EXECUTIVE SUMMARY

Market-oriented education reforms’ rhetoric trumps reality

The impacts of test-based teacher evaluations, school closures, and increased charter-school access on student outcomes in Chicago, New York City, and Washington, D.C.

By Elaine Weiss and Don Long
ABOUT BBA

The Broader Bolder Approach to Education is a national campaign that acknowledges the impact of social and economic disadvantage on schools and students and proposes evidence-based policies to improve schools and remedy conditions that limit many children’s readiness to learn. BBA was launched in 2008 by the Economic Policy Institute, but is guided by outside co-chairs and an independent Advisory Council that shape policies distinct from those of EPI.

ABOUT THE AUTHORS

Elaine Weiss has been national coordinator of the Broader, Bolder Approach to Education (BBA) since 2011. Her public policy background has focused on early childhood education, specifically on policy mechanisms to help low-income families and children escape poverty. In her role with BBA, Weiss has coordinated and led events at the Economic Policy Institute and on Capitol Hill, developed research and advocacy materials, launched a social media presence for BBA, and worked with numerous allied organizations to increase the visibility of comprehensive strategies in education policy. Elaine has a J.D. from Harvard Law School and a Ph.D. in public policy from the George Washington University Trachtenberg School of Public Policy and Administration.

Don Long has been a consultant for BBA since November 2011. In that capacity, he has conducted research, drafted and edited documents, and identified much of the evidence for this report. Prior to working as an independent education consultant in Virginia, Long was director of the State Collaborative on Assessment and Student Standards (SCASS) and program manager at Pearson Educational Measurement in Austin, Texas. Long has a Master of Public Affairs degree from the LBJ School of Public Policy at the University of Texas, Austin.
Pressure from federal education policies such as Race to the Top and No Child Left Behind, bolstered by organized advocacy efforts, is making a popular set of market-oriented education “reforms” look more like the new status quo than real reform. Reformers assert that test-based teacher evaluation, increased school “choice” through expanded access to charter schools, and the closure of “failing” and underenrolled schools will boost falling student achievement and narrow longstanding race- and income-based achievement gaps. This report examines these assertions by assessing the impacts of these reforms in three large urban school districts: Washington, D.C., New York City, and Chicago. These districts were studied because all enjoy the benefit of mayoral control, produce reliable district-level test score data from the National Assessment of Educational Progress (NAEP), and were led by vocal reformers who implemented versions of this agenda.

KEY FINDINGS

The reforms deliver few benefits and in some cases harm the students they purport to help, while drawing attention and resources away from policies with real promise to address poverty-related barriers to school success:

- Test scores increased less, and achievement gaps grew more, in “reform” cities than in other urban districts.
- Reported successes for targeted students evaporated upon closer examination.
- Test-based accountability prompted churn that thinned the ranks of experienced teachers, but not necessarily bad teachers.
- School closures did not send students to better schools or save school districts money.
- Charter schools further disrupted the districts while providing mixed benefits, particularly for the highest-needs students.
- Emphasis on the widely touted market-oriented reforms drew attention and resources from initiatives with greater promise.
- The reforms missed a critical factor driving achievement gaps: the influence of poverty on academic performance. Real, sustained change requires strategies that are more realistic, patient, and multipronged.
In most large urban districts studied, test score gains among minority students narrowed race-based achievement gaps, and low-income students had gains comparable to their affluent peers. This contrasts with reform cities, where achievement gaps grew as poor and minority students’ scores fell further behind those of their peers.

Source: National Center for Education Statistics, Trial Urban District Reading Assessment, 2005 and 2011; National Assessment of Educational Progress scores for District of Columbia Public Schools provided by D.C. budget consultant Mary Levy in 2012.
Test scores increased less, and achievement gaps grew more, in “reform” cities than in other urban districts. Leaders promised that the reforms would raise test scores, especially those of minority and low-income students, and close race- and income-based achievement gaps. Analysis of the most reliable, comparable data—National Assessment of Educational Progress scores—shows that the rhetoric did not match the reality. While test scores increased and achievement gaps shrank in most large urban districts over the past decade, scores stagnated for low-income and minority students and/or achievement gaps widened in the reform cities.

