The Case Against Michelle Rhee

How persuasive is it?

PHOTO ILLUSTRATION/JOHN WEBER

By PAUL E. PETERSON

Recently, two separate studies—one by Alan Ginsburg, a former director of Policy and Program Studies in the U.S. Department of Education, the other by a committee constituted by the National Research Council (NRC)—have sought to discredit the work of Michelle Rhee, former chancellor of schools for the District of Columbia.

According to Ginsburg, Rhee was no more effective—probably even less effective—than her predecessors. Not surprisingly, his argument was quickly picked up by American Federation of Teachers president Randi Weingarten. In a *Wall Street Journal* interview, she asserts that Michelle Rhee "had a record that is actually no better than the previous two chancellors." In a blog post dated March 29, 2011, Diane Ravitch makes the same point: "The gains under Rhee were no greater than the gains registered under her predecessor Clifford Janey, who did not use Rhee's high-powered tactics, such as firing massive numbers of teachers." Yet the evidence Ginsburg musters to support such claims falls well short of its mark.

In the second study, the NRC committee does not deny that student performance in the District of Columbia improved under Michelle Rhee's chancellorship between 2007 and 2010, but it says there is no scientific evidence that proves the work of the chancellor is responsible for those gains. "The problem was the [test score] changes that seem to be going in the right direction can't be attributed to the specific changes in the system," the study committee's co-chair Robert M. Hauser told an *Education Week* reporter. While it is certainly true that one cannot, in the absence of experimental evidence, establish a connection between policy changes and test-score outcomes, Hauser added a carefully worded slap at Rhee: "All districts should be cautious about generalizing from the kind of aggregate overview data that have been used to suggest successes of changes made in the district to date." The reporter is then informed that "students' NAEP scores started to improve before the overhaul law passed, as noted in a report last month by Alan Ginsburg."

The NRC study bears the more prestigious imprimatur, but it is the Ginsburg study that is most likely to be cited in future discussions of merit pay, teacher tenure, and the like. So our fact-checking of the two studies begins with his contribution to the discussion.

The Ginsburg Report

Alan Ginsburg, though now retired, was until very recently the ultimate Washington insider. For more than a generation he was known as the Department of Education's data-collection guru, the person inside the bureaucracy who understood best what information to collect and how to collect it. So it is of considerable interest that Ginsburg has now chosen to give aid and comfort to Weingarten and other union leaders by leveling a hard-core attack on "The Rhee DC Record."

To an *Education Week* reporter, Ginsburg insisted that his critique of "The Rhee DC Record" is not "intended to be anti-Rhee." He is reported as saying that he acted only because "he believes they [his findings] should serve as a check on a policy of mass dismissals of teachers as a way to improve districts. 'For me, it's the much larger question in this country of building a large teaching force.'" It is nonetheless quite disconcerting that he—and those who rely on his work—say that she was engaged in "large-scale firing" and "mass dismissals" when in fact she released in 2010 just 241 teachers for low performance.

Ginsburg excludes any and all information coming from the D.C. exams, known as the Comprehensive Assessment System (CAS), required by the federal law known as No Child Left Behind. He explains that decision on the grounds that "performance levels for 2006 and afterwards are not comparable with those from prior years." But that does not preclude a comparison of Rhee's record for the years beginning in 2007 with the situation in the year before she arrived. Had Ginsburg taken a look at that information, he would have found an acceleration of the gains in the percentage of students deemed proficient. Before Rhee's tenure, or between 2006 and 2007, the percentage increase in proficiency was about 1 percentage point in reading and 4 percentage points in math. But between 2007 and 2010, the gains in percent proficient were 9 percentage points in reading and 15 percentage points in math.

District Performance on National Assessment of Educational Progress

Although these gains are impressive, a *USA Today* investigative team has expressed concerns that, at least in some schools, those test-score results might have been improperly inflated. No conclusive evidence of cheating has yet been established, but it may well be prudent to focus, as Ginsburg does, on the performance of D.C. students on the National Assessment of Educational Progress (NAEP), commonly known as the nation's report card. That is a low-stakes test taken only by a representative sample of students, none of whom answer all the questions and for whom no results are reported by student, teacher, or school. As the NAEP is not part of any accountability system, incentives to cheat on the test are minimal, and no allegations of cheating have been made.

Had D.C. students gained as much every year between 2000 and 2009 as when Rhee was in charge, the gap in 4th-grade math between D.C. and students nationwide would have narrowed from 34 to 7 points.

