity that can undermine a study; for example, fidelity to a home visiting program could be compromised if the home visitor did not complete the prescribed number of visits or if the visits were too short on average. Many factors can derail implementation fidelity, including:

- Staff Buy-In: Front-line staff, managers, and other key implementation stakeholders may not always accept the reasoning underlying the various components of an intervention or the evidence that demonstrates its effectiveness. Without buy-in, these stakeholders may deliver the intervention without enthusiasm or covertly change the content of the intervention to correspond more closely with their clinical experience or personal beliefs.
- Knowledge Retention: Front-line staff and other key implementers may not recall important details regarding an intervention due to inadequate initial or follow-up training, insufficient opportunities to implement the intervention in regular practice settings, or simple forgetfulness. When staff members forget crucial aspects of an intervention they improvise or revert to past practices, thereby reducing fidelity.
- Staff Turnover: Frequent turnover among staff members charged with implementing a new intervention can result in continual pressure to vet, hire, and train new personnel under considerable time constraints, which can in turn encourage an organization to modify, truncate, or skip important aspects of the process for training, monitoring, and coaching staff in implementing the intervention.

- Participant Resistance: Fidelity depends on targeted participants embracing an intervention as fully as the staff charged with implementing it. If staff members do not fully engage the target population and convince them of the intervention's benefits, prospective participants may resist active involvement and consequently forego the full "dose" or treatment effect. Alternatively, staff may begin to modify the intervention in response to client resistance, which also undermines fidelity.

Given these challenges, it is essential to monitor implementation fidelity to ensure the integrity of both the intervention itself and of the RCT; without this assessment, it may prove difficult to determine whether and how a new program or service, as conceived and developed, affects child and family outcomes. Suggestions for monitoring and maintaining fidelity to a given intervention are discussed in more detail later in this brief.

Threats to external validity. A treatment that is found to be effective in one child welfare setting or with one target population may or may not be valid in other settings or with other populations. Inferential statistics allow for the examination of the impact of a treatment beyond the variability associated with the individuals in the study, but provide no information about the impact of the treatment on other types of individuals who are not part of the study (i.e., generalizability). A good RCT design, therefore, ideally includes diverse groups of people and seeks to assess whether the intervention is comparably effective among all classes of eligible participants or in different practice settings (Shadish, Cook & Campbell, 2002). Alternatively, an initial efficacy trial conducted in one practice setting with a particular population can be followed up with additional studies involving new populations and settings that permit broader conclusions regarding the intervention's effectiveness.

Placebo effects. In medical research, RCTs often compare the effectiveness of a particular medicine to a compound that lacks active ingredients. Decades of research have confirmed that just being given a pill and informed that it will improve health or reduce disease symptoms contributes to improved (whether real or imagined) health outcomes. This phenomenon is called the placebo effect, and variations of it can occur in studies of child welfare interventions as well. Related phenomena include "Hawthorne" effects, in which control group subjects improve or otherwise modify their behavior simply in response to the fact that they are being studied, or "Pygmalion" effects, whereby subjects perform better in response to heightened expectations. To minimize placebo or other reflexive effects in the evaluation of a human service intervention, researchers in an ideal RCT require that participants in the control condition engage in activities that are comparable in duration and engagement potential as those provided to subjects in the experimental condition. This helps to ensure, for example, that a program designed to reduce adolescent depression by teaching participants to control downward spiraling thoughts works because of the actual skills taught to regulate emotions, rather than just the act of "participating in training" or "engaging with caring adults."

Alternatives to Pure RCTs

When a pure RCT design is not feasible due to one or more of the special considerations highlighted above, methodologically rigorous quasi-experimental alternatives may be considered. Three design options are described below.

Cluster randomization. Situations exist in which it is not possible to assign participants to conditions on a completely random basis. For example, if two teenage siblings are placed in the same foster home and both qualify for participation in an adolescent pregnancy prevention program, the results of the study would be weakened if one sibling happened to be randomly assigned to the treatment condition (eligible for the prevention program) while the other sibling is assigned to the control condition (services as usual). The unavoidable interactions between the siblings would likely influence the behavior of the child in the control condition through exposure to the knowledge and skills acquired by the sibling in the treatment condition, thus attenuating differences in observed outcomes between the treatment and control groups as a whole. This phenomenon is referred to as "design contamination" due to the blurred distinction between the treatment and control groups with respect to the interventions they are exposed to.

