



Charter School Leadership Council

## **Charter School Achievement: What We Know**

*Prepared by Bryan C. Hassel, Public Impact  
January 2005*

### **EXECUTIVE SUMMARY**

Recent reports by the American Federation of Teachers and the U.S. Department of Education have raised lingering questions about the performance of public charter schools. To answer these questions, and provide a full and fair picture of how charter schools are actually doing, the Charter School Leadership Council commissioned an extensive review of the available research on charter school achievement.

The following report summarizes and evaluates 38 comparative analyses of charter school and traditional public school performance, including a study-by-study look at central findings and methodological strengths and weaknesses. All of these studies fell into one of two categories: 1) snapshot pictures of one or more points in time; or 2) longer-term measures of change over time. In addition, they all met four basic criteria for rigor and relevance: they are recent (2000 or later), compare charter vs. traditional public school performance, use serious (though often flawed) analytical methods, and examine some significant segment of the charter sector.

Here are the key findings:

#### **STUDY QUALITY: The quality of available research varies widely.**

The stronger studies typically offer information about how much value charter schools are contributing to their students; study an adequate number of students and schools to be meaningful; use sound comparisons when assessing relative performance of traditional public schools vs. charter schools; and “disaggregate” analysis to show how well different kinds of students and schools are doing. Many of the studies fall short on one or more of these standards.

#### **SNAPSHOTS: The results are mixed and of limited use.**

Of the 38 studies, 17 look only at a snapshot of performance at one or more points in time. Nine show charter schools generally underperforming traditional public schools. The other eight show comparable, mixed or generally positive results for charter schools. These studies, however, fail to examine how much progress students and schools are making over time, and they are thus of limited use in drawing conclusions about the effectiveness of charter schools.

#### **CHANGE-OVER-TIME: The results, while far from conclusive, are encouraging.**

The other 21 studies make some attempt to look at change over time in student or school performance. Nine actually follow individual students over time, the ideal way to examine change. The rest use other methods, such as looking at changes in school-wide or grade-wide performance. Of these 21 studies:

- Nine find that overall gains in charter schools were larger than other public schools

*Studying Achievement in Charter Schools: What Do We Know?*

- Three find charter schools' gains higher in certain significant categories of schools, such as elementary schools, high schools, or schools serving at risk students
- Five find comparable gains in charter and traditional public schools
- Three find that charter schools' overall gains lagged behind

Seven studies examine whether individual charter schools improve their performance with age (e.g. after overcoming start-up challenges). Of these, five find that as charter schools mature, they improve. The other two find no significant differences between older and younger charter schools.

**RECOMMENDATIONS:** A number of conclusions about the state of charter school research – and how to improve it -- emerge from this review:

1. We need better research about how well students in charter schools are performing.
2. We need more and better research about why some charter schools perform so much better than other charter and non-charter schools.
3. We need much more attention focused on evaluating chartering as a policy. Knowing how well charter school students on average are performing does not answer the most important questions policymakers have about where to go with their charter policies.
4. Charter schooling represents an experiment worth continuing – and refining to improve quality further over time.

## Studying Achievement in Charter Schools: What Do We Know?

Prepared by Bryan C. Hassel  
Public Impact  
for the Charter School Leadership Council  
January 31, 2005

Accountability is a cornerstone of the charter school idea. In return for autonomy over key aspects of school operations, charter schools agree to be held accountable for results – to have their performance measured and to face the consequences if they fail to live up to expectations. As public schools, all charter schools participate in state assessment programs. They administer tests to all of their students, report the results, receive labels, and become subject to whatever sanctions arise from inadequate performance. The No Child Left Behind Act makes clear that charter schools, like all public schools, must make “Adequate Yearly Progress,” or else.

As a result, more and more data have become available over time about how well individual charter schools are doing when it comes to student achievement on standardized assessments.<sup>1</sup> Along with more data has come a raft of academic studies, state evaluations, and other efforts to answer the question “How well are charter schools doing?” The charter sector has been subject to an unprecedented level of scrutiny and transparency related to school performance. Just as individual schools are to be held accountable for results, the very idea of charter schools is being asked to prove itself, as well it should.

