



October 05, 2011

### [Gold-Standard Program Evaluations, on a Shoestring Budget](#)

*Note: This is a guest post by Jon Baron, President of the Coalition for Evidence-Based Policy, and Chairman of the National Board for Education Sciences*

In today's tough economic climate, quality evaluations of education reforms - to determine which are truly effective in improving student achievement, graduation rates, and other key outcomes - are especially important. They enable us to focus our limited resources on strategies that have been proven to work.

Well-conducted randomized controlled trials are generally recognized as the most reliable method (the "gold standard") for evaluating a program's effectiveness. However, widespread misconceptions about what such studies involve - including their cost - have often limited their use by education officials.

In plain language: Randomized controlled trials in education are studies that randomly assign a sample of students, teachers, or schools to a group that participates in the program ("the program group") or to a group that does not ("the control group"). With a sufficiently large sample, this process helps ensure that the two groups are equivalent, so that any difference in their outcomes over time - such as student achievement - can be attributed to the program, and not to other factors.

Such studies are often perceived as being too costly and administratively burdensome to be practical in most educational settings. In fact, however, it is often possible to conduct such a study at low cost and burden if the study can measure outcomes using state test scores or other administrative data that are already collected for other purposes. Costs are reduced by eliminating what is typically the study's most labor-intensive and costly component: locating the individual sample members at various points in time after program completion, and administering tests or interviews to obtain their outcome data. In some cases, the only remaining cost is the researcher's time to analyze the data.

For example, the following are two recent randomized trials that were conducted at low cost, yet produced findings of policy and practical importance:

- [Roland Fryer, recent winner of the MacArthur "Genius" Award, conducted an evaluation of New York City's \\$75 million Teacher Incentive Program](#) in which 396 of the city's lowest-performing public schools

were randomly assigned to an incentive group, which could receive an annual bonus of up to \$3000 per teacher if the school increased student achievement and other key outcomes, or a control group. Three years after random assignment, the study found that the incentives had no effect on student achievement, attendance, graduation rates, behavior, GPA, or other outcomes. Based in part on these results, the city recently ended the program, freeing up resources for other efforts to improve student outcomes.

The study's cost: Approximately \$50,000. The low cost was possible because the study measured all outcomes using state test scores and other administrative records already collected for other purposes.

•[Eric Bettinger and Rachel Baker conducted an evaluation of InsideTrack college coaching](#) - a widely-implemented mentoring program for college students designed to prevent them from dropping out of school. This was a well-conducted trial, which randomized more than 13,000 students at eight colleges. The study found that the program produced a 14 percent increase in college persistence for at least two years, and a 13 percent increase in likelihood of graduating college.

The study's cost: Less than \$20,000. The low cost was possible because the study measured its key outcomes using administrative data that the colleges already collected for other purposes - i.e., their enrollment and graduation records - rather than by collecting new data through individual surveys.

In recent years, federal and state policy, as well as improvements in information technology, have greatly increased the availability of high-quality administrative data on student achievement and other key educational outcomes. Thus, it has become more feasible than ever before to conduct gold-standard randomized evaluations on a shoestring budget. Equipped with reliable evidence, education officials can have much greater confidence that their spending decisions will produce important improvements in student outcomes.

-Jon Baron

[\*The Coalition for Evidence-Based Policy\*](#) is a nonprofit, nonpartisan organization whose mission is to increase government effectiveness through the use of rigorous evidence about "what works."