By Lauren Akers,
Alexandra Resch,
and Jillian Berk
Mathematica Policy Research

Making the Most of Opportunities to Learn What Works: A School District's Guide

If you are a school district leader, principal, or other education official, you are always deciding how best to use your school's or district's resources. You want to use those resources on programs that improve outcomes for your students, teachers, and schools—programs that have been proven effective. Unfortunately, many programs and curricula have no proof of effectiveness, and even when they do, the evidence may come from studies with weak research designs or from a different context from that of your district. For example, findings from a study done in small, rural schools may not be relevant to an educator in a large, urban school.

Facing these challenges, some districts may forge ahead with untested or unproven programs, potentially spending resources and years using something that doesn't work. But there is a reliable and accessible tool—known as a randomized controlled trial, or RCT—that you can use to test the effectiveness of a program before rolling it out. If the program is not effective, you can find alternatives and invest resources elsewhere. If it is effective, you can continue or expand the program with confidence.

WHAT IS AN RCT?

An RCT is a type of study that demonstrates whether an *intervention*—a program, policy, or other change—causes a certain outcome. All RCTs start with two basic components: a *treatment group* (those who will receive the intervention) and a *control group* (those who will not).[1] The study is "randomized" because people are randomly assigned to each group, ensuring that the groups are as similar as possible at the beginning of the study. This is important to ensure that you are comparing "apples to apples." (For a definition of the italicized terms, see sidebar.)

To see how this works in practice, consider how one district took advantage of an opportunity to learn whether an intervention worked. The district had a summer reading program for kindergarten and first-grade students who were at high risk for reading difficulties. The district wanted to expand the program to include students who were only at moderate risk. It did not know whether the program would work for those students. To find out, the district partnered with researchers at a university to conduct an RCT. The researchers identified students at moderate risk for reading difficulties and randomly selected half of them—the treatment group—to receive an invitation to the program. The other half—the control group—were not invited to participate. That fall, the researchers analyzed the data and found that the program improved reading outcomes for moderate-risk students.[2] District leaders used this evidence to make an informed decision about whether to permanently expand the program to moderate-risk students.

Like this district, you may be able to take advantage of opportunities to conduct RCTs and use the findings to inform how you allocate resources. The remainder of 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) follow the key steps involved in conducting an RCT.

Key Terms in RCTs

Intervention: The policy, program, or practice being evaluated.

Random assignment: The statistical process by which study participants—students, teachers, or schools—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).

May 2014
REL 2014-048

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

WHY SHOULD WE USE RCTS?

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

Why Do We Need to Compare Two Groups to Find Out What Works?

We need to know whether outcomes for those who participated in the intervention differ from those who did not. Consider the example of the RCT conducted of the summer reading program. It compared two groups of students at moderate risk for reading difficulties: a treatment group, which was invited to take part in the reading program, and a control group, which was not. But what if the district had decided to enroll all moderate-risk students at a single pilot school in the program, leaving no students for comparison? If students' reading skills improved, would it be reasonable to conclude that the program caused the improvement? (Figure 1)

Let's suppose the district had looked at how moderate-risk students were doing at other schools. They might have found that on average moderate-risk students in the district were improving in reading. (Figure 2) The district would not know if the summer program caused the improvements. For example, other district-wide programs might have helped to improve reading. Looking around, the district might have observed that free books were distributed to students at the end of school, a summer reading program was operating in local libraries, and a media campaign encouraged parents to read with their children during the summer. Whether the district's summer reading program was the reason why student performance improved remains a question.

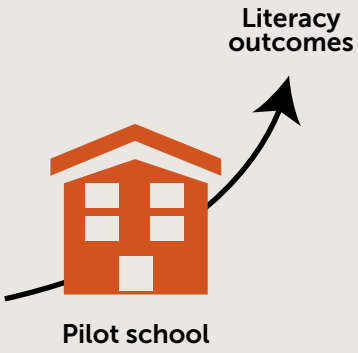


Figure 1

Reading outcomes improved at the pilot school—does that mean the program worked?