At first glance, Ginsburg does not seem to have much of a case against Rhee. D.C. scores on the NAEP shifted upward during the first two years Rhee was in office. In both 4th-grade math and reading they jumped by 6 points, and in 8th-grade math they leaped by 7 points, though they slipped a point in 8th-grade reading (see Figure 1).

But Ginsburg says those gains are actually no greater than the ones students had been making in prior years, when superintendents Paul Vance and Clifford Janey were in charge. He reports, "With respect to the distribution of DC's total gains in NAEP scores over grades 4 and 8 between 2000-09, Vance accounted for a 46% share of the total gain, Janey 30% and Rhee 24%."

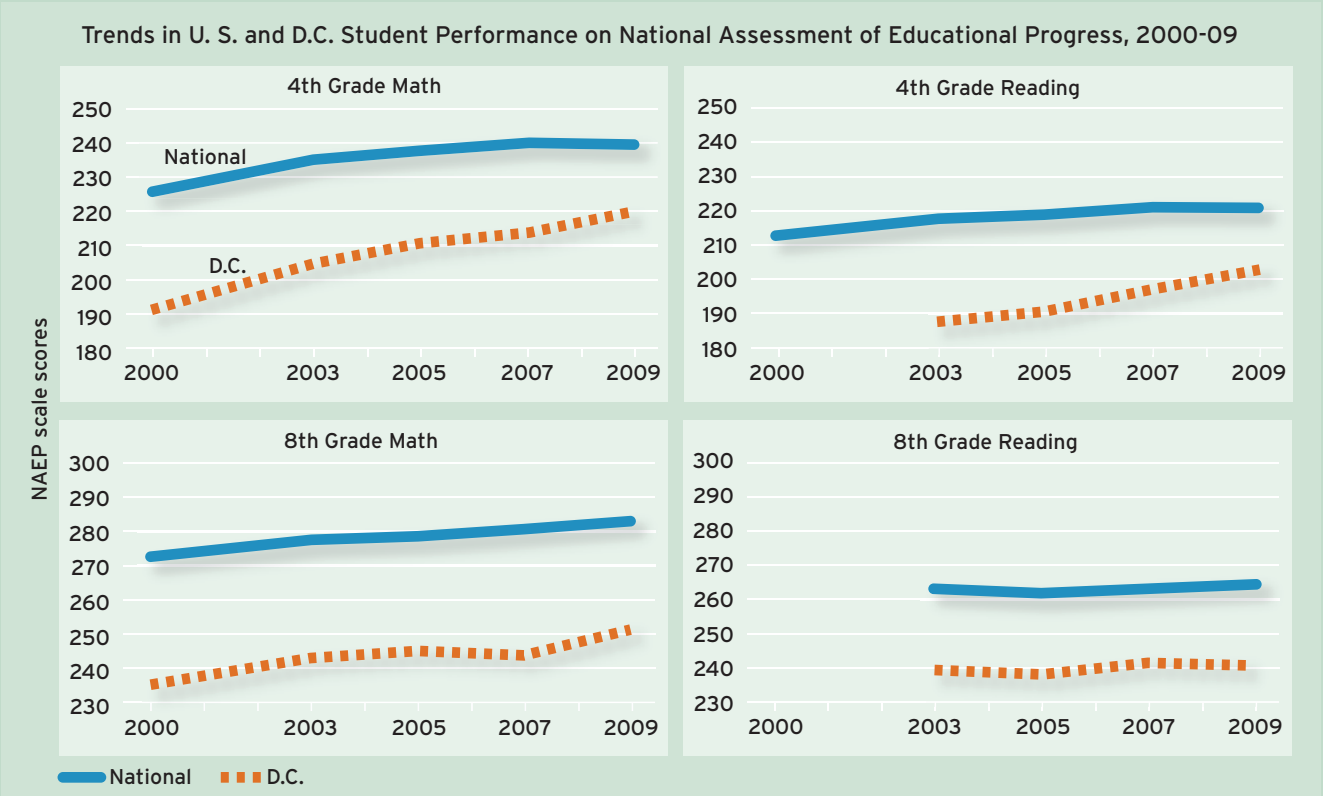
Though headline-grabbing numbers, they are quite misleading. Between 2000 and 2009, Rhee was in office for only two years, while Vance was in office for three, and Janey for four. If gains were rising at the same rate over the nine-year period, then each superintendent should account for 11.1 percent of the gains for each year in office: Vance 33.3%, Janey 44.4%, and Rhee 22.2 %. So based on Ginsburg's own calculations, Rhee outperformed her immediate predecessor.