Cluster randomization can be used to avoid or mitigate the problem of design contamination. In this procedure, eligible participants are first randomly assigned to one or more groups or "clusters." The members of these random clusters are then moved as necessary to minimize the likelihood of contamination, for example, by placing family members or classmates together in the same cluster. While constituting a deviation from pure random assignment, membership in the clusters *as a whole* is still determined on a random basis. One downside of cluster randomization is that the clusters themselves, rather than just the individual members of the clusters, must be included as a variable in subsequent statistical analyses. However, the loss of statistical power that results from the use of clusters rather than individuals as the unit of analysis can be mitigated through the use of multi-level statistical techniques such as Hierarchical Linear Modeling (Raudenbush & Bryk, 2002).

Matched case designs. This design alternative seeks to approximate the high degree of comparability achieved across groups through random assignment by matching each case designated to receive an intervention on a case-by-case basis with a comparison case that looks *as much like* that intervention case as possible; matching occurs using a variety of previously selected demographic and case-related variables, such as gender, age, maltreatment type, and placement history. A variety of statistical methods can be used to carry out the matching process, including propensity score matching, in which cases are matched based on a composite "propensity score" to minimize differences across any one matching variable. Matched case designs can be a good option when insufficient resources are available to serve all potentially eligible children, or when legal or contractual obligations require that all eligible children in a given child welfare serve area be offered a given service. Children living in other jurisdictions that have characteristics similar to those of children and families in the target jurisdiction can be matched and compared with one another over time. A significant challenge in implementing matched case designs involves identifying and collecting detailed data on presenting problems, demographics, and other case characteristics that are available for both the intervention and matched comparison groups that also ensure that these two groups are in fact as much alike as possible.

Waitlist/overflow designs. Waitlist designs are useful when legitimate concerns regarding an RCT exist or when legal or contractual obligations require that all eligible cases be served, but the number of eligible cases exceeds an organization's capacity to provide the intervention of interest at a given time. The hallmark of a waitlist design is that all eligible cases receive the intervention eventually, with receipt staggered across a series of time stages. The wait-listing process can be implemented either through randomization or using a "first come, first served" rule. With the first method, an initial wave of eligible families is

randomly assigned to a treatment or to a control group, with the treatment group receiving the intervention under study and the control group receiving serves as usual or an unrelated placebo intervention (as with an RCT). Pre-intervention data are collected from both groups prior to randomization and then during and/or after the intervention is completed. After a specified time period (e.g., three months), the cases assigned to the wait list control group are offered the intervention and any resulting changes in their outcomes are observed. A second wave of eligible cases is then randomized and the process repeats itself until the study's conclusion. Using a "first come, first served" protocol, when initial service capacity is reached the remaining eligible participants are placed on a waitlist and are observed along with participants that are currently receiving the intervention. As service slots open up, participants on the waitlist are assigned to the intervention group and retested to assess changes in the outcomes of interest.

Waitlist designs can be challenging to implement. First, they generally work best with very discrete and time-limited interventions, such that the observation of outcomes in the wait-listed group can be completed within a reasonable time frame. In addition, the risk of design contamination is increased if wait-listed clients seek a similar treatment while waiting or if the wait-listing process is in some way manipulated. For example, caseworkers who want to access an intervention on behalf of their clients may delay making a referral when they know there are no program vacancies; conversely, a supervisor may delay the referral of particularly problematic cases when openings do exist. These challenges highlight the importance of educating caseworkers, supervisors, and other program "gatekeepers" about maintaining the integrity of referral and wait-listing procedures.