Reviewing all of these emerging studies of achievement in charter schools, however, is enough to make one’s head spin. As studies accumulate, each with its own unique methodological take on the basic question, contradictory findings proliferate. In fall 2004, for example, we were treated to two nationwide analyses of charter school achievement, one purporting to show that charter schools outperformed district schools, and one purporting to show the reverse. And these two were just the latest in an increasingly rapid volley of studies that show charter schools to be working well, or not.

At some level, mixed results are inevitable. The charter sector is host to a vast diversity of schools, utilizing all manner of educational and organizational approaches. The charter is but a shell, into which the operators place an instructional and management program. Asking about the quality of “charter schools” as a group is a bit like asking about the quality of “new restaurants” or “American cars” – any overall generalization will mask the great diversity within.

In short, there is really no simple answer to the question “how are charter schools doing?” At any point in time, some will be doing well, and some poorly. What we really want to know is how well *chartering*, as a policy, is working for a state. Is it producing new and better schools? How are the schools being chartered different from district schools? Are good charter schools expanding and being copied, while poor schools close or stagnate? Is the quality of chartering getting better over time? Is the presence of chartering inducing non-charter public schools to improve?<sup>2</sup>

In light of that set of questions, comparing the test scores of charter vs. district public schools cannot provide all of the answers. But it can shed some light on important issues of performance and progress by the students enrolled in this new form of public school. This report aims to help those interested in charter schools make sense of the dizzying array of studies about charter achievement in two ways:

- **Setting out some criteria** that observers can use to judge how sound a particular study's comparison of charter vs. district schools is; and
- **Summarizing and providing commentary** on many of the most recent comparative analyses of charter and district achievement. This summary includes an overview of some trends and patterns that appear across studies, as well as a study-by-study look at central findings and methodological strengths and weaknesses.

### **What Makes a “Good” Study of Achievement in Charter Schools?**

Research methodology is a highly complex field, and this report does not endeavor to touch on all the intricacies of method that might arise in a study of charter achievement. Instead, it outlines a set of high-level, essential criteria that ought to be applied to any study that seeks to compare charter and district achievement. When analyzing any particular study, it will be important to go beyond these basic criteria to look at specific methodological issues related to the particular study's approach.

- **Value-added Analysis.** For a given charter school, what we really want to know is whether students are better off for having attended it. The best way to find out is to examine the learning of individual students over time, seeking to determine how much “value” schools are “adding” to student learning. It is quite common in educational studies to compare two groups of schools or students based on a snapshot of their performance at a point in time. For example, many studies compare the percentage of charter school vs. district school students “making grade level.” Such comparisons can be very misleading because they fail to take into account changes in student performance over time. Consider two middle schools, both with 70% of their children at grade level at the end of an academic year. Suppose, though, that in one of the schools, only 15% of the children entered the school at grade level. In the second school, 80% did. These schools appear identical on the simple snapshot measure, but in fact their performance is dramatically different. Researchers can mitigate these challenges by comparing schools and students that are as similar as possible, but doing so is challenging (see bullet point about “sound comparisons,” below).

Ideally, then, a study would follow students over time in charter and district schools – ideally randomly assigned to attend them – and determine how much growth or gain students were experiencing. Such longitudinal analyses have their own methodological challenges, but if done well they shed more light on the central questions: how much are students learning in their schools. The longer the study can follow a given student, the better.

Many studies of school performance look at change over time, but are not able to follow individual students. For example, a study might examine the change in the percentage of a school's students meeting grade level standards from one year to the next. These approaches are inferior to following individual students; change from one year to the next in a school's overall performance reflects, in part, changes in composition of the student body – not just growth by individual students. While these studies are not true measures of “value added,” they provide more insight than simple snapshots.

- **Adequate Sample.** Does the study include a sufficient sample of charter schools/students to allow for generalization? Studies that are restricted to a small number of schools, a subset of grade levels, or a subset of the student population are less valuable than studies that seek to look at all charter school students in the relevant jurisdiction. Often, sampling is necessary due to lack of data or the immensity of data-gathering that would be required to look at all students and schools. In such cases, the question becomes whether the approach to sampling introduces bias into the results.
- **Sound Comparison.** Does the study compare charter school schools'/students' performance to that of a relevant group of district schools/students? Does it use appropriate controls or other methods to make the comparison valid? The “gold standard” in this regard is random experimental design, in which students are randomly assigned to a “treatment” group (admitted to the charter school) or a “control” group (not admitted). Such a design minimizes the chance that charter school attenders are somehow different from non-attenders in ways that influence achievement, such as their motivation, the level of challenge they bring to the school, or the engagement of their families.

