

Brief Research-to-Results

Child
TRENDS[®]

...information for program providers who
seek to evaluate their programs.

January 2008

Publication #2008-03

RANDOM ASSIGNMENT EVALUATION STUDIES: A GUIDE FOR OUT-OF-SCHOOL TIME PROGRAM PRACTITIONERS

Part 5 in a Series on Practical Evaluation Methods

Kristin Anderson Moore, Ph.D., and Allison Metz, Ph.D.

OVERVIEW

Program providers work extremely hard to offer services that they believe will enhance the lives and future prospects of children and youth. But how can they be sure that their program is having the intended impacts? How can funders and policy makers be sure that their investments in programs are paying off in substantially better outcomes for children and youth? How can parents and children know that they are choosing to be involved in a worthwhile activity?

Random assignment studies are only one kind of research, and they are not appropriate for every situation. But when the question is “What works?” and rigorous evidence about cause and effect is needed, a random assignment experimental evaluation can be very useful. In addition, a random assignment evaluation can provide credible evidence about the size of impacts found for the program participants.

WHAT IS A RANDOM ASSIGNMENT STUDY?

A **random assignment study** uses a lottery system to randomly assign participants to the **experimental group** (often called the **treatment group**) or to the **control group**. Participants in the treatment group can participate in the program or activities being evaluated. Participants in the control group cannot. Because a lottery system is used, participants have an equal chance of being in the treatment group or in the control group.

The key to a random assignment experimental study is that members of both the experimental group and the control group are, as groups, the same or very similar. It’s that simple. Researchers use many other non-experimental methods to try to make people as similar as possible, but even very complex non-experimental studies cannot compete with a simple, well-implemented random assignment study for determining whether a program affects outcomes.

WHY ARE RANDOM ASSIGNMENT STUDIES IMPORTANT?

Randomly assigning participants to the treatment or control groups has two important advantages over all other evaluation designs:

- **Randomization is the best way to create equivalent groups.** That is, participants with varied characteristics (e.g., age, sex, race, and motivation) will be equally (or very close to equally) represented in both the treatment and control groups. Groups and organizations such as classrooms can be randomly assigned, as well as individuals.

- *A random assignment experimental study is the only way to be sure about cause and effect.* For example, if doctors want to know whether a medication *causes* patients to be cured, they will do a random assignment study in which the experimental group gets the medication and the control group does not. Patients are chosen randomly, through a lottery, to decide who will be in the experimental group. Knowing that the two groups are equivalent on all characteristics except that one group did not receive treatment enables doctors to compare outcomes for the treatment and control groups. The doctors can conclude whether or not the medication resulted in a cure. This important advantage has led researchers to describe random assignment experiments as “the gold standard.”

In the out-of-school time field, the “treatment” isn’t a pill or a medical procedure, but the opportunity to be in a certain program. A lottery is held to randomly assign eligible children or youth to be in the experimental group or the control group. The experimental group has the opportunity to be in the program. For example, those in the experimental group may have an opportunity to have a mentor in Big Brothers Big Sisters, or to participate in Quantum Opportunities activities, or to be a part of a Children’s Aid Society program. In a well-implemented experiment, every child or youth has the same chance to be in the experimental and control groups, and the treatment and control groups will be equivalent.¹

Every program provider knows that the children and youth who are actively involved in their program have special characteristics, such as motivation, that distinguish them from the children who never show up, or the children who drop out. If researchers compare these more motivated children who choose to come to a program and want to be highly involved with a sample of children who are not as motivated, the motivated children are probably going to look better on many “outcome” measures. However, as much as we might like to, we cannot conclude that the program caused these fortunate children to have positive outcomes. The characteristics that led them to choose to attend the program (or to “self-select” into greater participation) may be the real reasons that they have good outcomes.

