

When Is It Possible To Conduct a Randomized Controlled Trial in Education at Reduced Cost, Using Existing Data Sources?

A Brief Overview

Coalition for Evidence-Based Policy

A Project Sponsored by

THE COUNCIL FOR
Excellence
IN GOVERNMENT



February 2007

This publication was produced by the [Coalition for Evidence-Based Policy](#), under a contract with the U.S. Education Department's Institute of Education Sciences (Contract #ED-02-CO-0022). However, the views expressed are those of the Coalition and do not necessarily represent the positions or policies of the Institute of Education Sciences.

We wish to acknowledge the valuable comments we received, in preparing this document, from the following individuals: Robert Boruch, Sandra Gregerman, Rebecca Herman, Robinson Hollister, Rebecca Maynard, Stephen Nettles, Denis Newman, Howard Rolston, Robert Slavin, and Joseph Torgesen.

This publication is in the public domain. Authorization to reproduce it in whole or in part for educational purposes is granted.

We welcome comments and suggestions on this document (jbaron@excelgov.org).

Purpose and Overview

Purpose: To advise researchers, policymakers, and others on when it is possible to conduct a high-quality randomized controlled trial in education at reduced cost.

Well-designed randomized controlled trials are recognized as the gold standard for evaluating the effectiveness of an intervention (i.e., program or practice) in many diverse fields, such as medicine, welfare and employment, psychology, and education.¹ But they have been relatively rare in education, in part because of a perception among policymakers, researchers, and others that such studies are too costly and too administratively burdensome on schools to be practical.²

This guide explains that, in many circumstances, it may in fact be possible to conduct a randomized controlled trial at modest cost and with minimal burden, by measuring outcomes using school-administered test scores or other administrative data that are already collected for other purposes. The cost may be as low as \$50,000 under certain conditions.³

Trials using such a reduced-cost approach can often answer research questions that are directly relevant to policymakers and educators, such as:

- *In our school district, what effect will this new math curriculum have on student test scores compared to the curriculum currently in use?*
- *What effect will our efforts to improve teacher quality – e.g., by revising the certification requirements for new teachers, or by adopting a new professional development program – have on student test scores, special education placements, and grade retentions?*
- *What effect will this new career track for struggling high school students have on attendance, graduation rates, and post-graduation employment and earnings?*
- *What effect will this new violence prevention program have on student suspensions and arrests?*

Such trials may not be able to tell you what effect the intervention has on a variety of other outcomes, or *why* the intervention has (or does not have) an effect. Obtaining answers to these questions may require more comprehensive and costly studies. But reduced-cost trials can often provide policymakers with the information they need to decide whether an intervention merits continuation, revision, expansion, and/or further investigation in more comprehensive studies.

Overview: This guide includes two main sections:

- ▶ **Conditions that offer the opportunity to conduct a randomized controlled trial at reduced cost – namely:**
 1. High-quality administrative data (e.g., school-administered test scores) are available to measure key outcomes the intervention seeks to affect (e.g., reading achievement); and
 2. Top school district and/or school officials will work with you as partners in the trial, providing you with ready access to such data and facilitating the random assignment.
- ▶ **Examples of well-designed randomized controlled trials conducted at reduced cost.**

Conditions that offer the opportunity to conduct a randomized controlled trial at reduced cost

What follows is an overview of conditions that offer the opportunity to conduct or sponsor a well-designed randomized controlled trial at reduced cost, to evaluate an educational intervention. This is intended as a discussion of general principles, rather than as a detailed guide to conducting such a trial.

Importantly, the costs discussed here are the costs of the trial itself (e.g., designing the study, conducting the random assignment, collecting and analyzing outcome data), and do not include the costs of the educational intervention that the trial is evaluating (e.g., curriculum materials, teacher training).

Condition 1

Administrative data (e.g., school-administered test scores) are available to measure key outcomes that the intervention seeks to affect (e.g., reading achievement).