✓ Between 2005 and 2011, in large, urban districts, Hispanic eighth-graders gained six points in reading (from 243 to 249), black eighth-graders gained 5 points (from 240 to 245), and white eighth-graders gained 3 points (from 270 to 273). In District of Columbia Public Schools, however, Hispanic eighth-graders’ scores fell 15 points (from 247 to 232), black eighth-graders’ scores fell 2 points (from 233 to 231), and white eighth-graders’ scores fell 13 points (from 303 to 290).

✓ New York City ranked second to last among 10 large, urban districts in NAEP test score gains from 2003 to 2011 (averaged across fourth and eighth grade reading and math). New York City students gained 4.3 points, half the urban district average gain of 8.8 points. Only Cleveland students had a smaller average gain (1 point).

✓ In Chicago Public Schools (CPS), white and Asian students made modest gains in reading between 2003 and 2009, but Hispanic students gained little and black students gained nothing, so achievement gaps between white and minority students grew at both the fourth and eighth grade levels, as measured by NAEP scores. Nationally, race-based achievement gaps in reading narrowed considerably among fourth-graders and remained about the same among eighth-graders.

“While test scores increased and achievement gaps shrank in most large urban districts over the past decade, scores stagnated for low-income and minority students and/or achievement gaps widened in the reform cities.”
Reformers in DC, NYC, and Chicago reported “success” in large test score gains and shrinking achievement gaps. When the data were recalibrated, broken down by subgroup, and compared with reliable numbers, however, the gains vanished and gaps grew.

NYC MAYOR BLOOMBERG CLAIMED TO HAVE CUT THE RACE-BASED ACHIEVEMENT GAP BY 50% FROM 2003 TO 2011

IN REALITY, THE GAP CLOSED BY 1%

Reported successes for targeted students evaporated upon closer examination. Reformers in all three cities claimed that they had boosted student achievement and closed achievement gaps. But when state test scores were recalibrated to make standards consistent, compared with NAEP scores, and disaggregated by race and income, gains vanished or turned out to have accrued only to white and high-income students.

✓ As Washington, D.C., schools chancellor, Michelle Rhee announced that all subgroups of students had improved their reading and math scores between 2007 and 2010, with low-income and minority high school students posting double-digit gains in “proficiency.” But those gains, based on an arbitrary DC Comprehensive Assessment System “proficiency” level, were illusory. NAEP scores showed minimal-to-no improvement for low-income and minority students, and some losses. Moreover, higher scores were due in most cases not to actual improvements for any group, but to an influx of wealthier students. For example, average fourth-grade NAEP reading scores rose from 198 to 201, or by 1.5 percent, from 2007 to 2011. But during that period, scores for white and Hispanic students fell by 3 points, and black students’ scores stagnated, so only new students who brought higher scores to the pool could account for the small overall gain.

✓ New York City Mayor Michael Bloomberg claimed to halve the white/Asian to black/Latino achievement gap in city schools from 2003 to 2011, but scores on state-administered tests, averaged across fourth and eighth grades in reading and math, show that the achievement gap had stagnated; it was 26.2 percentage points in 2003, versus 25.8 percentage points in 2011 (a 0.01 standard deviation change). Columbia University professor Aaron Pallas, who calculated the 1 percent reduction, noted, “The mayor has thus overstated the cut in the achievement gap by a factor of 50.”

✓ President Obama and Education Secretary Arne Duncan (when he was CPS CEO) have both cited large increases in elementary school reading “proficiency” of 29 percentage points—from 38 percent of students in 2001 to 67 percent in 2008. CPS used these figures in January 2009 brochures. When scores were adjusted for changes in tests and procedures, however, the percentage of elementary and middle-school students deemed proficient (“at or above grade level”) had grown by about 8 percentage points, while the percentage of proficient high school students had grown only a point and a half.
Using Test Scores to Evaluate Educators and Schools Led to Teacher Churn and Inexperience

Reforms in DC, NYC, and Chicago that used student test scores to evaluate, reward, and fire teachers and to target schools for closure delivered increased turnover and fewer experienced, qualified teachers, but no improvement in student achievement.