If it is well-implemented, a random assignment study will generally have just as many motivated children in the control group as are in the experimental group. Then, over time, the outcomes for children in the experimental group can be compared with outcomes for children in the control group. If the experimental group is doing better, it will be possible to conclude that it is the program that produced these benefits for the children in this group.

HOW IS RANDOM ASSIGNMENT DONE?

As in a lottery, people, groups, or organizations who apply for a program are randomly assigned to either be in the experimental group or to be in the control group, as noted. Sometimes there may be more than one treatment group; for example, some participants in one experimental group may be assigned to a one-year program, while a second experimental group gets assigned to a two-year program. Outcomes for both groups are usually compared with a control group that receives either no treatment or an innocuous and unrelated treatment. And there can be no exceptions. If somehow youth with particular characteristics (for example, very motivated or severely at-risk youth) are moved into or out of the experimental group, the integrity of the study is compromised and the results will be inconclusive.²

When random assignment is done at the level of a group or organization, such as a classroom, treatment is provided to all of the classrooms in the treatment group and none of the classrooms in the control group. This design is often chosen when it is difficult or impossible to randomly assign individuals, for example, when it is only feasible to offer a curriculum to an entire classroom. To be confident of the results, an independent evaluator can assist a program to design study procedures and identify the necessary sample size.

Sometimes a random assignment study is designed such that control group members get the treatment six months later, or a year later, in a delayed treatment condition. This may occur if there is strong opposition to random assignment for some reason. The downside of this strategy is that long-term impacts cannot be assessed. For example, if the delay is six months, then impacts can only be assessed for those six months.

Members of the control group may, of course, join a similar program elsewhere. If applicants who are assigned to the control group end up receiving services that are similar or identical to those the treatment group is receiving, it becomes difficult for the researchers to determine the potential impact of the treatment program. (Impacts will identify the effect of the program compared with a similar program, rather than the effects of the program compared with no program.) While this situation weakens the contrast between the control group and the treatment group, it cannot be prevented. Because some applicants assigned to the control group may seek out alternative services, it is useful for researchers to collect data on the types of services that members of the control group are receiving.

IS RANDOM ASSIGNMENT ALWAYS NECESSARY?

Random assignment experiments are often difficult, expensive and time-consuming; and they have a specific purpose—to provide strong causal evidence about whether the program produces impacts, and the magnitude of those impacts, in the population served. However, we must acknowledge what one researcher calls the “rhinoceros in the living room”:³ random assignment evaluations are not always the right kind of study for a particular program at a particular point in time.

A program that is just getting started will benefit more from doing a careful **implementation study** that develops a logic model and assesses whether services are being delivered as planned to the intended target populations, as well as identifies specific barriers or facilitators to implementation.⁴ **Monitoring outcomes** over time can also be very useful to programs that seek preliminary evidence about whether the expected outcomes are occurring and appear to be improving over time.⁵

In addition, if a program is replicating a proven approach, then it is not usually necessary to conduct another experimental study. Rather, in this case, it is critical to confirm that the core components of the intervention are being replicated and implemented with fidelity. However, if the population, community, or social context is different, a new evaluation may be warranted. Also, if replication is unlikely, if rapid information is needed, if funding is not available for a rigorous experimental evaluation, or if community opposition or ethical considerations preclude random assignment, another approach to evaluation may be warranted.

DIFFERENCES BETWEEN BIOMEDICAL AND SOCIAL EXPERIMENTS

When considering random assignment studies, there are key differences to remember between **biomedical experiments** and **social experiments**:

Many biomedical experiments are “**double-blind**” studies. This means that neither the person participating in the study nor the researchers know whether the person is in the treatment group or the control group. Person A gets a pill and Person B gets a pill that appears identical. No one knows which person got the placebo or the “phony pill” with no active ingredients (except for the record-keeper, who keeps all of the data confidential). This approach is impossible for an out-of-school time experimental study, in which everyone knows whether or not a child is in the treatment group because he or she has been assigned to have a tutor or a mentor or to be in an activity or a class. Nevertheless, the fact that eligible children are randomly assigned to be in the treatment group or the control group is extremely valuable. This value stems from the fact that random assignment addresses the problem of bias due to self-selection into a program, which is extremely important for studies of long-term involvement in an out-of-school time program.