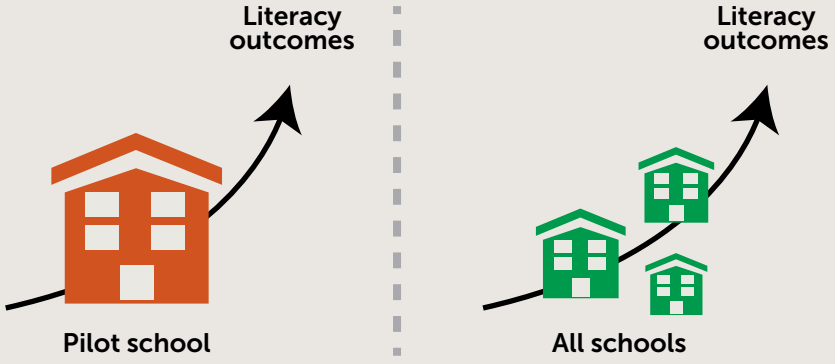


Figure 2

Reading outcomes improved at all schools, not just the pilot school, so we can't say that the pilot program caused the improved reading at the pilot school.

on page 1

[1] A control group may be assigned to receive no intervention, to continue with "business-as-usual" (which may entail receiving an intervention already in use), or to receive an alternative intervention.

[2] Zvoch, K., and J. J. Stevens. "Summer School Effects in a Randomized Field Trial." Early Childhood Research Quarterly, vol. 28, no. 1, 2012, pp. 24–32.

How Does Random Assignment Let Us Compare Apples to Apples?

To know whether a program has the desired effect, you can't just compare two groups; you must compare two similar groups. That's where random assignment comes in. It creates an "apples-to-apples" comparison. In the example, moderate-risk students were randomly assigned to the treatment and control groups. Because assignment was random, differences between the two groups arise simply by chance, except for the crucial one: exposure to the program. Because the two groups were not different except for exposure to the program, differences in outcomes had to be caused by exposure to the program.

Without Random Assignment, It's Easy to End Up Comparing Apples to Oranges

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

For example, suppose the district had asked parents of students at moderate risk to sign up their children for the summer reading program. After participating in the program, students whose parents signed them up had better reading outcomes than other students. Did the summer program cause the improvement? Maybe the students whose parents signed them up for the program were different from other students. Perhaps they had more books at home or parents who read with them more often. Or perhaps the parents knew that their children already liked to read and thought that the children would enjoy the program. These resources and attitudes may have caused the improvement, rather than the program. The point is that, without random assignment, we still have questions about whether it was the program or other factors that made a difference. (Figure 3)

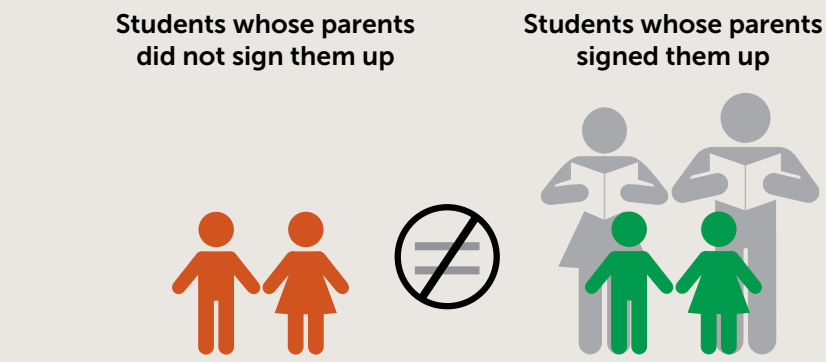


Figure 3
Students whose parents volunteered them for the program may have more involved parents, so it's not a fair comparison.

WHEN ARE SOME GOOD OPPORTUNITIES TO CONDUCT AN RCT?

You may encounter many situations in which conducting an RCT would be beneficial. Here are some opportunities to look for:

- **Pilots of a new program or curriculum.** A district considering a new program or curriculum might ask schools to volunteer to pilot it. There's the opportunity: test the intervention by randomly assigning volunteer schools to a group that pilots the program and a group that carries on with business as usual.
- **Limited resources to roll out a new program or a preference to implement it in stages.** A district might not have the resources to roll out a new program in all schools, or might prefer to launch it in stages. There's the opportunity: test the intervention by randomly choosing schools for the first stage of a staggered rollout. Schools chosen for later rollout now are a control group. The district can conduct an RCT during the first stage of the rollout by comparing outcomes

of the schools who implement the program during the first stage to those who will implement it later.

- **Communication efforts.** A district might try to influence students or parents through materials such as information about college prep resources. There's the opportunity: the district could use three approaches to send materials to parents—such as email, postal mail, and giving the materials to students to give to their parents. By randomly choosing which method is used, districts can learn which is most effective.

- **Excess demand for a new program.** A district might see excess demand for a program that has a limited number of slots. There's the opportunity: the district can use a lottery—which is a type of random assignment—to assign slots to a portion of those who are interested in the program. Evaluations of charter schools and voucher programs have used this lottery mechanism. Before-school or after-school programs with excess demand also can be studied using lotteries.