More significantly, Ginsburg ignores the fact that the D.C. NAEP sample in 2009 did not include students attending charter schools not authorized by the district, while in 2007 all charter school students were included. Because charter schools outside district control were outperforming district schools, the latter appeared to be doing better in 2007 than they actually were. NAEP corrected its data-collection procedures in 2009, but, except for 8th-grade math, it failed to provide the data that allow for an apple-to-apple comparison between 2007 and 2009. For 8th-grade math, NAEP explains that had NAEP followed the same policy in 2007 that it adopted in 2009, 8th-grade math scores under Rhee would have increased by 7 points, a statistically significant gain, not just the 3 points that are officially reported.

Similar underreporting of gains may have occurred on the 4th- and 8th-grade reading exams and the 4th-grade math tests, but NAEP unfortunately does not tell us how large they were. Its report only says that giving us that information would not alter the findings as to the statistical significance of gains. So in the analysis below, I provide the corrected results for 8th-grade math, but I cannot provide corrected results for the other exams.

Upward Trend (Figure 1)

D.C. scores on the NAEP shifted upward on three of four tests after Rhee took office in 2007.



SOURCE: National Assessment of Educational Progress, National Center for Education Statistics, U.S. Department of Education

Closing the Gap between District and National Performance

Most importantly, Ginsburg did not adjust for national trends in student performance occurring between 2000 and 2009. Unless one adjusts for national trends, one does not know whether gains in the district are due to district-specific events or to some larger developments in the nation, such as changes in the economy, or the waning effectiveness of No Child Left Behind, or permutations in the design and administration of the NAEP examination, or some other large-scale factor.

The most straightforward way of adjusting for national trends is to look at the extent to which D.C. closed the gap between its students' performances and those of students nationwide. Once that adjustment is made, it can be shown that Rhee did considerably better at that task than did her predecessors (see Figure 2). For example, during the Rhee years, 4th-grade students, in both reading and math, gained an average of 3 points each year relative to the scores earned by students nationwide, a gain twice that of Rhee's predecessors.

These numbers seem small, but they add up. In 2000, the gap between D.C. and the nation in 4th-grade math was 34 points. Had students gained as much every year between 2000 and 2009 as they did during the Rhee era, that gap would in 2009 have been just 7 points. Three more years of Rhee-like progress and the gap is closed. In 8th-grade math, the gap in 2000 was 38 points. Had Rhee-like progress been made over the next nine years, the gap would in 2009 have been just 14 points, with near closure in 2012. In 4th-grade reading, the gap was 30 points in 2003 (scores are unavailable for 2000); if Rhee-like gains had taken place over the next six years, the gap in 2009 would have been cut in half.

None of this proves that Rhee could sustain the gains observed over a two-year period. That is too short a time to draw conclusions about a leader based on NAEP results alone. Also, no improvement in 8th-grade reading is detected. The overall results do, however, cast doubt on Ginsburg's claim that Rhee did no better than her predecessors.

But perhaps the other report, the one issued by a committee of the prestigious National Research Council, makes a more persuasive case that Rhee's performance is less than it seems.