Recommended Steps for Implementing an RCT

Despite the challenges and special considerations discussed above, child welfare organizations can conduct RCTs (as well as designs that approximate random assignment) successfully with careful planning and oversight of the random assignment and data collection process. In developing a protocol for implementing an RCT, organizations should in all cases adhere to several guiding principles:

- Ensure that random assignment occurs immediately after a case is determined to be eligible for the intervention under investigation; this minimizes the chances of skewed results if, for example, a person is exposed to other experiences or services that might affect the outcomes of interest before they are offered the new intervention.

- Have a member of the research team, rather than program staff, implement and monitor random assignment; this will maximize the integrity of the process by minimizing the chances of interference with the established random assignment protocol and increase the likelihood that any problems are identified and rectified in a timely manner.

- Follow the "once assigned, always assigned" rule: To avoid design contamination, once a case is assigned to the experimental or control group it must maintain that status throughout the duration of the study.

Within the framework of these general guidelines, the following steps are recommended for the successful implementation of an RCT:

1. **Educate stakeholders and obtain agreement.** A plan to implement an RCT must be explicitly articulated and agreed to by key organizational stakeholders before the research study or evaluation begins; it is unlikely that an RCT will succeed without the buy-in (if not outright enthusiasm) of front-line staff and key supervisory personnel. Gaining organizational consent to implement an RCT requires the research team to actively engage staff at all levels of the organization in order to educate them about the benefits of this design and determine exactly how the random assignment process will unfold. Engaging and obtaining the support of top leadership is especially important for buffering the RCT from potential interference or "sandbagging" by personnel at all levels of the organization. A contract or work plan for the evaluation should include an explicit statement that an RCT will be conducted and should carefully delineate both the random assignment process and the specific roles that members of the research team and organizational staff will play in implementing it.

2. **Establish inclusion/exclusion criteria.** Before active study recruitment can begin, discrete criteria must be established to determine who is or is not eligible to participate in the intervention; these standards are essential for ensuring that the intervention is only offered to persons or families for whom it is appropriate and most likely to be beneficial. Eligibility may be determined using both *inclusion* and *exclusion* criteria. Inclusion criteria include personal and case characteristics (e.g., race, gender, age, income level, health status, maltreatment risk level) that are especially relevant to the program or service of interest. For example, in an RCT examining the effectiveness of a sex education program on reducing first-time pregnancy among children in foster care, inclusion criteria could include child age (e.g., between the ages of 13 and 18), gender (female), and placement status (e.g., placed in a relative or non-relative foster home). Exclusion criteria include variables or traits that make the intervention inappropriate for a particular individual or family, or that the intervention may not be designed to address effectively. In the example of the teenage pregnancy prevention program described above, exclusion criteria could include a prior pregnancy, inappropriate age (e.g., 12 or younger and 19 or older), male gender, and residence in a group home or other congregate care facility.

   From a research standpoint, exclusion criteria can also simplify the study by reducing the number of variables that must be tracked and analyzed. Once inclusion and exclusion criteria have been identified, they should be operationalized by developing a screening form, such as a checklist, that can be used by the person(s) responsible for determining potential participants' eligibility. Eligibility screening may be conducted by front-line personnel in the organization that is implementing the intervention or by a member of the research team. If eligibility screening will occur in person, it is often preferable to have case managers or other front-line staff conduct it in the interests of maximizing prospective participants' trust and interest in the intervention.

3. **Develop a random assignment protocol and procedures.** Once eligibility criteria have been established, the actual procedures for carrying out random assignment must be determined. The specific method will depend in part on the complexity of the study and the organization's resources. In theory, random assignment can be done with something as simple as a flip of a coin, drawing names written on slips of paper from a container, or rolling a die. However, because these methods are often too easily compromised and are difficult to document for accuracy, it is generally preferable to automate the process by using a random number table or computer software that

automatically assigns participants to a study condition on a random basis[5]; these methods will maximize the integrity of the random assignment process while minimizing random assignment violations. This is also a good point at which to determine an appropriate assignment ratio; if the sample size is large enough, a ratio other than 1:1 (e.g., 2:1, 3:1) may be considered in the interests of increasing enrollment into the program under study.