Districts do not have to conduct RCTs on their own. They can partner with researchers in universities or research organizations. But schools and districts play a key role in identifying the opportunities to conduct RCTs. By the time research partners learn about these opportunities, it may be too late to create an RCT. Developing an ongoing "thought-partner" relationship with researchers could help you identify opportunities early on. Regular meetings with researchers can provide districts and schools with a forum to talk about current initiatives and get feedback about research opportunities, while allowing researchers to discuss their results from the field and help identify potential research opportunities. (The next section provides more information about finding and working with a research partner.)

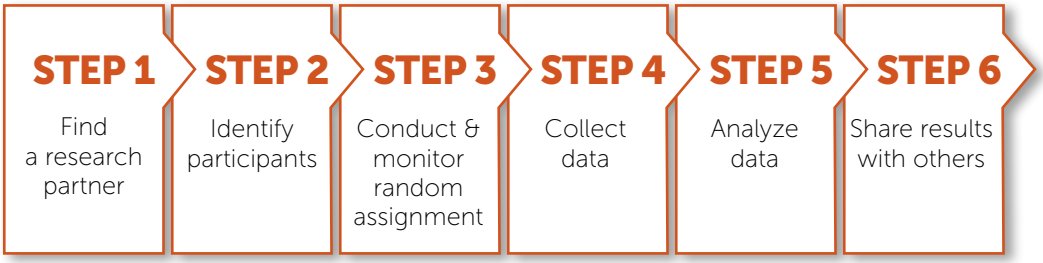
Districts do not have to conduct RCTs on their own. They can partner with researchers in universities or research organizations. But schools and districts play a key role in identifying the opportunities to conduct RCTs.

Where Can I Find the Existing Evidence on a Program?

Before you begin an RCT, you may wish to look for existing evidence about the program or policy you're considering. An excellent resource is the U.S. Department of Education's What Works Clearinghouse (WWC). The WWC reviews research on programs, policies, practices, and products to identify high quality studies on what works in education. Ultimately, the WWC aims to equip educators with the information they need to make evidence-based decisions. The Clearinghouse can provide districts with rigorous evidence on some curricula and programs. Visit http://ies.ed.gov/ncee/wwc/ for summaries of its findings.

HOW DO I CONDUCT AN RCT?

There are several steps involved in conducting a successful RCT:



Step 1: Find a research partner. Districts may have staff that can conduct RCTs. Those that don't may want to recruit a research partner. To find a partner, districts could recruit researchers at local universities (the district in the example did that). Districts could also reach out to researchers who have submitted requests to conduct a study or collect data in the district. You can also consult the U.S. Department of Education's Regional Educational Laboratories (RELs), whose primary mission is to collaborate with state departments of education, school districts, and others on the use of research and evidence. The RELs can act as a research partner or can assist you in identifying one. Find out more at http://ies.ed.gov/ncee/edlabs.

Step 2: Identify participants. You need to identify participants—the schools, classrooms, teachers, or students who will be in the study. The type of participants you identify will depend on the program you are testing and what you are trying to find out. In the earlier example, the district identified moderate-risk students who scored within a certain range on the spring reading assessment. Or districts could invite schools to participate in a staggered rollout—for example, letting them know that some schools will start the program in the first stage and the rest will start in the second stage.

Identifying participants can be easy. To test the effectiveness of different kinds of mailings to parents, the participants already are known. The same is true when an excess number of students apply to a charter school or summer program. The whole group becomes the participants.

How Can District and School Administrators Encourage People to Participate in an RCT?

- Emphasize the benefits of the RCT, such as enabling the district and schools to better serve students and to more effectively allocate its resources.
- Discuss the "costs" of not doing the RCT, such as continuing to use—and spend resources on—a program that may not work.
- Convey that in the case of overenrollment, limited resources, or both, randomization is a fair and transparent way to distribute resources or services.
- Explain that in the case of a randomized staggered rollout, the control group will still get to participate in the intervention during stage 2 of the rollout.
- Assure stakeholders that the study is designed to impose minimal disruption for schools, classrooms, and students.
- If needed, create exemptions from the study. It's important not to overdo the number of exemptions. (For more information, see Step 3.)

However, for some studies, districts need to plan their participant recruitment strategy carefully. Generally, studies provide better answers when they have more participants. To recruit enough participants, districts may need to work with their staff or with research partners (see sidebar).