**After 2 Years of Reforms**

33% OF DCPS Teachers Left

**After 4 Years**

52% OF DCPS Teachers Left

Test-based accountability prompted churn that thinned the ranks of experienced teachers, but not necessarily bad teachers. Reformers said that using student test scores to evaluate teachers, and to reward and fire them based on those scores, would improve the quality of teachers in low-income schools. The report finds, rather, that narrow, unreliable metrics turned off great teachers, increased churn, and drained experience from teacher pools, with no boost to student achievement.

- District of Columbia Public Schools’ IMPACT system, which bases teacher evaluations (and dismissals) heavily on test scores, is associated with higher teacher turnover. The share of DCPS teachers leaving after one year increased from 15.3 percent in 2001–2007 (before IMPACT began in 2009) to 19.3 percent in 2008–2012; the share leaving after two years increased from 27.8 percent to 33.2 percent; the share leaving after three years increased from 37.5 percent to 42.7 percent; and after four years fully half (52.1 percent) of teachers left the system, up from 45.3 percent. Few teachers reach “experienced” status, generally considered at least five years and, by some experts, seven years or more.

- New York City spent $50 million from 2007 to 2010 on awards to teachers who substantially raised test scores in high-needs schools. In 2011, it ended the program after a RAND study confirmed “mounting evidence that all those bonuses weren’t having much of an effect.” The Schoolwide Performance Bonus Program, intended to “motivate educators to change their practices to ones better able to improve student achievement” failed to improve student achievement at any grade level, school progress report scores, or teachers’ reported attitudes and behaviors.

- CPS used test scores to close “failing” schools, forcing out many experienced teachers. CPS also laid off 1,300 teachers from 2001 to 2008, citing budget shortages. The district never provided teachers due process hearings, and in 2010, a group of teachers, including 749 who had tenure, won a discrimination suit. They had alleged that they had been replaced with less experienced, younger, whiter teachers. The judge held that CPS had violated their rights and ordered the district to work with the union to recall them, noting that most had not received unsatisfactory reviews.

“New York City spent $50 million from 2007 to 2010 on awards to teachers who substantially raised test scores in high-needs schools. In 2011, it ended the program after a RAND study confirmed ‘mounting evidence that all those bonuses weren’t having much of an effect.’”
SENDING STUDENTS FROM ONE ‘LOW-PERFORMING’ SCHOOL TO ANOTHER DID NOT IMPROVE ACHIEVEMENT

Reforms closed “failing” schools in DC, NYC, and Chicago, promising better student outcomes. But students stagnated or lost ground, as new schools were no better, and moves also meant instability and longer (and sometimes more dangerous) commutes.

**Only 6% of students whose schools were closed in Chicago moved to better schools that could support them.**

*Source: Marisa de la Torre and Julia Gwynne, When Schools Close: Effects on Displaced Students in Chicago Public Schools, Consortium on Chicago School Research, 2009.*
School closures did not send students to better schools or save school districts money. Reformers closed schools deemed “failing” so students could transfer to “better-performing” schools. But most students whose schools were closed went to schools that were no better, and the disruption (some students moved multiple times) was exacerbated by longer commutes and spikes in gang violence as established lines were crossed.

- DCPS’s initial reported cost of $9.7 million to close 23 underenrolled schools in 2008 grew to $39.5 million, with added moving expenses, demolitions, patrols, new transport costs, and others quadrupling the pricetag.¹⁴

- Michelle Rhee had also noted that students in the schools slated for closure were struggling, but the students went, on average, to schools with lower test scores and lower odds of making “adequate yearly progress” (AYP).¹⁵

- Between 2002 and early 2012, the NYC Department of Education closed 140 schools, all of which served the students with the highest needs.¹⁶ About 15 percent of those were large, comprehensive public high schools that were broken up into smaller, themed schools. Most of the students who would have attended the closed high schools were not admitted to the smaller schools but went to other large comprehensive high schools, “which consequently became academically overwhelmed, making them additional targets for closure.”¹⁷ Of 34 large Brooklyn, Manhattan, and Bronx high schools, 26 significantly increased enrollments—by 150 to more than 1,100 students—as other high schools were closed from 2002 to 2007. In 19 of these 26 schools, attendance declined; in 15, graduation rates declined, and in 14, both attendance and graduation rates declined.¹⁸