In addition to establishing the specific procedures for executing random assignment, this step also involves the institution of procedures for maintaining the confidentiality of participants' personal information (for example, ensuring that identifying information is stripped from data files before they are shared with research staff responsible for conducting data analysis), and developing criteria for defining and documenting violations of the random assignment process. Violations of random assignment may be acceptable in rare and limited circumstances, for example, an extremely high-risk case for which no other suitable service options are available. The study team may wish to establish a maximum number of cases (e.g., 10) as exceptions that can be assigned to the treatment condition; this can enhance credibility of the RCT with service staff without compromising the overall research design. These excepted cases would not be included in subsequent data analysis.

4. **Recruit participants.** An RCT will only be as effective as the methods used to recruit study participants; without sufficient enrollment there will be little data to determine the intervention's effectiveness. Specific recruitment methods depend in part on the nature of the intervention under investigation and the characteristics and accessibility of the intervention's target population. One approach is to advertise about the study in the communities in which the target population lives, for example, through newspaper advertisements, flyers, and Web-based social media. This "direct marketing" method is commonly used with discrete interventions that potential participants will be highly motivated to engage in (e.g., a smoking cessation program). An approach that is more common in child welfare settings involves working with front-line staff and supervisors in the organization in which the intervention will be implemented to develop a protocol for explaining the research study to eligible persons and encouraging their enrollment. Study recruitment can often be integrated into a front-line worker's routine outreach and case management activities.

5. **Obtain informed consent.** As part of the recruitment process, researchers must gain the informed consent of potentially eligible individuals to participate in research regarding an intervention, regardless of their eventual assignment condition (treatment or control group). Consent should be obtained in writing using a consent form and protocols that have been reviewed and approved by a third-party institutional review board (IRB). A critical issue with respect to RCTs involves the timing of random assignment and the request for informed consent: Should the study be explained to eligible subjects and their consent solicited before random assignment occurs, or is it acceptable to randomly assign eligible subjects first and then seek their consent? Opinions regarding the appropriate approach vary among child welfare professionals and researchers. Some argue that it is important to obtain consent before random assignment in the interests of full transparency, whereas others

[5]Free software for generating random numbers or randomly assigning cases is available online at sites such as Stat Trek (www.stattrek.com) and GraphPad (www.graphpad.com).

contend that post-assignment consent is acceptable since subjects are giving their permission to participate in research regardless of their assignment status. In addition, post-assignment consent reduces the number of eligible subjects who may withhold consent because of the risk of being assigned to the control group. Numerous factors may influence the approach selected, including the vulnerability of the target population and the risks and benefits associated with the intervention and associated research activities. In either case, a research team should present its reasoning for a given approach in its IRB application and obtain the IRB's guidance regarding the most appropriate alternative.

6. **Develop participant retention strategies**. After eligible participants have been assigned, it is imperative to minimize attrition from all study groups. Whether the study involves an RCT or a quasi-experimental design, several techniques are available to keep participants engaged throughout the duration of the research process:

   a) Collect as much contact information as possible from potential participants before random assignment occurs, including physical addresses, land line and mobile telephone numbers, e-mail addresses, and social media accounts (e.g., Facebook and Twitter). In addition, collect similar contact information of the participants' closest family members to provide alternative means of finding and reaching out to participants as the study progresses. Assure participants that all of their contact information will remain strictly confidential.
   b) Continually remind participants of their importance in advancing knowledge regarding effective interventions. Their active engagement in the study will help child welfare agencies and practitioners understand the real needs of people like them and contribute to the development of more effective services.
   c) Maintain ongoing contact with participants during all phases of the intervention and research process. Notes sent via regular mail, text messages, e-mail, Facebook, or Twitter can be used to remind participants of upcoming service appointments and data collection activities (e.g., surveys, interviews, focus groups). In selecting appropriate contact methods it is important to consider factors such as confidentiality (e.g., are online media sufficiently secure to guarantee the anonymity of study participants and the confidentiality of data?) and accessibility (e.g., will all participating families have access to a computer?). Maintain a positive tone with these reminders to ensure participants understand that their ongoing involvement is appreciated. Thoughtful communications that are not directly related to the study can also foster engagement and motivation to participate; for example, when gathering demographic data, note participants' birth dates and send them traditional paper or electronic birthday cards.
   d) Creating a catchy project name and logo can create a sense of identity and solidarity among study participants regardless of whether they are assigned to a treatment or control condition. If funds allow, order and distribute t-shirts, pens, pencils, or similar promotional items with the project's logo to maintain interest in and a sense of ownership among participants. A project Facebook page is another low-cost and accessible strategy for keeping participants engaged and informed about the intervention itself as well as about the progress of the research study.
   e) If funds allow, provide participants tangible incentives for their involvement in the study; examples include small cash payments, gift cards, or a chance to win a drawing for a prize. Incentives must be substantial enough to be meaningful to participants but not so large that they become "coercive," i.e., a subject feels like

