OPRE Report #2015-98 • October 2015

Learning What Works: **A Guide to Opportunistic Experiments for Human Services Agencies**

Learning What Works: A Guide to Opportunistic Experiments for Human Services Agencies

OPRE Report #2015-98 • October 2015

Jonathan McCay • Alicia Meckstroth • Lauren Akers • Alexandra Resch • Michelle Derr • Jillian Berk

Submitted to:
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services
Project Officer: Emily Schmitt

Contract/Task Number: HHSP23320095642WC / HHSP23337043T

Submitted by:
Mathematica Policy Research
1100 1st Street, N.E. Suite 1200
Washington, DC 20002-4221
Project Director: Michelle Derr
Reference Number: 40153.233

This report is in the public domain. Permission to reproduce is not necessary. Suggested citation: McCay, Jonathan, Alicia Meckstroth, Lauren Akers, Alexandra Resch, Michelle Derr, and Jillian Berk. (2015). *Learning What Works: A Guide to Opportunistic Experiments for Human Services Agencies*, OPRE Report #2015-98, Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Disclaimer: The views expressed in this publication do not necessarily reflect the views or policies of the Office of Planning, Research and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services.

OPRE

MATHEMATICA Policy Research

Contents

A Guide to Opportunistic Experiments for Human Services Agencies ... 4

What is a randomized controlled trial (RCT)? 5

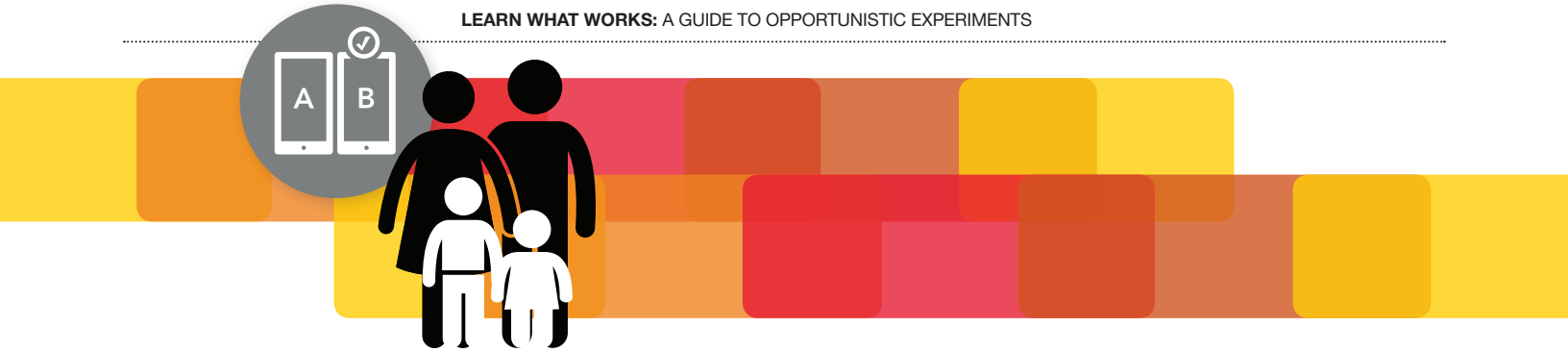
Why should you use RCTs? 6

What are some good opportunities to conduct an RCT? 7

How do you conduct an RCT? 8

What are some common questions and concerns about RCTs? 11

Conclusion 12



Learning What Works

A GUIDE TO OPPORTUNISTIC EXPERIMENTS FOR HUMAN SERVICES AGENCIES

If you are a program director or administrator at a human services agency that serves low-income families, you must constantly decide how to use and invest your agency's resources most effectively. You want to use those programs and service delivery strategies that have been shown effective at moving clients and their families toward greater self-sufficiency, stability, and well-being. Unfortunately, many programs and strategies have little or no proof of effectiveness, and the available evidence may come from studies without a strong research design or from a context different from your own. Faced with these challenges, it may seem sensible to forge ahead with the approaches and strategies you know best, even though they may lack rigorous evidence of effectiveness. The risk in this approach is that you could spend years and resources using a strategy that isn't effective at improving outcomes for your clients and their families.