- Although Arne Duncan closed Chicago public schools deemed “underperforming” in order to move students to better schools, the closings had almost no effect on student achievement because almost all displaced elementary school students transferred from one low-performing school to another, according to a study of 18 schools closed between 2001 and 2006. Only the 6 percent who moved to better schools with greater resources had improved outcomes.¹⁹

“Most of the students who would have attended the closed [NYC] high schools were not admitted to the smaller schools but went to other large comprehensive high schools, ‘which consequently became academically overwhelmed, making them additional targets for closure.’”
Reformers in DC, NYC, and Chicago depicted charters schools as solutions to “failing” schools and “dropout factories.” City-level outcomes painting a much more complex picture—with some of the neediest students left out altogether—are mirrored by a national study of charter students’ performance relative to their regular school equivalents.

CHARTER STUDENTS

34% DID WORSE

49% STAYED THE SAME

17% DID BETTER

Charter schools further disrupted the districts while providing mixed benefits, particularly for the highest-needs students. Reformers say charters offer better options and outcomes for students in “failing” public schools. But charter outcomes in these cities and across the country are uneven. Charters serve fewer of the highest-need students and can disrupt districts logistically and financially. High-performing charters may also spend more per student.

Rhee transferred operations of two DCPS high schools and one elementary school to outside charter organizers. None of the three schools improved their performance under new charter management, and both high schools have since been reconstituted again for very poor performance. Overall, charters in Washington, D.C., seem to slightly outperform regular neighborhood schools, but they serve fewer high-needs students.

Two widely cited reports found most students benefited significantly by attending NYC charter schools rather than regular district schools, but two other reports questioned those reports’ methodologies and findings. It is clear, however, that New York City charters benefit from more funding per student and better facilities in co-located spaces. While they serve more minority and low-income students, they serve fewer students who are special needs, very poor, or English language learners (ELL), and these high-needs students are costlier to serve. Comparing charters with nearby public schools illustrates stark differences. At Samuel Stern public school, where 86 percent of students qualify for free lunch and 19 percent are ELL, per-pupil spending is $12,476. At nearby Harlem Day charter school, 62 percent of students qualify for free lunch, and there are no ELL students, but per-pupil spending is $19,632.

The Chicago Public School system uses its own “value-added” metric to measure school performance, with schools scoring lower on the distribution identified for closure. By this measure, if students in the types of schools most likely to be closed moved to charters, they would move to lower-performing schools. Specifically, students who moved from high-poverty regular public schools at the 47th percentile in performance would go to charter schools at the 40th percentile, and those moving from intensely segregated schools at the 43rd percentile would end up in charters at the 33rd percentile. Random-lottery enrollment schools, which, unlike charter schools, do not “select out” students via a challenging application process, outperform their demographically comparable charter counterparts: Students who moved to charters would drop from the 52nd percentile to the 40th percentile.
BROADER STRATEGIES OFFER MORE PROMISE THAN MARKET-BASED REFORMS

While reformers continue to advance market-oriented policies with few benefits, more holistic strategies with real promise in the three cities have failed to receive the attention or funding needed to have a real impact.

100 new small NYC schools created in 2003–2005 focused on ensuring strong, consistent student-teacher relationships; leveraging community partners for extra staff, coaching, and resources; and providing hands-on learning experiences, such as internships at law firms and seeding oyster beds. These schools reportedly increased the share of ninth-grade students on track to graduate and high school students’ college readiness.

To attract more high-quality teachers to Chicago Public Schools, then-CEO Arne Duncan identified the strongest teacher-preparation programs and encouraged CPS to hire from them, moved recruitment dates up, established job fairs to boost recruiting ability, and offered new teachers higher starting salaries. This improved teacher quality and reduced inequities across districts.

Duncan worked to improve low-income and minority students’ college readiness by increasing their access to AP courses, putting college counselors in low-income high schools to help students choose courses and schools and match their goals with skills, and holding principals accountable for ensuring that students applied for financial aid, which nearly doubled in one year. Budget cuts have since removed counselors from almost all schools.