she must participate because of the generous gift even if she would actually prefer not to. The appropriate type and size of the incentive depends on the characteristics and circumstances of the target population of interest.

f) Make it easier for participants to engage in research activities by addressing potential logistical barriers. For example, offer participants transportation, child care, and meals to facilitate their participation in interviews and other structured data collection activities, and adjust the timing of data collection to their work and school schedules. Going to participants' homes to collect data rather than having them travel to an office or other outside location is another way to increase the convenience of the research process. Caseworkers or other service providers whom the participants know and trust can be charged with reminding them to keep their data collection appointments, as long as confidentiality is not compromised. These steps not only remove practical impediments to research subjects' involvement but act as additional incentives by demonstrating that their time is valuable and that their participation is meaningful.

g) Conduct focus groups before, during, and after the study to collect feedback on the intervention itself as well as regarding the research process. Early focus groups can ensure that intervention and research activities are sensitive to the logistical realities and time commitments of participants by addressing questions regarding the duration and spacing of activities. Focus groups conducted in the middle of the study can offer valuable information regarding aspects of the intervention and research process that are working well, as well as those that may need refinements. Focus groups conducted after the completion of the study can provide insights into participants' perceptions of the intervention and research study as whole, as well as their recommendations for improving similar services and research activities in the future.

7. **Measure and ensure implementation fidelity.** As noted earlier, the maintenance of fidelity to an intervention's core components is essential to the integrity of the intervention itself and to the RCT. Consequently, the development of procedures for tracking and ensuring fidelity to the intervention represents an integral part of the entire research process. Fidelity measurement begins by establishing fidelity criteria for the intervention (i.e., what specific activities must occur at what duration and frequency) followed by the development of data collection tools for assessing fidelity (e.g., structured observations, case record reviews). Frequent and intensive staff training is critical to ensuring fidelity to the intervention; the format and content of this training depends largely on the intervention itself. With some programs, the developers require that certified trainers conduct the training (e.g., Trauma-Focused Cognitive Behavioral Therapy) to ensure a high degree of consistency in the delivery of the educational content about the intervention. In other cases, the program developer may offer a manualized training protocol that an agency's training staff can deliver directly to program staff. In either case, all personnel involved in the intervention should be trained, the effectiveness of the initial training should be evaluated, and the training must be reinforced to ensure that it is applied consistently throughout the course of the intervention. The training evaluation should include an objective assessment of trainees' mastery of the knowledge and skills contained in the training (Antle, Barbee, & van Zyl, 2008).[6] Afterwards, trainees' fidelity to the intervention

[6]For more information on this subject, see JBA's evaluation brief titled *Measuring Implementation Fidelity*, available online at http://www.jbassoc.com/reports/section.aspx?category=8.

should be measured during implementation by an evaluator or other objective observer according to the previously established fidelity criteria.