Such data can reduce a trial's cost by eliminating what is typically the most labor-intensive and costly part of a trial – namely, locating the individual sample members at various points in time after the intervention is completed, and administering surveys, tests, interviews, and/or observations to obtain their outcome data. If, instead, key outcome data are readily available for most or all sample members from administrative data sources, the cost of data collection can be reduced to a nominal amount. This can be true even for trials involving large numbers of students, including trials that randomize whole schools or classrooms, rather than individuals.

A. To use administrative data to measure outcomes, you ideally need to obtain such data for individual sample members (e.g., individual students), not just for aggregated groups (e.g., classrooms).

This is because most trials seek to estimate the intervention's effect on the individuals (i.e., students) who participate in the intervention.⁴ To obtain a valid estimate of this effect, you need to compare outcomes for the individuals in the intervention group to outcomes for individuals in the control group (this is true even if the trial randomizes classrooms or schools rather than individual students). Thus, aggregate outcome data on classrooms or schools will suffice for a trial with a short-term follow-up, where the individual students remain in their original classroom or school for the duration of the study (e.g., one school year). But if you wish to estimate the intervention's effect over a longer period of time – as students move on to new classrooms and schools – you need access to their individual data.

B. Fortunately, administrative entities (e.g., states, school districts, individual schools, juvenile courts) collect many types of data on individuals, not just aggregated groups.

Examples of data that administrative entities typically collect on individuals include: scores on standardized achievement tests, disciplinary suspensions, attendance, grade retentions, special education placements, attendance, high school graduation, employment and earnings after graduation, and criminal arrests. If a researcher can gain access to such data on individuals, he or she can often use it to measure outcomes for individual sample members even as they move to new classrooms or schools, as long as they stay within the jurisdiction of the entity (e.g., state or school district) collecting the data.

C. Whether an administrative entity (e.g., state or district) will provide you with ready access to such data is often the key factor in whether a reduced-cost trial is possible.

The Family Educational Rights and Privacy Act (FERPA) allows researchers that are conducting studies for, or on behalf of, educational agencies or institutions to gain access to individual student scores and other educational records without parental consent, provided the researchers take specific precautions to protect student privacy that are described in the Act (see the relevant statutory provision in endnote 5).⁵ Similarly, non-educational agencies, such as state unemployment insurance agencies and juvenile courts, are typically allowed to provide researchers with data on individuals (e.g., their wages, arrests), provided the researchers take similar required actions to safeguard privacy.

But even though such agencies are allowed to provide individuals' data to researchers who observe these precautions, whether they actually will do so in a particular case depends on a number of factors.

1. As a general matter, the following conditions increase the likelihood that you can gain ready access to the agency data on individuals that you need for your study:

- **You are requesting data from an agency that has a direct stake in your study (e.g., the state agency or school district that is sponsoring it),** rather than from another agency (e.g., a school district participating in a national, federally-funded study). This is because the agency with a stake in your study has a self-interest in giving you the data needed to carry it out in a cost-effective way.
- **The data you need for your study are all kept by one agency (e.g., school district) rather than several (e.g., the individual schools in the district),** because it is usually easier to negotiate access with one entity than with several.
- **You are able to engage a top agency official (e.g., school district superintendent or key state official), as a partner in your study** – specifically, an official who is in a position either to grant you access to the data directly or to intercede with others to grant you such access. (The importance of such a partner is discussed in more detail below.)
- **Agency officials and staff regard you as a trustworthy party who will faithfully observe the legally-required privacy precautions.** The guide cited in endnote 6 summarizes the precautions, and discusses steps you can take to reassure school district officials and others that you will observe them.⁶

2. If you find that you cannot readily gain access to administrative data on individuals, two possible courses of action are –

- **To use *aggregate* data on schools or classrooms to measure outcomes.** As noted above, such data will suffice for a trial with a short-term follow-up, where the individual students remain in their original school or classroom for the duration of the trial. Aggregate school-wide test scores are often readily accessible on school or district websites.
- **To undertake efforts to persuade agency officials to provide you with access to data on individuals – efforts such as those discussed in endnote reference 6.** Of course, such efforts are not costless, and thus may raise your trial's cost of data collection

above nominal levels. In the end you may decide that collecting outcome data yourself (e.g., by testing and/or interviewing sample members) is the more cost-effective approach.