There are several limits on the use of randomized studies in this area. First, such studies are expensive to run relative to approaches that rely on statistical analysis of publicly available data. Second, not all states require their charter schools to run lotteries. Third, even in states that require lotteries, only oversubscribed schools run them. Leaving undersubscribed schools out of an analysis because they do not have lotteries undermines the study's generalizability. Finally, students who are unsuccessful in a lottery may attend a district school, a private school, or another charter school. To the extent that the aim of the study is to compare charter and district performance, decisions by students to attend private and other charter schools muddy the waters considerably.

More commonly, researchers will use large databases of information about test scores, student demographics, and school characteristics to carry out comparisons. They will seek to compare charter school students' performance to that of students who are similar demographically and/or who are attending schools that charter school students would likely have attended in the absence of charter schools.

As the University of Washington's Paul Hill has noted, such comparisons are fraught with peril, and no study can provide the perfect comparison.<sup>3</sup> However, some will be better than others, working harder to compare charter school and students to the most relevant counterparts possible, in light of the available data. Sound comparative methods

are especially important for studies that provide just a snapshot of performance at a point in time.

- **Appropriate Disaggregation.** A related question is whether the study adequately differentiates between the performance of different kinds of schools and students. The charter sector is likely to be diverse and serve multiple student populations. Schools are chartered by different kinds of entities; have different levels of funding; take different approaches in their learning programs and organizational arrangements. Student populations differ by race, income, special needs, degree of initial academic challenge, and other factors. Since different kinds of schools and students may experience different success rates, it is vital for a study to analyze different groups separately where possible in addition to whatever kind of aggregate analysis is conducted.

One particularly important kind of disaggregation in research about charter schools is examining the performance of schools at different points in their life-cycles. Early-stage charter schools may have a different performance pattern than more mature charter schools, and studies should endeavor to sort out these differences.

### **The Studies: Emerging Patterns & Questions**

Dozens of studies looking at charter schools have been issued since the first charter school opened in 1992. This review examines 38 analyses that meet several criteria.<sup>4</sup> All of the included studies:

- **Are recent.** They were all released in or after the year 2001. The charter sector has grown and changed enormously since then, making earlier studies less relevant to today's charter school policy debates.
- **Include comparisons** of charter school students' achievement on standardized tests with that of students in district schools. As argued above, other kinds of studies are critical to understanding how well chartering is working. But the focus here is on test-based measures of charter school student achievement.
- **Use serious methods.** While the methodological quality of the studies varies greatly, all the research reviewed here represent reasonable attempts to analyze data about student achievement in charter and district schools. This is admittedly not a very high bar. While more higher quality research, such as the planned federally funded study that will use random assignment to gauge charter school value-added, are needed, this report gleans what it can from the research that exists, however imperfect. In addition, reviewing flawed studies helps highlight strengths and weaknesses in today's research and point the way to better evaluations in the future.
- **Examine a significant segment of the charter sector.** All but two of the studies included examined national data, multi-state data, or statewide data. One study (Metis Associates) was included because it examines all charter schools in Kansas City, MO, a city with one of the largest charter school "market shares" of any city. Another (Hoxby & Rockoff) looks at Chicago charter schools operated by a single management organization.

It was included because it is the only study to date that uses randomized experimental design to attribute value-added to charter schools.