To learn about the effectiveness of a program or strategy within your agency, there is a reliable and accessible tool—*known as a randomized controlled trial (RCT)*—that you can use to test the effectiveness of a change in a program or strategy before rolling it out to all your clients. If the program or strategy is shown to be ineffective, you can make adjustments or try a different approach. If it is effective, you can continue or expand it with confidence. There are numerous options along a continuum of research approaches—both quantitative and qualitative—for learning about your program and making improvements. However, an RCT—because it uses random assignment—is the most rigorous and reliable approach both for learning what works to improve outcomes and developing strong evidence of program effectiveness.

This guide is intended to help your agency introduce an accessible, reliable, and efficient approach for conducting low-burden RCTs—sometimes known as *opportunistic experiments*—by identifying common situations that are well-suited for an experiment and then walking you through the research process. The guide explains the benefits of using random assignment to conduct an experiment, answers common questions and concerns, and provides practical, step-by-step guidance on how to conduct an RCT in a way that both minimizes cost and disruption and provides an opportunity to inform ongoing program improvements in a timely way.

WHAT IS A RANDOMIZED CONTROLLED TRIAL (RCT)?

An RCT is a type of study that demonstrates whether an *intervention*—a program, strategy, policy, or process—causes a certain outcome. All RCTs start with two basic components: (1) a *treatment group* (those who will receive the intervention), and (2) a *control group* (those who will not receive the intervention). A control group may be assigned to continue with "business as usual" (which could be an existing intervention or strategy), to receive an alternate intervention, or to receive no intervention at all. The study is "randomized" because people are randomly assigned to each group so that the groups are as similar as possible at the beginning of the study and any differences between groups are due to chance. *Random assignment* is a critical step to ensure that you compare "apples to apples." (See box below.)

To see how this works in practice, consider how one agency took advantage of an opportunity to learn whether an intervention worked. An urban human services agency in a Minnesota county piloted an integrated case management and service delivery model that co-located mental health counseling, vocational rehabilitation, primary health care, and employment services for Temporary Assistance for Needy Families (TANF) clients with disabilities. The agency partnered with a research team funded by the Administration for Children and Families (ACF) within the U.S. Department of Health and Human Services to assess whether the integrated model increased services uptake, employment, earnings, and other outcomes for clients. Nearly 400 eligible TANF clients were randomly assigned to a treatment group (those receiving the pilot integrated model) or a control group (those receiving the existing case management and service model). Within 16 months of launching the pilot, the research team examined administrative data and found, for example, that within one year, treatment group participants had significantly higher earnings than control group participants.[1]

Like this agency, you may be able to take advantage of such an opportunity to conduct an RCT and use the findings to effectively target your allocation of scarce resources. After further exploring the benefits of using an RCT, this brief provides more detail on conducting an RCT, including how to: (1) identify opportunities to conduct an RCT while minimizing the time and resources required; (2) gauge the feasibility of conducting an RCT; and (3) understand the steps involved in conducting an RCT.

Key terms in RCTs

Intervention: The program, policy, or other change being evaluated.

Random assignment: The statistical process by which study participants—clients, staff, or local sites—are randomly assigned to either a treatment or control group. The only systematic difference between the groups is whether they receive the intervention. Microsoft Excel® includes a random-number generator that is often used for random assignment.

Treatment group: The participants randomly assigned to receive the intervention.

Control group: The participants randomly assigned to not receive the intervention (also sometimes referred to as a comparison group or the counterfactual).

[1] Farrell, M., et al. *The TANF/SSI Disability Transition Project: Innovative Strategies for Serving TANF Families with Disabilities*. OPRE Report 2013-51, Washington, D.C., U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation, December 2013.

WHY SHOULD YOU USE RCTS?

Random assignment, a defining aspect of RCTs, provides an "apples-to-apples" comparison of what happens with and without a given program or strategy.