D. Administrative data that you use to measure outcomes must be (i) a valid measure of the outcomes the intervention seeks to improve, (ii) a uniform measure for all sample members, and (iii) complete.

For example, if you wish to use a school-administered standardized test to measure the effect of an algebra curriculum for middle-school students, you need to be sure that the test is (i) a valid measure of middle-school algebra skills – that is, highly-correlated with true algebra skill levels; (ii) applied uniformly to sample members (e.g., administered at approximately the same time in the school year); and (iii) administered to all (or nearly all) sample members.

Similarly, if you wish to use school disciplinary records to measure the effect of an intervention to prevent misconduct, you need to be sure that the records are (i) a valid measure of such misconduct – that is, highly-correlated with true acts of misconduct rather than being a reflection of arbitrary teacher judgment; (ii) a uniform measure for all sample members (e.g. based on approximately the same standards of conduct); and (iii) complete – that is, free of any serious instances of missing data. Thus, school records of suspensions are usually a better measure of misconduct than records of referral to detention, because suspensions tend to be more highly-correlated with serious misconduct and to be applied based on a uniform set of criteria, and are more often complete than records of detentions.

The requirement for the administrative data to be a uniform measure for all sample members can sometimes be a barrier to using such data in a multi-site trial – for instance, one spanning several school districts that each administer the state’s achievement tests at different times of the year, or have different disciplinary standards (to continue with the above examples). In such cases, researchers seeking to use these data would need to invest effort to make the data comparable across sites – effort that may well raise the cost of using administrative data above nominal levels and thus diminish its cost advantage versus collecting original outcome data.

E. The following are examples of administrative data that can be, or have been, used to measure outcomes in randomized controlled trials of various educational interventions.

The endnotes provide references to trials where these data have been used. These trials evaluated a broad range of educational interventions, with widely varying goals – goals such as improvement in academic achievement, prevention of delinquent or criminal behavior, improvement in young adult employment and earnings, and increase in high school graduation rates.

- Standardized state reading test scores⁷;
- School records of disciplinary suspensions⁸;
- School records of assignments to special education, and of high school graduation⁹;
- School records of grade retentions, assignments to summer school, and absenteeism¹⁰;

- Quarterly employment and earnings data reported by employers to state unemployment insurance agencies¹¹; and
- Juvenile court records, and local and state police records, of arrests for criminal behavior.¹²

Condition 2

Top school district and/or school officials will work with you as partners in the trial, facilitating the random assignment and providing you with access to administrative data.

A. The engagement of such officials is essential to conducting a trial at reduced cost.

The main reason is that it facilitates what can otherwise be a labor-intensive effort to persuade teachers, parents, program providers, and others that random assignment is ethical, practical, and desirable. In most cases, you will still need to make this case to the various parties, but the task is invariably easier if a top official joins you to reinforce the message and underscore the importance of the trial to the district and/or school.

The cost of making this case for random assignment is a main reason why randomized controlled trials are often more expensive than comparable comparison-group studies without random assignment. Strong, visible support from top district and/or school officials in making this case can reduce this additional expense to a modest or nominal amount – in some cases, tipping the cost advantage to the randomized controlled trial.¹³

A second reason the engagement of top officials is essential is that their support is usually needed for you to gain access to the administrative data that will be used to measure outcomes, particularly data on individual students, as discussed above. In some instances, the officials may need to intercede with district or school staff whose excessive caution or overly-narrow legal interpretations might otherwise thwart your access to such data.

B. The funder of the educational intervention being studied in the trial often can help greatly in securing the partnership of such top officials, and thereby reduce the cost a trial.

The funder can do this by strongly urging districts or schools receiving the funds to participate in a trial to evaluate the intervention's effectiveness, or even requiring them to participate as a condition of the funding award. As an example, the U.S. Education Department required each school district awarded a grant in the Department's Striving Readers program in 2006 to have an independent research team conduct a randomized controlled trial of the district's project.¹⁴ In other programs, the Department has given a competitive priority to grant applicants proposing to evaluate their project in such a trial,¹⁵ or has encouraged states to do this in making sub-grants to school districts.¹⁶

Examples of well-designed randomized controlled trials conducted at reduced cost

Example 1: A trial in Seminole County, Florida to evaluate the effectiveness of three remedial reading interventions for struggling readers in grades 9 and 10.