Many other analyses were reviewed but ultimately excluded. Notably, the report does not include studies that provide only simple comparisons of a state's charter school and district test scores, without conducting statistical analysis to create a reasonable comparison. For example, the analysis does not include the January 2005 *Boston Globe* story showing that the percentage of students passing the state tests in urban charter schools higher than in district schools.<sup>5</sup>

The included studies differ from one another in many ways, but probably the most important is what kind of outcome they examine. It is possible to divide the studies into three groups as shown in Table 1:

- **Panel studies.** Nine of the studies follow individual students over time to see how their test scores change from year to year (left column of Table 1). These student “panel” studies are the most likely to identify the schools’ “value-added.”
- **Snapshot studies.** A second group of 17 studies (right column of Table 1), by contrast, only look at a snapshot of one or more points in time. While some of these studies attempt to control for student background characteristics, most are not as powerful as the panel studies in gauging how much value the schools are adding. Snapshots may reveal more about the starting levels of students entering the schools than they do about how much learning the schools are producing. The better the controls, the more likely snapshots are to shed light on value-added. For example, Hoxby & Rockoff’s experimental design makes it more meaningful because the two groups compared in the snapshot randomly assigned by lottery.
- **Other change studies.** Finally, 12 studies (middle column) look at change over time, but through some method other than following individual students over time (for example, looking at changes in average school-wide scores from year to year). While these studies contain more information about the effects of the schools than do most of the snapshots, they are not as powerful as the panel studies. Change over time in school-level averages could well be due to changes in what students attend schools rather than how much learning the schools produce.<sup>6</sup>

A more complete description of each study and its methods is in Table 2. The Appendix contains the full references, including web links where available.

**Table 1. Different Approaches to Comparing Charter and District Achievement**

Panel studies (following individual students over time)	Other change studies (e.g., looking at changes in school average results over time)	Snapshot studies (looking at results in one or more points in time)
Bifulco & Ladd (NC) Booker et al. (TX) Florida Department of Education (FL) Gronberg & Jansen (TX) Hanushek, Kain, & Rivkin (TX) Noblit & Dickson (NC) Sass (FL) Solmon & Goldschmidt (AZ) Zimmer et al. (CA)	Bettinger (MI) Greene, Forster, & Winters (multi) Loveless (multi) Metis Associates (KC, MO) Miron & Horn (CT) Miron & Nelson (MI) Miron, Nelson & Risley (PA) NY Board of Regents (NY) Raymond (CA) Rogosa (CA) Slovacek et al. (CA) Zimmer et al. (CA)	Bifulco & Ladd (NC) Colorado Department of Education (CO) Eberts & Hollenbeck (MI) Finnigan et al. (multi) Florida Department of Education (FL) Gronberg and Jansen (TX) Hoxby (national) Hoxby & Rockoff (Chicago) Legis. Office of Ed. Oversight (OH) Loveless (multi) Nelson, Rosenberg, & Van Meter (national) Nelson & Miron (IL) Noblit & Dickson (NC) Raymond (CA) U.S. Department of Ed. (national) Witte et al. (WI) Zimmer et al. (CA)

Looking at all of these studies as a group, here are some observations that emerge about student performance in charter schools:

**Diversity of outcomes.** Like their instructional and organizational designs, schools’ results vary widely from one to the next. Some charter schools score at or near the top of the heap in their cities and states; it is these schools that show the great promise of chartering as a mechanism for creating new, excellent schools. Others lie at the bottom of the heap; the existence of these schools suggests the need for stronger up-front chartering and ongoing accountability, but it is also just a natural phenomenon in any open system. Other charter schools fall within the “normal” range of performance for schools in their states. In this context, any attempt to discuss “the average charter school” is destined to mask this wide diversity. More illuminating would be an analysis of what proportion of charter schools fall in the top quintile of schools, the second quintile, and so on, and what happens to schools over time in these categories. Do the best schools expand and replicate? Do the worst get better or close?

**Evidence of added value.** Of the 38 studies reviewed, 17 look only at a snapshot at one or more points in time, and these return a mix of results. Nine show charter students generally underperforming district schools (Bifulco & Ladd; Eberts & Hollenbeck; Finnigan et al.; Gronberg & Jansen; Legislative Office of Education Oversight in Ohio; Loveless; Nelson et al.;

Noblit & Dickson; and U.S. Department of Education). However, most of these studies tell us little about whether charter schools are “adding value.”

Twenty-one made some attempt to look at change over time in student or school performance. Nine actually followed students over time. Another 12 examined growth by looking at changes in school-wide or grade-wide scores. What do these studies tell us about the gains students are making in charter schools?