Random assignment, a defining aspect of RCTs, provides an "apples-to-apples" comparison of what happens with and without a given program or strategy. In this section, we walk through this idea step by step.

Why do you need to compare two groups to find out what works?

We need to know whether outcomes for those who received the intervention differ from outcomes for those who did not. Consider the example of the RCT of the integrated case management and service delivery model. It compared two groups of TANF clients with disabilities: (1) a treatment group, which received the integrated model, and (2) a control group, which continued with the "business as usual" approach. However, suppose the county agency had simply piloted the intervention with all clients with disabilities in the county, but left no clients for comparison? If clients' earnings increased, for example, would it be reasonable to conclude that the intervention caused the improvement?

Earnings outcomes improved for clients of the integrated case manage ment and service delivery model—does that mean the intervention worked?

Next, let us suppose that the agency looked at TANF clients with disabilities across the state and found that, on average, their earnings were also increasing, just as they were in the county piloting the intervention. Without a control group with which to compare the impacts of the intervention, the county agency could not confidently conclude that its integrated service model caused the improvements because something else might have occurred in the state or county that improved outcomes. For example, a service provider may have held training for staff that improved job search coaching and contributed to clients' improved outcomes. Or, the state TANF agency may have distributed a new handout to clients that more clearly explained their eligibility for and how to access various services. Thus, it is unclear whether the integrated case management and service delivery model was the reason for improved client outcomes.

How does random assignment allow you to compare apples to apples?

To know whether an intervention truly caused the desired effect, it isn't enough just to compare two groups; you must compare two similar groups. This is where random assignment is critical. Random assignment ensures an "apples-to-apples" comparison. In the earlier example, clients were randomly assigned to a treatment or a control group. Because the assignment process was random, any noticeable differences between the two groups arose simply by chance, except for the crucial, intended one: receiving the intervention. Since the two groups were not different except for exposure to the intervention, we can confidently conclude that differences in outcomes were caused by exposure to the intervention.

Without random assign ment, clients in one group may differ from the other in a way that skews outcomes.

Without random assignment, it's easy to end up comparing apples to oranges

What would have happened if the agency had *not* randomly assigned clients? The agency might have found itself comparing apples to oranges.

For example, suppose the agency decided to test the integrated case management and service delivery model by asking clients with disabilities to volunteer to participate.

Earnings outcomes for clients with disabilities were improving across the state, not just in the pilot county, so we cannot say that the pilot program caused the improved outcomes.

After participating, the clients with disabilities who volunteered achieved better outcomes than other clients with disabilities who did not participate. Did the intervention cause the improvement? If the agency attributed the improvement to the intervention, this conclusion could be incorrect because, without random assignment, the difference between the groups could be the result of some factor other than the intervention. For instance, it is possible that the clients who volunteered to participate recognized the potential value of the intervention and were more motivated to get services or find work than those who did not volunteer. Or, the volunteers may have had different characteristics than the non-volunteers, such as a greater level of education, on average. Without random assignment, there will always be questions about whether it was the intervention or other factors that led to differences in the groups' outcomes.

WHAT ARE SOME GOOD OPPORTUNITIES TO CONDUCT AN RCT?

You may encounter many situations in which conducting an RCT would be both advantageous and feasible. You will need to thoughtfully consider how an experiment can best be used to inform and improve program practices and service delivery. Here are some opportunities to look for:

- **Pilots of a new program or service delivery process.** An agency considering a new program or process might ask its local sites to volunteer to pilot it. Here's the opportunity: Test the intervention by randomly assigning the volunteer sites to a treatment group that pilots the program or process and a control group that carries on with "business as usual."

- **Limited resources to roll out a new program, strategy, or process or a preference to implement it in stages.** If an agency doesn't have the resources to roll out an intervention to all of its sites, or prefers to launch the intervention in stages, then here's the opportunity: Test the strategy by randomly selecting sites for the first and second stages of a staggered rollout. Sites chosen for a second phase of rollout can serve as a control group. The agency can conduct an RCT during the first stage of rollout by comparing outcomes of the sites who implement the strategy during the first stage (treatment group) to those that will implement it in the second stage (control group).