Specifically, Seminole County Public Schools, Florida State University Learning Systems Institute, and the Florida Center for Reading Research recently launched a randomized controlled trial to evaluate the effectiveness of three interventions designed to improve the reading ability of struggling readers in grades 9 and 10. The interventions were Scholastic’s Read 180, the SRA Reach System, and the Reading Instruction through Strategy Enhancement (RISE) intervention developed by the Florida Department of Education. This trial is currently underway.

At the start of the school year, 1532 struggling readers entering 9th or 10th grade in 10 district high schools were randomly assigned¹⁷ within their school to remedial reading classrooms using one of the three interventions or to a control classroom, where the school’s usual (pre-existing) approach to remedial reading was used. Each intervention was administered for 90 minutes per day, for one school year.

The primary outcome measures will be the students’ scores at the end of the school year, and the end of the following school year, on the reading subtest of the Florida Comprehensive Assessment Test (FCAT), which is administered annually to all Florida public school students in grades 3-11. These test scores will be obtained for all sample members who remain in their school as well as those who transfer to another school within the district, thus permitting a high rate of sample retention. The researchers are also administering the Woodcock Reading Mastery Test (WMRT) to a subsample of about 350 students at the start and end of the first school year, in order to obtain a more finely-tuned measure of specific reading skills affected by the interventions.

A. The trial’s cost is very low – about \$62,000 – and would be even lower (\$44,500) if the study were done without administering the WMRT as a secondary outcome measure.

The trial’s costs – other than the cost of administering the WMRT— include: (i) the researchers’ time to design the study, conduct the random assignments, and analyze the data; and (ii) the partial salary of the on-site project manager to monitor schools’ adherence to the study protocol, including the random assignment. These represent the complete costs of conducting the trial to evaluate the three interventions (these costs do not include the cost of the interventions themselves – such as curriculum materials, teacher training, and the on-site project manager’s time facilitating implementation – which the school district is paying for separately as part of its remedial reading efforts).

B. The two critical factors enabling the trial to be conducted at low cost are:

- (1) The strong support of the school district superintendent for this study, including the random assignment.** This support meant that the researchers did not have to spend a great deal of time and effort to persuade school administrators, staff, and others to agree to the random assignment, or to provide access to student test score data.
- (2) The use of administrative data (i.e., scores on the FCAT reading subtest) as the main outcome measure.** This will enable the trial to obtain, at a nominal cost, nearly complete outcome data for the original sample of 1532 students over a two-year period.

Example 2: A trial of the University of Michigan's Undergraduate Research Opportunity Program (UROP).¹⁸

A. Description of the intervention: UROP, established in 1989, creates research partnerships between faculty members and undergraduates.

Specifically, UROP creates one-year research partnerships between faculty members and first and second-year undergraduate students – primarily under-represented minority students and women with interest in the sciences. The program gives students the opportunity to work closely with a faculty member in conducting literature reviews, formulating research questions, conducting studies, and, in some cases, co-authoring research presentations and journal articles. The program's goal is to reduce student attrition (i.e. students leaving school prior to graduating) by (i) providing program participants with a faculty mentor, and (ii) getting them excited about research early in their college careers.

B. UROP was evaluated in a large randomized trial, using administrative data from the university to measure the key outcome (percent of students leaving school before graduating).

The trial randomly assigned 1,334 freshmen and sophomores who applied to the program between 1990 and 1993 to either an intervention group that participated in the program, or a control group that did not. In the spring of 1994 (i.e. between one semester and three years after students completed the program), researchers used student enrollment data from the Registrar's office at the University to determine the percentage of students in each group that had left the school prior to graduating (i.e., percentage of student attrition). This data included information on each student's race, gender, and high school Grade Point Average, allowing researchers to examine the program's effects on a variety of subgroups. The researchers also surveyed students in the sample on their attitudes toward school.