Unlike many biomedical experiments, in which everyone in the experimental group actually receives the pill or the vaccine, in a social experiment, everyone is offered the program but not everyone will actually participate. Some people in the experimental group will come by for the first session and drop out; others will come occasionally; and others will come regularly. This is part of the treatment. Some people in the control group will find a different program. If they are not chosen for the treatment program, some will go elsewhere and join a different activity. Children in the experimental group, however, remain in the experimental group; and children in the control group remain in the control group. If crossover occurs, the results may not be valid. Since it is often the case that some members of the control group receive some kind of somewhat similar service, findings of program impact are weaker than they would be if the control group did not receive any services.

Program providers understandably prefer to be judged on the performance of those children who actually participate. However, to avoid self-selection bias, it is critical to assess all of the children or youth who are assigned to the experimental group and compare them to all of the children in the control group. The reason avoiding self-selection bias is so important is that, when evaluating the impacts of a program, you need to assess the entire program, and this includes determining how effective the program is in engaging and retaining participants. It may seem easier to just assess the impact of a pill or vaccine; but that is not what out-of-school time programs do.

ADDRESSING RESERVATIONS THAT PROGRAM PROVIDERS MAY HAVE

A program provider who works hard every day and believes in the value of his or her program is naturally going to have reservations about assigning children or youth to a control group that is not in the program. Even when random assignment is deemed the most appropriate evaluation approach, many program managers and practitioners feel it is unfair to deny potentially beneficial services to children or their families. This is completely understandable. However, many programs have more eligible children or families than they can actually serve. When this is the case, it presents a perfect opportunity to conduct a random assignment study because a lottery represents a fair and reasonable way to allocate program slots. At the same time, researchers need to recognize the concerns of practitioners and work together to assure that procedures are as fair as possible. For example, it may be possible for the experimental group to be larger than the control group. Moreover, as noted above, many times, participants assigned to the control group can be provided with treatment group services within six months or a year or so.

Also, from a larger viewpoint, it is very important to know whether the programs that we believe work actually do work in a real-world situation:

- Sometimes *researchers find that programs don't work* (for example, New Chance⁶) or that programs have unexpected negative impacts (for example, programs that use a Scared Straight approach⁷).
- In other instances, *researchers find that programs have unexpected positive impacts* (for example, early childhood programs that prevent teen parenthood more than a decade later, such as the Abecedarian program⁸ and the Perry preschool program⁹).
- Also, *experimental studies often find that programs work for one subgroup but not for another* (for example, the Children's Aid Society Carrera Program had impacts on sexual activity for girls but not for boys¹⁰).

All of these findings have been important. They have provided critical insights not only for the programs themselves, but also for funders and for the field as a whole. Most important, these studies are contributing to improved quality for out-of-school time programs. In the long run, this will enhance outcomes for children and youth.

NEXT STEPS: ADDITIONAL RESOURCES FOR YOUR PROGRAM

Child Trends – Lifecourse Interventions to Nurture Kids Successfully (LINKS). LINKS is a continually updated compendium of experimentally evaluated social programs for children.

Available online at: <http://www.childtrends.org/Lifecourse>

Coalition for Evidence-Based Policy – Evidence-Based Policy Help Desk. This Web site is full of resources for programs needing guidance on how to implement experimental evaluation designs.

Available online at: <http://www.evidencebasedpolicy.org/default.asp?sURL=ombII>

Gueron, J.M. (2000). The politics of random assignment: Implementing studies and impacting policy. This paper gives practical advice on how to implement randomized, experimental designs within a social policy context.