- **Communication efforts.** An agency might try to influence clients or potential clients through different approaches to communicating with them. For example, an agency might want to establish a better communication method to encourage clients' attendance at regularly scheduled appointments. Here's the opportunity: Test three approaches to sending meeting reminders to clients: (1) a text message, (2) a telephone call, or (3) an e-mail. By randomly choosing which method is used for a client, agencies can learn which strategy is most effective at improving clients' attendance.

- **Excess demand for a new program.** An agency might see excess demand for a new intervention, such as an intensive work-readiness training program, that has a limited number of slots. Here's the opportunity: Test the training program by using a lottery approach to enrollment—which is a form of random assignment—to assign slots to a portion of those who are interested in the program. This approach allows the agency to

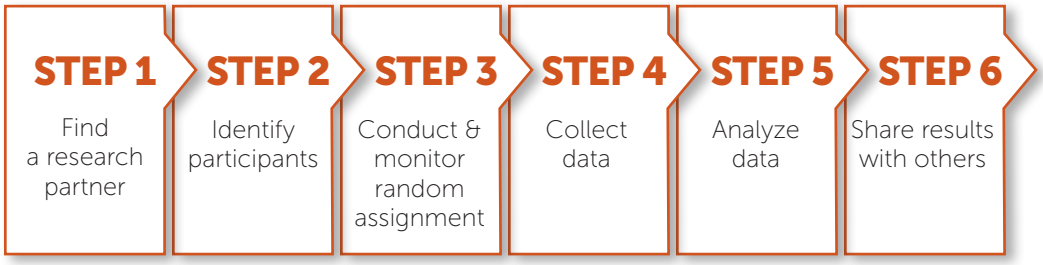
Agencies do not have to conduct RCTs on their own. You can partner with research ers at universities or research organizations. Agencies themselves, however, play a critical role in identifying opportunities in time to conduct RCTs.

use the lottery to randomly assign interested individuals to treatment and control groups. This is a fair process whereby each applicant has an equal chance of being selected.

Agencies do not have to design and conduct RCTs on their own. You can partner with researchers at universities or research organizations. The agency, however, plays a critical role in identifying opportunities in time to conduct RCTs. By the time a research partner learns about such opportunities on their own—usually once an intervention has been rolled out—it may be too late to create an RCT. Developing an ongoing relationship with researchers could help you identify opportunities proactively. Regular meetings with researchers can provide you with a forum to talk about current initiatives and planned changes, and get feedback about potential research opportunities and findings. (The next section provides more information about how to find and work with a research partner.)

HOW DO YOU CONDUCT AN RCT?

Once you have identified a question of interest, there are several steps involved in conducting a successful RCT:[2]



Where can you find the existing evidence on a program or strategy?

Before you begin an RCT, you may wish to look for existing research/evidence about the type of program or strategy change you're considering. One resource is the Self-Sufficiency Research Clearinghouse (SSRC), an initiative sponsored by the Office of Planning, Research and Evaluation (OPRE) within ACF. The SSRC is an active and growing virtual portal of research on low-income and TANF families; it can be accessed at https://www.opressrc.org. Another helpful resource is the "Building Better Programs" project of the Center on Budget and Policy Priorities, which provides an online resource bank of effective and promising human services program models. The site can be accessed at http://www.buildingbetterprograms.org/. In addition, OPRE recently launched the Employment Strategies for Low-Income Adults Evidence Review project to identify evidence-based programs and services that help low-income individuals get and maintain employment and achieve self-sufficiency. A related website and searchable database will be available in Fall 2015 at http://employmentstrategies.acf.hhs.gov.

