

Making Sense of Charter School Studies | A Reporter's Guide

By Lydia Rainey

Updated Spring 2015

Evaluating education research, and studies of charter schools in particular, is no simple assignment—either for reporters or education analysts. The pressure to release new findings and the digital availability of discussion and conference papers means that more and more research skips the traditional peer review process before being released to the public. This new role for reporters as unofficial arbiters of charter school studies would be difficult enough considering the complicated and often inadequate methods used in education research. But journalists must also consider that research is sometimes politically motivated, and especially so when it comes to controversial initiatives like charter schools.

This guide is designed to aid reporters in evaluating the quality of charter school achievement studies. It can also be used to interpret other charter school or general education research. Not surprisingly, a blue-ribbon Charter School Achievement Consensus Panel report found that no single research method or approach to assessing charter schools is problem free.¹ This guide draws from that panel's report, which provides several useful suggestions for education reporters. For example:

- Paradoxically, the studies that most accurately capture what is going on in their respective samples may not reveal much at all about what might be occurring in other schools outside those samples.
- The results of studies of one kind of charter school cannot and should not be generalized to all charter schools.
- Better studies can be done, but not without much better data than is currently collected by most states.

Existing research provides important findings about how students are doing in charter schools. For example, researchers have established that the quality of charter schools varies enormously, as does their impact on academic achievement. But it is also important to understand the limitations of each study. Researchers have yet to show if there is a consistent pattern of school characteristics among charter schools that boost or fail to raise academic achievement.

“Critical consumer” questions:

- Does the report compare apples to apples?
- How large and representative is the study's sample and comparison group?
- Does the report look at test scores from one or more than one year?
- Does the author overstate the significance of the findings?

1. Charter School Achievement Consensus Panel, *Key Issues in Studying Charter Schools and Achievement: A Review and Suggestions for National Guidelines* (Seattle, WA: Center on Reinventing Public Education, 2006).

“Critical consumer” questions to ask when evaluating a new study

A good place to start when covering a new study is to answer the “critical consumer” questions. Those questions are presented here in the context of charter school research. Commonly used study methodologies and their limitations are summarized in Table 1.

Does the report compare apples to apples?

A study’s design affects the reliability of its findings, and one of the most important design factors is the comparison group. The type of experimental research that determines the effects of a new drug requires that subjects be randomly assigned to either a group that receives the treatment or a control group. Indeed, many types of policy intervention, including housing vouchers, welfare-to-work programs, and initiatives to reduce recidivism among convicts have long been tested through randomized experiments, the “gold standard” of program evaluation. The Food and Drug Administration requires peer-reviewed, randomized, controlled trials before a medical device or drug can be marketed, and some 10,000 clinical research articles based on such experiments are published annually.

Unfortunately, education researchers rarely assess the impact of pedagogical shifts through the classic scientific experimental method. In its present incarnation, the K–12 education system generally does not allow students to be randomly assigned to either charter or traditional public schools. As a result, the next best alternative to traditional random assignment has been to create a comparison group composed of students who signed up for the admission lotteries at oversubscribed charter schools. The lottery provides a ready control group since students who were randomly denied admission to the charter school are likely to have the same kind of motivation and family support as students who were picked in the lottery. Properly done, this kind of study design approaches an apples-to-apples comparison.

In studies where students are not randomly assigned to charter or traditional public schools by lottery, researchers must compare test scores of students already in charter and traditional public schools. The danger with this design is that the charter students might differ in important but unmeasured ways from the traditional public school students to whom they are being compared. The students may have weaker skills or more involved parents, they may come from poorer families, or be recent immigrants just learning English. Any differences in outcomes could thus be due to differences in the children served, not due to the quality of instruction provided.

Researchers can turn this kind of study into an apples-to-apples comparison by controlling for some of these demographic differences. This method is commonly referred to as a “quasi-experiment.” Some important controls to look for include parent income (typically determined by free/reduced-price lunch status), race/ethnicity, urban/rural, grade level, number of years the school has been open, previous test scores, and prior education experience.

How large and representative is the study’s sample and comparison group?

Even when a study of charter school achievement employs an apples-to-apples comparison (or something close to it), the size and representativeness of the sample still matter. One of the most famous studies in all of education research—and one of the few to use random assignment—looked at the impact of the Perry Preschool Project in Ypsilanti, Michigan. The Perry Preschool study looked at children who enrolled in the intensive preschool program between 1962 and 1967. It had treatment and control groups that contained 58 and 65 disadvantaged black children, respectively. The evaluations of the Perry Preschool Project documented a slew of impressive benefits. Test scores for students in the treatment group were consistently and statistically significantly higher through age 14, and students had higher grades and were more likely to graduate from high school. They were also less likely to spend time in special education. Even after they finished school, the former Perry Preschool students were less likely to be on welfare or commit crimes. However, Head Start, the nation’s chief program for economically disadvantaged preschoolers, has not had the same impact on student achievement. A number of researchers have concluded that Head Start was unable to replicate the Perry Preschool initiative, both because Head Start provided a less intensive preschool program and because the comparatively small treatment group in the Perry Preschool program may have differed in important ways from the typical Head Start preschool population.²

2. W. Steven Barnett, “Benefit-Cost Analysis of the Perry Preschool Program and Its Policy Implications,” *Education Evaluation and Policy Analysis* 7, no. 4 (Winter 1985): 333-42.