The study found that UROP produced (i) a 25% overall decrease in student attrition (compared to the control group), which approached statistical significance (p-value = 0.17); and (ii) a statistically-significant 45% decrease in the attrition of African-American students.

C. The cost of this large, 4-year trial was modest – about \$50,000 per year – and would have been even lower (about \$43,000 per year) if the researchers had omitted the survey of students' attitudes.

The trial's costs included the researchers' time to design the study, conduct the random assignment, monitor the trial's implementation, and assemble and analyze the data. (The costs do not include the cost of the intervention itself – such as recruiting students and faculty into the program, matching them in research partnerships, and so on – all of which the university paid for separately through other funding sources.)

A critical factor enabling this trial to be conducted at modest cost was its use of the University's administrative data to measure the key outcome (student attrition). In addition to its low cost, this data source enabled the researchers to measure outcomes for almost the entire sample of students originally randomized¹⁹ – a follow-up rate rarely achieved in trials that collect their own data.

Additional Resources

What follows are other resources that can be helpful in sponsoring or conducting a well-designed randomized controlled trial at reduced cost:

- ***How To Solicit Rigorous Evaluations of Math and Science Partnerships (MSP) Projects: A User-Friendly Guide For MSP State Coordinators***²⁰
(<http://www.ed.gov/programs/mathsci/issuebrief.doc>)
This document outlines a strategy for state officials to solicit reduced-cost randomized controlled trials in the MSP program (a teacher professional development program). It includes suggested solicitation provisions. This strategy could be used, with appropriate adaptations, in other competitive grant programs that fund interventions to improve teacher quality.
- ***How To Conduct Rigorous Evaluations of Math and Science Partnerships (MSP) Projects: A User-Friendly Guide For MSP Project Officials and Evaluators***²¹
(<http://www.ed.gov/programs/mathsci/mspbrief2.doc>)
This is a guide on how to conduct a reduced-cost randomized controlled trial to evaluate a teacher professional development project or model funded by the MSP program. Many of its suggestions could be used, with appropriate adaptations, to carry out reduced-cost evaluations of other interventions to improve teacher quality (e.g., other professional development programs, alternate teacher certification requirements, new teacher recruitment strategies).
- ***Key Items to Get Right When Conducting a Randomized Controlled Trial In Education***²²
(http://www.whatworkshelpdesk.ed.gov/guide_RCT.pdf)
This document, intended for researchers and sponsors of research, describes items that are often critical to the success of a randomized controlled trial in producing valid evidence about whether an intervention is effective.

References

¹ See, for example, Office of Management and Budget, *What Constitutes Strong Evidence of Program Effectiveness*, http://www.whitehouse.gov/omb/part/2004_program_eval.pdf, 2004; U.S. Department of Education, “Scientifically-Based Evaluation Methods: Notice of Final Priority,” *Federal Register*, vol. 70, no. 15, January 25, 2005, pp. 3586-3589; the Food and Drug Administration’s standard for assessing the effectiveness of pharmaceutical drugs and medical devices, at 21 C.F.R. §314.12; “The Urgent Need to Improve Health Care Quality,” Consensus statement of the Institute of Medicine National Roundtable on Health Care Quality, *Journal of the American Medical Association*, vol. 280, no. 11, September 16, 1998, p. 1003; “Criteria for Evaluating Treatment Guidelines,” American Psychological Association, *American Psychologist*, vol. 57, no. 12, December 2002, pp. 1052-1059; and *Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination*, Society for Prevention Research, April 12, 2004, at <http://www.preventionresearch.org/sofetext.php>.

² Another reason randomized controlled trials are not more common in education may be a belief that it is unethical to deny an intervention to some individuals (i.e., the control group). We believe that in many (but not all) situations, a strong case can be made that a trial is indeed ethical, but such a discussion is beyond the scope of this guide. Two other resources that address this issue are Larry L. Orr, *Social Experimentation: Evaluating Public Programs With Experimental Methods*, Sage Publications, Inc., 1999, pp. 17-22; and *How To Conduct Rigorous Evaluations of Math and Science Partnerships (MSP) Projects: A User-Friendly Guide For MSP Project Officials and Evaluators*, produced by the Coalition for Evidence-Based Policy in partnership with the National Opinion Research Center, August 2005, pp. 11-12, at <http://www.ed.gov/programs/mathsci/mspbrief2.doc>.