[2] For more information on recognizing and conducting random assignment experiments, see Meckstroth, A., A. Resch, J. McCay, M. Derr, J. Berk, and L. Akers (2015). *Advancing Evidence-Based Decision Making: A Toolkit for Recognizing and Conducting Opportunistic Experiments in the Family Self-Sufficiency and Stability Policy Area*, OPRE Report #2015-XX, Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Step 1: Find a research partner. You may have staff in your agency who can conduct RCTs. If you do not, it can be helpful to collaborate with a research partner. To find a partner, you could recruit a researcher from a local university, contact a research organization, or reach out to researchers who are conducting a study or collecting data in your state.

Step 2: Identify participants. You need to identify participants—the sites, clients, or even staff (if, for example, the intervention is related to staff training)—who will serve as members of the treatment and control groups in the study. The type of participants you identify will depend on the program or strategy you are testing and what you are trying to learn. In the earlier example, the agency specifically identified TANF clients with disabilities to pilot an integrated case management and service delivery model. An agency could also rollout and test different communication strategies using its full caseload of TANF clients.

It can be easy to identify participants. For example, to test the effectiveness of different kinds of mailings to clients, the participants are already naturally defined. The same is true when an excess number of clients are interested in taking part in a program; the entire interested group becomes the participants (using the staggered rollout or lottery-style approaches previously described).

In certain circumstances, however, securing enough participants can be a challenge and agencies should plan their participant recruitment strategy carefully. In general, studies provide better, more reliable, and more precise answers when they have more participants. To recruit enough participants for the treatment and control groups, you may need to work with your staff, as well as with a research partner (see box below).

How can agencies encourage people to participate in an RCT?

- Emphasize the benefits of the RCT, such as enabling the agency to better serve clients and to more effectively allocate its resources.
- Discuss the "costs" of not doing the RCT, such as continuing to use—and spend valuable resources on—a program or strategy that may not be effective.
- Convey that, in the common situations of over-enrollment, limited resources, or both, randomization is a fair and transparent way to distribute resources or services.
- Assure stakeholders that the study is designed to impose minimal disruption for agencies, staff, and clients.
- If needed, create exemptions from the study. However, it is important not to overdo the number of exemptions as this may compromise the strength of your test. (For more information, see step 3.)

Step 3: Conduct and monitor random assignment. As previously emphasized, the assignment of participants to treatment and control groups needs to be random. This is not necessarily difficult: Microsoft Excel® spreadsheets have random-number generators that can be used to randomly assign groups. After you have created the groups, it is important to pay close attention to compliance with those assignments to maintain

integrity of the groups. For example, suppose some control sites scheduled for "stage 2" rollout start the intervention during "stage 1," because their clients are struggling and the sites want to offer the new intervention. This would violate the integrity of random assignment and may skew the findings, because the treatment and control groups are no longer random. That is, if these sites are struggling, then including them in the "stage 1" treatment group might lower outcomes of that group and raise outcomes of the "stage 2" control group (which would then include fewer struggling sites). What otherwise might have been shown to be an effective intervention could then appear ineffective because the integrity of random assignment was compromised.

Using random assignment does not mean agencies and service providers give up flexibility and discretion entirely. For example, a local site may want to admit relatives of clients who are offered a slot in the treatment group regardless of the lottery. Or, a site may wish to provide the intervention to a particular client or family due to special circumstances. Bypassing random assignment in these cases is effectively the same as excluding these participants from the study. Agencies should balance excluding individuals with what they want to learn from the study. For example, exempting TANF clients with multiple barriers from random assignment for a study of a new, more intensive job readiness and life skills program—possibly based on the logic that these clients need the most help—means the study cannot show whether the program helped the highest-need clients. Yet, knowing whether the program helped these clients might have been the study's biggest contribution.

Step 4: Collect data. For some RCTs, your agency may already collect the necessary data (such as administrative records of benefits received, employment in a given month or quarter, and earnings). You should first confirm within your agency that staff would not violate any privacy protections by using these data for a study. Other RCTs may require the collection of additional outcomes data. For example, a study of a program to improve clients' employment outcomes may benefit from administering short surveys to clients to collect data on the details of their employment experiences that might not be available in program administrative data (such as type of job held, wages and benefits received, hours worked, and number of months employed). In these cases, you can work with your agency staff and/or research partners to design the survey while minimizing cost and disruption.