Does the report look at test scores from one or more than one year?

One-year studies, or “snapshots,” have similar problems because these studies can also miss important information about students’ previous performance levels. With only one year of data, these studies cannot determine the “value-added” by charter schools since they cannot show whether students in charter or traditional public schools started at the same academic level. It is also possible that charter school students could have their achievement rise or fall in their first year at a charter school, only to have the trend reverse the next year—which would raise questions as to whether the charter school or some other force (e.g., a change in the student’s home life) was responsible for the second-year shift. At the very least, snapshot studies should control for previous student performance and academic history, or use an experimental model where students are randomly assigned to either charter or traditional public schools.

One way to avoid misleading results is to track student test score gains over two or more years. This is possible, however, only if researchers have multiple years of test scores both for the charter students and for the traditional school students to whom they are being compared.

Does the author overstate the significance of the findings?

Proving that a student’s attendance at one school versus another caused differences in test scores is very difficult; only high-quality studies that use extensive controls or compare the scores of winners and losers of charter school lotteries can determine if enrolling in a charter school improves a student’s test scores. Three research designs common in education cannot determine causality: correlation studies, “pre-post” comparisons, and case studies (see Table 1 for specifics). Beware of studies that claim charters outperform or underperform traditional schools when the studies rely on one of these methods.

It is also easy to confuse “statistical significance” with “effect size.” Statistical significance is a measure of confidence (how sure the researcher is that this measurement is accurate), while effect size (typically measured in standard deviations) is a measure of magnitude. Reporters should consider whether a finding is statistically significant and whether it has a relatively large effect size before assessing the impact of charter schools on student achievement. Although there is some debate, researchers generally consider an effect size of .1 of a standard deviation as slight, .2 or .3 as moderate, and .5 as large.³

Over-generalization is another common problem in research on student achievement in charter schools. Studies of charter schools often concentrate on schools with specific features (like those with waiting lists), or schools that serve students in a specific geographic area, (e.g., Kansas City). In these cases, the outcomes of a study most likely do not apply to charter schools as a group. High-quality studies that use lotteries to divide students into comparison groups fall into this category since they can only use schools with waiting lists. Results from these studies do not say much about the fate of charter schools that have no waiting lists since these schools might be lower in quality than schools with waiting lists. Similarly, because charter laws vary greatly from state to state, charter schools in a more permissive state like Arizona can be quite different from those in a less permissive state like Georgia.

In short, charter school evaluations typically entail a trade-off: the more precise the results, the less reliably they can be attributed to all schools. A study looking at a narrowly defined set of schools or students with similar attributes might do a good job of determining the impact of those schools on student performance, but it cannot answer the question of how charter schools as a group are doing compared to traditional public schools as a group. On the other hand, studies that compare scores from a broad mix of schools or students will yield more generalizable results, but they do a poor job of controlling for hidden differences among students that may significantly affect performance—a trait common in many state-mandated evaluations of charter schools.

3. For a general discussion of effect size, see Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. (Hillsdale, NJ: Lawrence Erlbaum Associates, 1988); as applied to education, see Eric Hanushek, “[Is the ‘Evidence-Based Approach’ a Good Guide to School Finance Policy?](#)” prepared for Washington Learns.

Things to consider when covering a below-average study

Other questions to consider:

- Who conducted the study and who paid for it?
- Did the report go through a quality-assurance process (e.g., independent peer review)?
- What do qualified scholars who are not personally close to the author think of the report?
- Does the report contain any unexpected or surprising findings, given the author's previous research?
- What is the author's professional reputation? What are the author's professional credentials and relevant experience?

Understanding and assessing a study's quality are the first steps to take when covering a study that employs a weak methodology. But journalists can enrich their coverage of a subpar study with additional reporting, including:

LOOK AT RESULTS FROM OTHER STUDIES

When covering a below-average study, it is worthwhile to consider the study alongside the findings of other reports. If the results from one of the studies are very different, this should raise a red flag, and more investigation into methods or data is warranted. If the results are similar, a pattern of performance may emerge.

CONSIDER THE LOCAL ANGLE

Charter schools vary tremendously by locality and by authorizer or district. Comparing your local context to that of the schools examined in the study might expose striking similarities—or differences—that could affect how applicable the results may be to your area's schools.