³ Indeed, the costs in some cases may be negligible if sponsors of the trial can persuade (i) a capable statistician to donate a modest amount of time to design the study (e.g., estimate the minimum sample size needed) and analyze the outcome data, and (ii) an academic researcher to offer the services of his or her students to assemble the data and monitor the trial’s implementation (e.g., any deviations from random assignment). Many academic researchers would welcome an opportunity for their students to participate in a trial.

⁴ Most trials seek to estimate an intervention’s effect on the individuals (e.g., students) who receive it. To estimate this effect, the trial needs to compare outcomes for the individual students in the intervention group to outcomes for individual students in the control group – even those who move on to other classrooms or schools during the course of the study. Failure to do so may well undermine the equivalence of the intervention and control groups because those students who move may differ in important ways from those who stay (e.g., in motivation, educational achievement level, demographics).

Some trials, however, seek to estimate an intervention’s effect on a group (e.g., school) rather than on individuals within the group (e.g., students). In such a trial, you do not need to obtain access to individual student data in order to estimate the intervention’s effect and can instead rely on aggregate school data. For example, a trial of an intervention to reduce school violence may be interested in estimating the intervention’s effect on violent incidents at the school, even if the effect is achieved by getting violence-prone students to leave the school or by enrolling more students without a history of conduct problems. To estimate this effect, the trial needs to compare aggregate outcomes for schools in the intervention group to aggregate outcomes for schools in the control group, rather than outcomes for the individual students in the schools.

⁵ FERPA allows “organizations conducting studies for, or on behalf of, educational agencies or institutions for the purpose of developing, validating, or administering predictive tests, administering student aid programs, and improving instruction” to gain access to individual student educational records, including test scores, without the consent of their parents “if such studies are conducted in such a manner as will not permit the personal identification of students and their parents by persons other than representatives of such organizations and such information will be destroyed when no longer needed for the purpose for which it is conducted.” (Title 20 U.S.C. Section 1232g(b)(1)(F))

⁶ *How To Conduct Rigorous Evaluations of Math and Science Partnerships (MSP) Projects: A User-Friendly Guide For MSP Project Officials and Evaluators*, op. cit., no. 2, pp. 11-13.

⁷ Rouse, Cecilia Elena and Alan B. Krueger. “Putting Computerized Instruction to the Test: A Randomized Evaluation of a ‘Scientifically-based’ Reading Program.” *Economics of Education Review*, Vol. 23, Issue 4, August 2004, p. 323, <http://www.ers.princeton.edu/workingpapers/5ers.pdf>.

⁸ James-Burdumy, Susanne, Mark Dynarski, Mary Moore, John Deke, Wendy Mansfield, Carol Pistorino, and Elizabeth Warner, *When Schools Stay Open Late: The National Evaluation of the 21st Century Community Learning Centers Program Final Report*, prepared by Mathematica Policy Research for the U.S. Department of Education's Institute of Education Sciences April 2005, <http://www.mathematica-mpr.com/publications/pdfs/21stfinal.pdf>.

⁹ Schweinhart, Lawrence J., Helen V. Barnes, and David P. Weikart. *Significant Benefits: The High/Scope Perry Preschool Study Through Age 27* (High/Scope Press, 1993).

¹⁰ Decker, Paul T., Daniel P. Mayer, and Steven Glazerman, *The Effects of Teach for America on Students: Findings from a National Evaluation*, Mathematica Policy Research, June 9, 2004, <http://www.mathematica-mpr.com/publications/pdfs/teach.pdf>.

¹¹ Burghardt, John, Terry Johnson, Charles Metcalf, and Peter Z. Schochet, *National Job Corps Study: Findings Using Administrative Earnings Records Data, Final Report*, Mathematica Policy Research, October 2003, <http://www.mathematica-mpr.com/publications/pdfs/jobcorpsadmin.pdf>.