- Nine studies find that *overall gains in charter schools were larger* (Booker et al.; Florida Dept. of Ed.; Greene et al.; Loveless; Metis; Miron & Horn; Miron, Nelson & Risley; NY Board of Regents; and Slovacek et al.).
- Three studies find charter schools’ gains higher than in district schools for *certain categories* of charter schools: at-risk schools in Texas (Gronberg & Jansen); elementary schools in Arizona (Solmon & Goldschmidt) and high schools in California (Raymond).
- Five studies find *comparable gains* (both Zimmer et al change analyses and Rogosa, all in CA; Hanushek et al. and Sass, once they account for the age of the charter school).
- Three find that charter schools’ gains *lagged behind* those in districts generally – the two NC studies (Bifulco & Ladd and Noblit & Corbett) and Miron and Nelson (Michigan).

So while the change-over-time picture is somewhat mixed, in general it is very encouraging about the gains students are making in charter schools. Only in North Carolina and Michigan do these analyses show that district students are out-gaining charter students overall. Most of the other studies show charter students or schools out-gaining their district peers, at least in some significant categories of schools.

**Schools gaining ground over time?** Some of the studies (Bifulco & Ladd; Miron & Horn; Miron et al.; Booker et al.; Hanushek et al.; Sass; Gronberg & Jansen) explicitly examine a more specific question: do charter schools get better as they age?<sup>7</sup> Do schools improve as they overcome initial start-up issues? Bifulco & Ladd finds that they do not: gains experienced by charter school students in NC are still lower even in more mature schools. Miron et al. finds only small differences based on first year of operation. The other five, though, find that as charter schools mature, they do better. Booker, Hanushek et al., Sass, and Gronberg & Jansen specifically find that rates of individual student growth in charter schools rise as schools get older.

## Recommendations

A number of recommendations emerge from this review. First, we need **better research on how well students in charter schools are performing**. Specifically, we need many more studies that track individual students over time, ideally in a randomized experimental setting, but at the very least using sophisticated methods to attribute value-added. These studies will provide the best information about how well individual schools are working for children.

We would also benefit from research that looks at other outcomes, such as dropout / completion rates in high schools, post-graduation outcomes like college persistence, attendance rates,

satisfaction levels, performance in subjects other than reading and math, and how well charter schools perform on aspects of their accountability plans other than standardized test scores. Also important is the question of “productivity.” Charter schools typically receive lower funding than districts do – what does this mean for the level of outcomes produced for each dollar provided to charter vs. district schools? These other measures are especially important if standardized test comparisons suggest that charter and district performance is similar, as they do in many of the examined studies.

Second, we need more and better research on **why some charter schools perform so much better** than other charter schools and non-charter schools. Moving beyond average performance, we see a significant number of charter schools that appear to perform far better than other schools. In all likelihood, many of the same factors that research has shown differentiate schools generally are at work in the charter sector. But there may be unique characteristics of leadership, organization, or program that are particularly valuable in the charter sector. Authorizers, policymakers, educators, and parents would benefit from understanding these.

Third, we need much **more research attention on evaluating chartering as a policy**. Knowing whether how well the average charter school student is performing does not answer most of the key questions that confront policymakers, like whether and how to expand the number of charter schools; how to change the way authorizing works; how to change charter schools’ funding and regulatory regimes; and how to stimulate the supply of more high-quality charter schools. Evaluating these dimensions is time-consuming and costly, but necessary if research is going to provide policymakers with useful, actionable information about how to make chartering work better as a strategy for improving public education.

The results to date **suggest important areas for action** by policymakers and practitioners. The existence of high quality charter schools and high growth rates for charter schools, at least in many states and studies, suggests that chartering holds promise as an approach to getting better schools. What we have is an experiment worth continuing – and refining. The existence of poor quality charter schools makes clear that we have more to learn about how to generate success with this policy. Together, these findings suggest a challenging agenda for policymakers and practitioners: to tap the full promise of chartering by continuing to expand the number of charter schools, while getting smarter about authorizing, accountability, and supply-creation. If we can do that, we can hope that a greater preponderance of tomorrow’s charter schools will match the outstanding quality of today’s best.