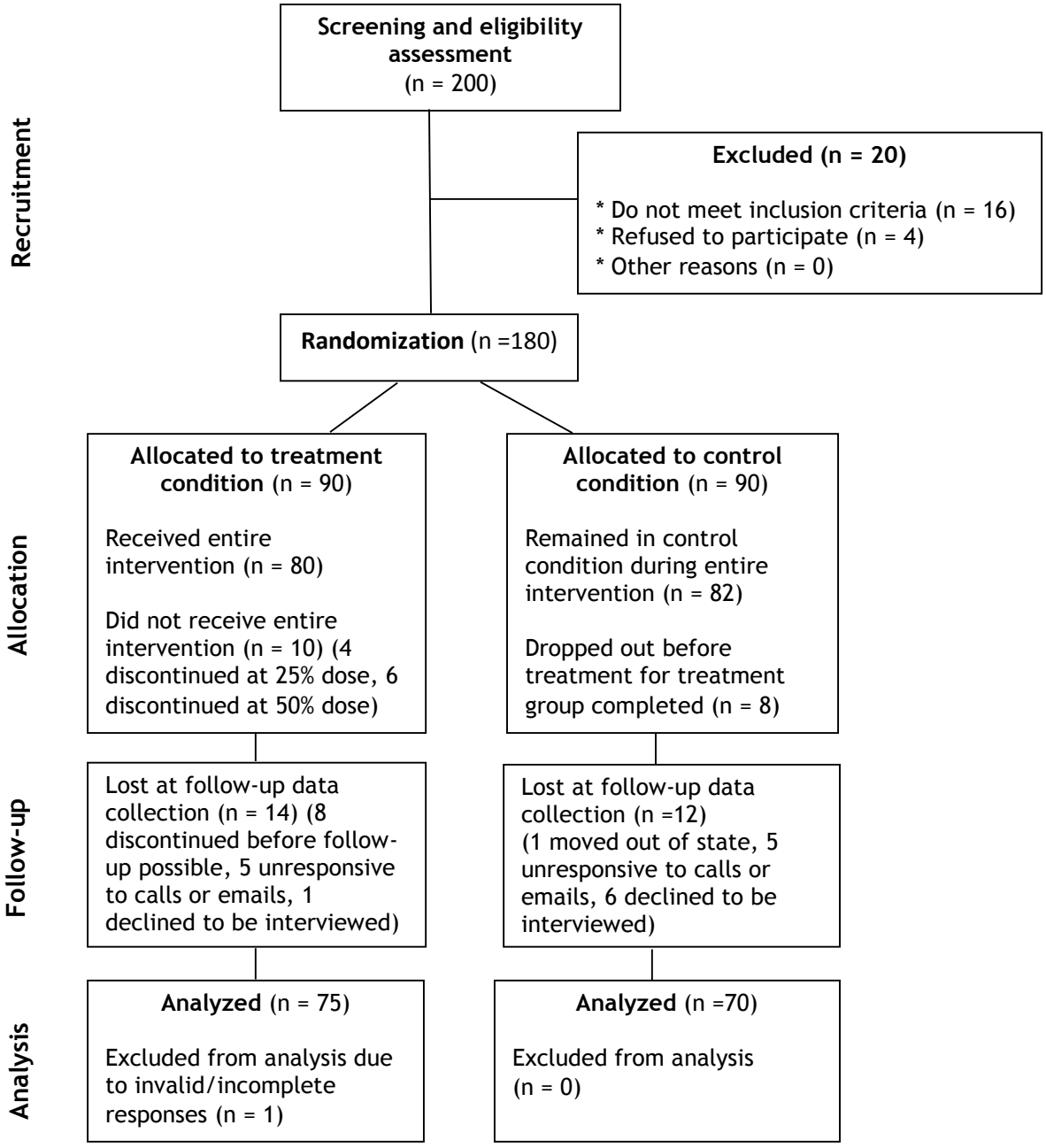
8. **Track the status of the control group.** The specific services and supports that control group subjects receive must be carefully monitored and documented, with special attention paid to the presence of design contamination if control group subjects are intentionally or inadvertently exposed to the intervention of interest or to substantively similar services. If the control condition is "treatment as usual," it is imperative that staff members who are trained in the intervention not share their new knowledge and skills with their untrained colleagues; therefore, regular follow-up with personnel assigned to work with cases in the control condition is an important component of an RCT protocol. If the control condition is an alternative service or activity that does not directly address the problem addressed by the primary intervention, personnel assigned to work with cases in this condition must be thoroughly trained in this alternative service and implement it with fidelity.