Step 5: Analyze data. An important benefit of using random assignment is the simplicity of the analysis: To assess an intervention's impact, you calculate the difference in the average outcome you are measuring between the treatment and control groups. It doesn't have to be complicated! In the ongoing example, the agency compared average earnings for members of the treatment group ($2,882) to average earnings for members of the control group ($1,235) after one year of the intervention. These simple analyses can also be conducted for "subgroups" of clients within the treatment and control groups, such as male clients or those with prior work experience. As long as the characteristics used to define these subgroups are ones that cannot change (such as gender, race or ethnicity) or ones that were measured at baseline (such as time on TANF or work experience before enrollment), subgroups can function as smaller RCTs within the larger one.

Step 6: Share results with others. Research is most useful when its lessons are shared with others. Agencies can report their results to local sites through internal channels or in group forums or meetings. You may view an RCT as a means to inform your own agency's decisions, but disseminating the results can also help other states and agencies that face the same or similar decisions. Publishing results or presenting findings, for example, via state or regional webinars and federal technical assistance networks can also facilitate knowledge-sharing and spur innovations that may benefit many agencies. Ultimately, disseminating results from your research helps create a community of evidence-based decision making in the policy and program area of family self-sufficiency and stability, thereby improving services and strengthening families.

WHAT ARE SOME COMMON QUESTIONS AND CONCERNS ABOUT RCTS?

An RCT may raise questions or concerns for you, your agency staff, and the clients you serve. Is the study fair? How much will it cost? How long before we have findings? The following table provides answers to some of these frequently asked questions.

Answers to Common Questions and Concerns about RCTs

Is it fair to deny some clients, staff, and sites access to a new program or resource?	If you don't confidently know whether an intervention works, then you don't know whether you are denying clients, staff, or local sites access to anything that will actually help them. Conducting an RCT is the best way to find out. Random assignment is also an equitable way to decide who receives the intervention because all participants have the same chance to receive it.
If clients, staff, or local sites are assigned to the control group, can they ever receive the intervention?	Yes. Assignment to the control group is not permanent; it only lasts for the duration of the study. Moreover, agencies can assure sites in the control group that they are next to receive the intervention as soon as resources are available. For example, in a staggered rollout, control sites could receive the intervention several months later than treatment sites.
Aren't RCTs expensive?	Not necessarily. The two largest expenses associated with RCTs usually involve identifying participants and collecting data. However, RCTs can be cost-effective if (1) agencies initiate the research—signifying that they are already interested in the intervention and willing to implement it—thus reducing efforts to identify participants, and (2) the RCT relies on data already collected by the state or agency. The costs of conducting a study should also be considered in light of the time and resources an RCT could save the agency—for example, avoiding the implementation of an ineffective program can lead to big savings.
Aren't RCTs disruptive for staff and clients?	Not necessarily. Using existing data minimizes disruption for agencies. Testing an intervention that would have been implemented with or without the RCT also involves minimal disruption.
Will we have to wait years to find out the results?	Not necessarily. Studies that focus on short-term impacts can produce results quickly. For example, clients who learn a skill set may be more likely to obtain a sustainable job and increase their earnings, which are long-term outcomes. However, these same clients may also obtain a child support order or successfully complete a job-coaching, goal-setting course. In fact, the whole point of the intervention could be to improve short-term outcomes. An RCT can examine short-term outcomes while also shedding light on longer-term outcomes.

CONCLUSION

At every level of our human services system, leaders need to know which programs and policies are most effective as they allocate scarce resources. Agency administrators and staff who work directly with families can be the first to identify opportunities to conduct RCTs and learn about an intervention's impact or a strategy's effectiveness. You can help discover what works and not only improve your program, but also share findings with states and agencies nationwide.

OPRE

MATHEMATICA
Policy Research