Available online at:

http://www.evidencebasedpolicy.org/docs/Gueron-Politics_of_Random_Assignment.pdf

MDRC – Methodology Initiative. MDRC is developing, testing, and plans to disseminate better strategies for evaluating social programs, including random assignment evaluation designs.

Available online at: http://www.mdrc.org/rsch_method.htm

University of Michigan – Building Capacity to Evaluate Group-Level Interventions. The University of Michigan has made available software and free consultation on conducting randomization at a group level. This service is funded by the William T. Grant Foundation, whose goal is to strengthen youth-related programs and their capacity to rigorously assess their program impacts.

Available online at: <http://sitemaker.umich.edu/group-based/home>

U.S. Department of Education – Random Assignment in Program Evaluation and Intervention Research: Questions and Answers. The Department of Education has produced a fact sheet that gives realistic guidance on the implementation of random assignment.

Available online at: <http://www.ed.gov/rschstat/eval/resources/randomqa.html>

REFERENCES

Julnes, G., & Rog, D.J. (Eds.). (2007). Informing federal policies on evaluation methodology: Building the evidence base for method choice in government sponsored evaluation. *New Directions for Evaluation*, 113.

Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: Sage Publications.

ACKNOWLEDGEMENTS

The authors wish to thank Martha Zaslów, Ph.D., and Brian Wilcox, Ph.D., for their thoughtful reviews and helpful comments on this brief.

- ¹ The treatment and control groups may differ slightly, especially if sample sizes are small; thus, we say “equivalent” rather than “identical.” Researchers should always assess equivalence when data are analyzed.
- ² Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: A systemic approach* (7th ed.). Thousand Oaks, CA: Sage Publications
- ³ Hunter, D. E. K. (2006). Daniel and the rhinoceros. *Evaluation and Program Planning*, 29, p.181.
- ⁴ Bowie, L., & Bronte-Tinkew, J. (2007). Process evaluations. (*Research-to-Results* Brief). Washington, DC: Child Trends.
- ⁵ Bronte-Tinkew, J., Moore, K. A., & Shwalb, R. (2006). Measuring outcomes for children and youth in out-of-school time programs: Moving beyond measuring academics. (*Research-to-Results* Brief). Washington, DC: Child Trends
- ⁶ Quint, J.C., Bos, J.M., & Polit, D.F. (1997). *New Chance: Final report on a comprehensive program for young mothers in poverty and their children*. New York: Manpower Demonstration Research Corporation. (The LINKS write-up of the New Chance program is available at: <http://www.childtrends.org/Lifecourse/programs/NewChance.htm>)
- ⁷ Petrosino, A., Turpin-Petrosino, C., & Buehler, J. (2003). Scared Straight and other juvenile awareness programs for preventing juvenile delinquency: A systematic review of the randomized experimental evidence. *Annals of the American Academy of Political & Social Science*, 589, 41-62.
- ⁸ Campbell, F.A., Ramey, C.T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian project. *Applied Developmental Science*, 6(1), 42-57. (The LINKS write-up of the Abecedarian program is available at: <http://www.childtrends.org/Lifecourse/programs/CarolinaAbecedarianProgram.htm>)
- ⁹ Schweinhart, L. J., Barnes, H. V., & Weikart, D. P. (1993). *Significant benefits: The High/Scope Perry preschool study through age 27. Monograph of the High/Scope Educational Research Foundation, 10*. Ypsilanti, MI: High/Scope Press. (The LINKS write-up of the Perry preschool program is available at: <http://www.childtrends.org/Lifecourse/programs/HighScope-PerryPreschoolProgram.htm>)
- ¹⁰ Philliber, S., Kaye, J., & Herrling, S. (May, 2001). *The national evaluation of the Children’s Aid Society Carrera-Model Program to prevent teen pregnancy*. New York: Philliber Research Associates. (The LINKS write-up of the CAS-Carrera program is available at: <http://www.childtrends.org/Lifecourse/programs/CAS-Carrera.htm>)