¹² The following is actually an example of a randomized controlled trial of a youth violence prevention program rather than an educational intervention; it nevertheless illustrates that official arrest records can be used to measure youth criminal behavior: Borduin, Charles, Barton J. Mann, Lynn T. Cone, Scott W. Henggeler, Bethany R. Fucci, David M. Blaske, and Robert A. Williams, "Multisystemic Treatment of Serious Juvenile Offenders: Long-Term Prevention of Criminality and Violence," *Journal of Consulting and Clinical Psychology*, vol. 63, no. 4, 1995, pp. 569-578.

¹³ Other differences in cost between a randomized controlled trial and a high-quality comparison-group study of similar size and structure include the following: (i) the comparison-group study often has higher costs of design and analysis, because of the need to use sophisticated statistical methods to ensure close matching of the intervention and comparison groups prior to the intervention, and to adjust for differences between the two groups in estimating the intervention's effects; and (ii) the randomized controlled trial often has the added expense of monitoring whether the students and teachers randomly assigned to the intervention or control group actually end up in their assigned group (and don't, for example, circumvent the process).

¹⁴ *Fiscal Year 2005 Application for New Grants for the Striving Readers Program*, U.S. Department of Education, Office of Elementary and Secondary Education, <http://www.ed.gov/programs/strivingreaders/2005-371a.pdf>.

¹⁵ U.S. Department of Education, "Scientifically-Based Evaluation Methods: Notice of Final Priority," op. cit., no. 1.

¹⁶ *How To Solicit Rigorous Evaluations of Mathematics and Science Partnerships (MSP) Projects: A User-Friendly Guide for MSP State Coordinators*, produced by the Coalition for Evidence-Based Policy in partnership with the National Opinion Research Center, May 2005, <http://www.ed.gov/programs/mathsci/issuebrief.doc>.

¹⁷ The study randomly assigned individual students to intervention or control classrooms, but did not randomly assign teachers to the classrooms. Thus, the study will produce a scientifically-valid estimate of the effect that a student's classroom assignment (to one of the three intervention classrooms or a control classroom) has on his or her reading achievement – which is valuable, policy-relevant information. But because teachers were not randomly assigned, one cannot rule out the possibility that any observed effect is caused partly by the classroom teacher rather than the reading intervention, because it's possible that teachers in the intervention and control classrooms differ systematically in motivation or other characteristics that affect their students' reading outcomes. Thus, if the study shows that assignment to one of the intervention classrooms produces an effect, it will be important to confirm that result in a study that also randomly assigns teachers to intervention and control conditions. Randomly assigning teachers would not add significantly to the trial's cost assuming the researchers have the strong support of the school district leadership for such a step.

¹⁸ Nagda, Biren A., Sandra R. Gregerman, John Jonides, William von Hippel, and Jennifer S. Lerner, "Undergraduate Student-Faculty Research Partnerships Affect Student Retention," *The Review of Higher Education*, vol. 22, no. 1, Fall 1998, pp. 55-72.

¹⁹ One modest flaw in the study was that it did not collect and analyze outcome data for the 54 students assigned to the intervention group who did not actually participate in a research partnership (in violation of the "intention to treat" concept). The trial did obtain complete outcome data for all other sample members (1,280 students), and presumably could have obtained complete outcome data for these 54 as well, at little or no additional cost.

²⁰ *How To Solicit Rigorous Evaluations of Mathematics and Science Partnerships (MSP) Projects: A User-Friendly Guide for MSP State Coordinators*, op. cit., no. 16.

²¹ *How To Conduct Rigorous Evaluations of Math and Science Partnerships (MSP) Projects: A User-Friendly Guide For MSP Project Officials and Evaluators*, op. cit., no 2.

²² *Key Items to Get Right When Conducting a Randomized Controlled Trial In Education*, produced by the Coalition for Evidence-Based Policy in partnership with the What Works Clearinghouse, December 2005, http://www.whatworkshelpdesk.ed.gov/guide_RCT.pdf.