Michelle Rhee expanded DCPS’s full-day voluntary prekindergarten program to serve 3- and 4-year-olds at all income levels, and the district adopted a holistic curriculum designed to nurture all domains of children’s development. Though third-graders who had participated had higher test scores than their nonparticipating peers, pre-K is not even a component of the agenda on which Rhee’s advocacy group, StudentsFirst, grades every state’s education system.
Emphasis on the widely touted market-oriented reforms drew attention and resources from initiatives with greater promise. Less-publicized strategies for boosting student achievement were piloted in these cities but not widely replicated or expanded to scale because leaders and funders focused on the market-oriented reforms. These promising but overlooked reforms are more multifaceted and holistic than reforms that seek quick fixes and rely on narrow, unreliable metrics.

Real, sustained change requires strategies that are more realistic, patient, and multipronged. In each city, the initiatives showing more promise than the touted reforms demonstrate that achievement gaps can only be closed when the opportunity gaps driving them are addressed. The hands-on experiences and consistent, intensive teacher-student relationships of New York City’s small schools must replace reform’s test preparation, novice teachers, and churn. Heavy reliance on college- and career-readiness test metrics should give way to CPS-style college- and career-readiness supports: helping students choose courses and schools, access AP courses, and match their skills with career goals; and holding schools accountable for scholarship applications. DCPS’s high-quality prekindergarten program, which is designed to nurture all aspects of children’s development, should serve as a model for all cities and students, not be sidelined in “reform” agendas.

“These promising but overlooked reforms are more multifaceted and holistic than those that seek quick fixes and rely on narrow, unreliable metrics.”
Lack of consistent physical and mental health care is a major driver of the opportunity gaps associated with growing up in poverty. Low-income children miss many more days of school due to preventable illnesses, relative to their wealthier peers—a reality largely dismissed in reform agendas.

The reforms missed a critical factor in achievement gaps: the influence of poverty on academic performance. In all three cities, a narrow focus on market-oriented policies diverted attention from the need to address socioeconomic factors that impede learning. In 2010, student eligibility rates for free- and reduced-price meals were 67 percent in Washington, D.C., 72 percent in New York City, and 77 percent in Chicago. Failing to provide supports that alleviate impediments to learning posed by poverty ensures continued low student test scores and graduation rates, and large gaps between average scores of white and affluent students and scores of minority and low-income students.

Districts that recognize the impact of poverty and address it head-on find the greatest success. Though it is higher-income, Montgomery County, Md., serves a student body that is as ethnically diverse as any of these urban districts, and has a large and growing share of low-income students. In contrast to the reformers, however, Montgomery County Public Schools (MCPS) Superintendent Joshua Starr staunchly opposes using test scores to evaluate teachers, employing a peer-assisted review system that focuses on teacher support, development, and collaboration. MCPS has no charter schools. Rather, it channels extra resources, including targeted professional development for qualified teachers, smaller classrooms, and intensive literacy, to the neediest schools. It has developed a holistic, creative curriculum to nurture in-depth, critical thinking. This includes art, music, and physical education teachers in every school. MCPS also leverages the county’s mixed-use housing policies to integrate schools. Finally, it employs high-quality prekindergarten, health clinics, and afterschool enrichment to further close income-based opportunity gaps. As Starr highlights, all of this has produced some of the highest test scores among minority and low-income students of any district, smaller and shrinking achievement gaps, and high school graduation and college attendance rates that are the envy of the country.

Every school district has unique needs and resources. But providing all students with the enriching experiences that already help high-income students thrive would represent a big step forward, and away from narrow reforms that miss the mark.
Endnotes


3. The 10 districts studied were the 10 urban districts that participated in the Trial Urban District Assessment (TUDA) in 2003. By 2011, there were 21 TUDA districts.


17. Urban Youth Collaborative, No Closer to College: NYC High School Students Call for Real School Transformation, Not School Closings,


27. New America Foundation, Federal Education Budget Project, “PreK-12” data for CPS, DCPS, and NYCPS 2009–2010; febp.newamerica.net/k12


Broader, BOLDER Approach to Education