9. **Continually monitor the RCT process.** As the RCT progresses it is important to keep track of the status of each phase of the study (recruitment, randomization, initial and follow-up data collection, and analysis) to ensure that potential problems are identified and addressed in a timely manner. Schulz, Altman, and Moher (2010) have developed what they refer to as a Consolidated Standards of Reporting Trials (CONSORT) diagram to assist with the RCT monitoring process. An example of a CONSORT diagram appears in Exhibit 1 on the following page.

Conclusion

This brief has sought to summarize the key elements of RCTs, explain their importance as a research approach that advances the evidence base regarding effective child welfare programs, and provide suggestions for their successful implementation in the practice settings of CB discretionary grantees and child welfare organizations in general. The challenges facing state and local child welfare systems have historically been so profound that it has often proved difficult to put RCTs, or other systematic tests of efficacy, into practice as part of efforts to assess the impact of new interventions on vulnerable children and families. However, new research tools and resources have made methodologically rigorous evaluations of child welfare programs and practices more feasible than ever before. The situation in the child welfare field today is not unlike the medical profession at the start of the 20th century. Advancements were continually made, but many medicines and procedures were either ineffective or did more harm than good. Progress accelerated when physicians, nurses, and hospital administrators were persuaded to use rigorous scientific protocols, including the RCT, to evaluate the efficacy, safety, and cost-effectiveness of alternate forms of treatment. As better practices were developed, more money became available to fund additional innovations and further research, which contributed in turn to continually evolving "standards of care" and "best practices." Similarly, it is imperative for the child welfare field to eschew ideologically based practice in favor of evidence-based practice (Gambrill, 1999). The methodological foundation for evidence-based practice is the RCT. Despite their inherent challenges, RCTs are an essential element to future progress in the child welfare field and can be implemented successfully with careful planning and education for child welfare workers, administrators, and policymakers.

Exhibit 1: Sample CONSORT Diagram



References

Antle, B. F., Barbee, A. P., & van Zyl, M. A. (2008). A comprehensive model for child welfare training evaluation. *Children and Youth Services Review, 9,* 1063-1080.

California Evidence-Based Clearinghouse for Child Welfare (CEBC) (2013). Available at: http://www.cebc4cw.org/.

Chaffin, M., Hecht, D., Bard, D., Silovsky, J. F., & Beasley, W. H. (2012). A statewide trial of the SafeCare home-based services model with parents in child protective services. *Pediatrics, 129,* 509-515.

Cohen, J. A., Berliner, L., & Mannarino, A. (2010). Trauma focused CBT for children with co-occurring trauma and behavior problems. *Child Abuse & Neglect, 34,* 215-224.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18*, 23-45.

Domitrovich, C. E., & Greenberg, M. T. (2000). The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation, 11*, 193-221.

Fraser, M. W., Walton, E., Lewis, R. E., Pecora, P. J., & Walton, W. K. (1996). An experiment in family reunification: Correlates of outcomes at one-year follow-up. *Children and Youth Services Review, 18,* 335-361.

Gambrill, E. D. (1999). Evidence-based practice, an alternative to authority-based practice. *Families in Society, 80,* 341-350.

Gambrill, E. D. (2013). *Social work practice: A critical thinker's guide*. New York: Oxford Press.

Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation, 24*, 315-340.

Raudenbush, S. W. & Bryk, S. A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Press.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin Company.

Schulz, K. F., Altman D. G., & Moher, D. (2010). CONSORT group CONSORT 2010 statement: Updated guidelines for reporting parallel group randomized trials. *British Medical Journal, 340*, 332.

Zients, J. D. (2012). *Memorandum for the heads of executive departments and agencies*. Executive Office of the President, Office of Management and Budget. Available at: http://www.whitehouse.gov/sites/default/files/omb/memoranda/2012/m-12-14_1.pdf.