INCLUDE IMPORTANT CAVEATS

More often than not, authors are very up-front about what their research can and cannot conclude about charter school performance. To avoid overstating a study's findings, be sure to include the author's most important caveats in abbreviated form.

CONSIDER OTHER OUTCOMES

Digging deeper than test scores can yield useful information about the performance of charter and traditional public schools. Graduation or promotion rates, attendance, student transfers, participation in advanced coursework, school safety, or class size might complete the picture of how charter schools and traditional public schools are doing.

GET GOOD ADVICE

Sometimes it is hard to tell whether a study is reliable or newsworthy. In those instances, reporters can turn to experts for aid in interpreting a study's findings. Sources to contact might include:

- The Center on Reinventing Public Education (CRPE) at the University of Washington Bothell
Visit: crpe.org | email: crpe@uw.edu | call: 206-685-2214
- The [Center for Research on Education Outcomes](#) (CREDO) at Stanford University
- [Mathematica Policy Research](#)
- The [Department of Education Reform at the University of Arkansas](#)
- Research institutes with charter school expertise: the [Brookings Institution](#), [RAND](#), [Urban Institute](#), [National Center on School Choice](#) at Vanderbilt University
- Local universities: look for professors with expertise in methodology and knowledge of charter school research

TABLE 1. Problems With Common Methods for Measuring Achievement in Charter and Traditional Schools

PROBLEM OR CAUTION		CONSENSUS PANEL METHOD RATING
CASE STUDIES	<ul style="list-style-type: none"> o Cannot determine causation o Results not applicable to other charter schools 	Not rated
OBSERVATIONAL STUDIES	Compare school-wide average test scores for 1 year <ul style="list-style-type: none"> o Cannot determine causation o Overlooks most student characteristics o Only school-level variables used 	Poor
	Compare trends in school-wide average test scores for 2 or more years <ul style="list-style-type: none"> o Cannot determine causation o Overlooks most student characteristics o Only school-level variables used 	Fair
	Compare individual test scores for 1 year, using controls <ul style="list-style-type: none"> o Needs a large quantity of reliable, individual-level data 	Fair
	Compare trends in individual test scores for 2 or more years, using controls <ul style="list-style-type: none"> o Needs a large quantity of reliable, individual-level data 	Good
	Use individual test scores for students who switch to charter schools to determine if learning rates were higher or lower in traditional public schools <ul style="list-style-type: none"> o Can only use data for students who transfer to charter schools; these students may be different from students who started and stayed in charter schools 	Very good
“NATURAL” EXPERIMENTS	<ul style="list-style-type: none"> o Charter lotteries are designed for admissions, not for research, and may not be reliable o Results only applicable to over-enrolled schools that use lotteries 	Fair to excellent

For more information:

[Urban Charter School Study: Report on 41 Regions](#) (2015). Center for Research on Education Outcomes. credo.stanford.edu

[Special Education in Charter Schools: What We've Learned and What We Still Need to Know](#) (2014). Betheny Gross, Robin Lake. crpe.org

[Are Charter Schools Working? A Review of the Evidence](#) (2014). Patrick Denice. crpe.org

[A Meta-Analysis of the Literature on the Effect of Charter Schools on Student Achievement](#) (2014). Julian Betts, Y. Emily Tang. crpe.org

[Why the Gap? Special Education and New York City Charter Schools](#) (2013). Marcus Winters. crpe.org

[Inside Charter Schools: Unlocking Doors to Student Success](#) (2011). Betheny Gross. crpe.org

[Taking Measure of Charter Schools: Better Assessments, Better Policymaking, Better Schools](#) (2010). Julian Betts, Paul Hill. crpe.org

[Hopes, Fears, & Reality: A Balanced Look at American Charter Schools in 2006](#). crpe.org

See in particular:

Chapter 4. "Improving State and Local Assessments of Charter School Performance." Paul T. Hill, Julian Betts.

Chapter 6. "A Better Way: Measuring Charter School Success and Failure." Laura S. Hamilton, Brian Stecher.

[Making Sense of Charter Schools: Evidence from California](#) (2006). Ron Zimmer, Richard Buddin. rand.org

[Making Sense of Research: What's Good, What's Not, and How to Tell the Difference](#) (2003). Elaine K. McEwan, Patrick J. McEwan. corwin.com

About CRPE

Through research and policy analysis, CRPE seeks ways to make public education more effective, especially for America's disadvantaged students. We help redesign governance, oversight, and dynamic education delivery systems to make it possible for great educators to do their best work with students and to create a wide range of high-quality public school options for families.

Our work emphasizes evidence over posture and confronts hard truths. We search outside the traditional boundaries of public education to find pragmatic, equitable, and promising approaches to address the complex challenges facing public education. Our goal is to create new possibilities for the parents, educators, and public officials who strive to improve America's schools.

CRPE is a nonpartisan, self-sustaining organization affiliated with the University of Washington Bothell. Our work is funded through private philanthropic dollars, federal grants, and contracts.