Step 3: Conduct and monitor random assignment. Assignment to treatment and control groups needs to be random. This is not difficult: spreadsheets have random-number generators that can be used. After creating the groups, some attention to compliance is needed to maintain integrity of the assignment. For example, suppose some control schools scheduled for "stage 2" rollout start the intervention during "stage 1," perhaps because the schools are struggling and really want to get started. That would violate random assignment and may skew the findings. If these schools are struggling, then having 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 schools). What otherwise might have been an effective program may then appear ineffective because the integrity of random assignment was lost.

Using random assignment does not mean schools and districts give up flexibility. For example, a charter school may want to admit siblings of students who are offered a slot regardless of the lottery. Or a school may need to provide an intervention to a particular child—for example, a child under court protection may need to be in the after-school program.

Not using random assignment in these cases is the same as excluding them from the study. Districts need to balance exclusion decisions with what they need to learn from the study. For example, exempting low-skill readers from random assignment for a study of a new supplemental reading program—possibly based on the logic that these students need the most help—means the study of the program cannot show whether the curriculum helps low-skill readers. But knowing whether the program helped low-skill readers might have been the study's biggest contribution for the district.

Step 4: Collect data. For some RCTs, there may actually be no data collection burden: the needed data (such as scores on state assessments) may already be collected. Districts may need to check with the director of assessment or the research office to confirm that district staff do not violate any privacy protections by using these data. Other RCTs may require collecting data. For example, a study of a program to enhance student engagement may require administering surveys to students to measure their engagement. In these cases, districts can work with their staff or research partners to minimize cost and disruption.

Step 5: Analyze data. An important benefit of using random assignment is the simplicity of the analysis: to see an intervention's impact, you calculate the difference in the average outcome between the treatment group and the control group. It's arithmetic. In the above example, researchers compared the average score on a reading test for the treatment group to the average score for the control group. The simplicity extends to "subgroups" of students within the treatment and control groups, such as students scoring below a threshold in the previous year. As long as characteristics used to define the subgroups are ones that don't change (such as race or ethnicity) or ones that were measured at baseline (such as free lunch eligibility), then subgroups function as smaller RCTs within the larger one.

More complex RCTs may need more complex data analysis, such as when students are assigned to treatment and control groups in unequal proportions. In these cases, working with research partners may be valuable if district staff do not have the technical expertise.

Step 6: Share results with others. Research is most useful when its lessons are shared with others. Districts can report their results to district managers through internal channels or in larger forums, such as board meetings. A district may view an RCT as a means to inform its own decisions, but disseminating its results can also help other districts that face the same decision. Publishing results or presenting findings at professional networking events can generate broader knowledge and spur innovations that benefit all districts. Ultimately, disseminating results can help create a community of evidence-based decision making in education.

WHAT ARE SOME COMMON QUESTIONS AND CONCERNS ABOUT RCTS?

An RCT may raise questions or concerns among parents, teachers, and school staff. Is the study fair? How much will it cost? How long before we have findings? The table below provides answers to some frequently asked questions.

Answers to Common Questions and Concerns about RCTs

Is it fair to deny some schools, teachers, or students access to a new program or resource?	If we don't know whether an intervention works, we don't know whether we're denying schools, teachers, or students access to anything that will actually help them. Conducting the RCT is the best way to find out. Random assignment is also the most equitable way to decide who receives the intervention because all participants have the same chance to receive it.
If schools 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. And, districts can assure schools in the control group that they are next to receive the intervention when resources are available. For example, in a staggered rollout, control schools could receive the intervention a year later than treatment schools.
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) districts 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 district. The costs should also be considered in light of the time and money an RCT could save the district—for example, not implementing an ineffective program can mean big savings.
Aren't RCTs disruptive for teachers and students?	Not necessarily. Using existing data causes minimal disruption for districts. 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, students who read more proficiently than their peers may be more likely to graduate from high school and attend college, which are long-term outcomes. But these same students may also score higher on current tests and may act out less in class, which are short-term outcomes. An RCT can examine short-term outcomes while also shedding light on long-term outcomes.

CONCLUSION

At every level of our education system, leaders need to know which programs and policies are effective to allocate scarce resources well. District and school officials can be the first to identify opportunities to conduct RCTs and learn about an intervention's effectiveness. You can help discover what works and share these findings with schools and districts nationwide.

Akers, L., Resch, A., & Berk, J. (2014). Making the most of opportunities to learn what works: A school district's guide (REL 2014–048). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Analytic Technical Assistance and Development. Retrieved from http://ies.ed.gov/ncee/edlabs.