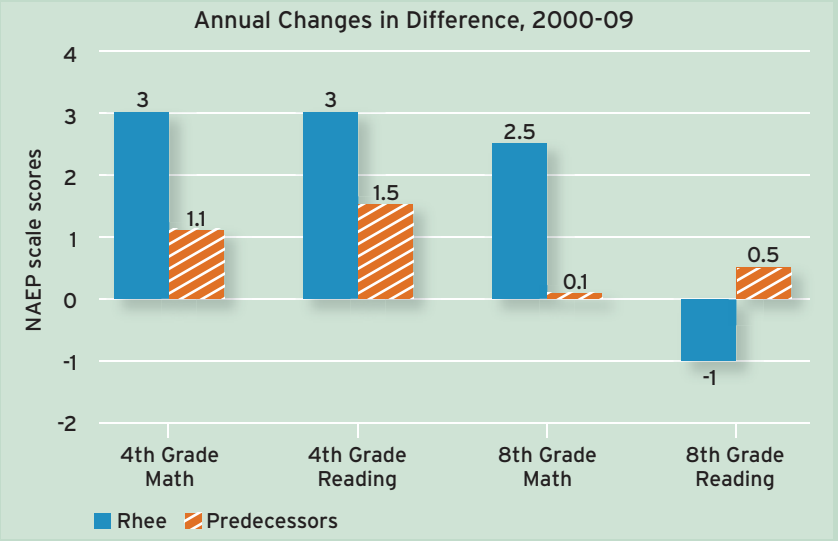
The National Research Council Report

The National Academy of Sciences dates its lineage back to the presidency of Abraham Lincoln, who asked three scientists to help in the "war against the rebellion." Operating under its aegis, the NRC has positioned itself as the only nonprofit organization that can sign contracts with federal agencies without submitting a competitive bid. In the hard sciences, NRC periodically issues major reports of public significance. But on too many occasions it exploits its reputation for objectivity by wandering into domains where scientific knowledge is thin.

NRC has expanded its operations beyond reports to federal agencies. In the case at hand, it acted on a 2007 request of the D.C. City Council "under the leadership of Vincent C. Gray" to carry out an independent evaluation of D.C. public schools. Despite the fact that Gray was already planning his run for mayor, NRC responded enthusiastically to his request by undertaking an energetic fundraising campaign that supplemented the council's own $325,000 in funding with a like amount from a variety of foundations and agencies, including the Spencer Foundation, the National Science Foundation (which

Closing the Gap (Figure 2)

During Rhee's tenure, 4th-grade students gained an average of 3 points each year on NAEP reading and math, relative to the gains made by students nationwide. The annual gains were twice those made during her predecessors' administrations.

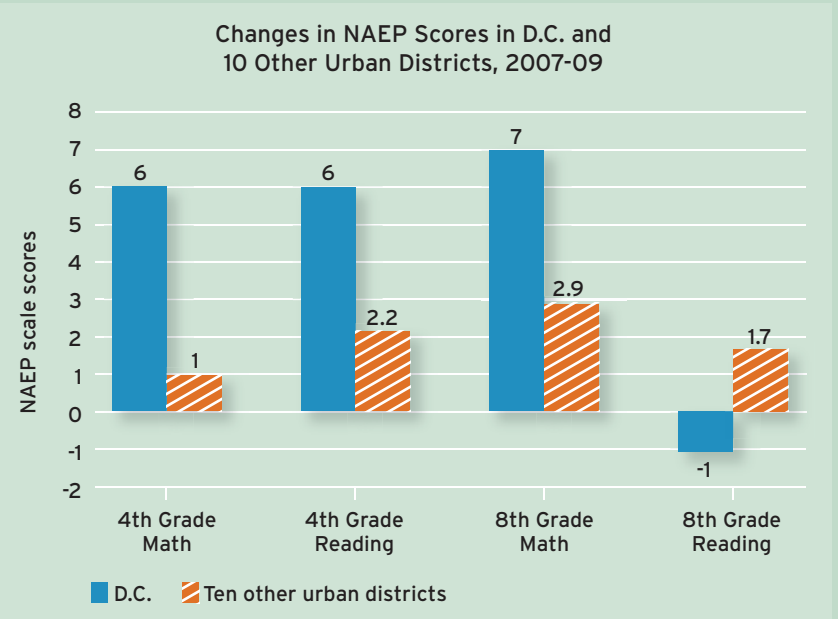


NOTE: Reading scores are only available from 2003 onward.

SOURCE: National Assessment of Educational Progress, National Center for Education Statistics, U.S. Department of Education

Better Than Average (Figure 3)

During Rhee's tenure, D.C. students made more progress than their peers from 10 other urban districts on three of four NAEP tests; only in 8th-grade reading did D.C. lag behind.



SOURCE: National Assessment of Educational Progress, National Center for Education Statistics, U.S. Department of Education

contributed $200,000), and the World Bank (which contributed $25,000).

With $650,000 in hand, NRC staff formed the 14-member, largely academic Committee on the Independent Evaluation of DC Public Schools, consisting of a variety of professors and researchers. Its co-chairs are Christopher Edley, the left-leaning dean of Berkeley law school and, as mentioned, Robert Hauser, former University of Wisconsin sociology of education professor, a liberal critic of accountability systems, who has recently assumed the leadership of NRC's division responsible for education reports.

Guidance for a Future Evaluation

The committee's official assignment was not to carry out an independent evaluation, as its title implies, but only to 1) "provide guidance on how to structure" that evaluation and 2) "provide feedback about implementation" of the Rhee reforms. As part of its "guidance," the committee calls for "systematic yearly public reporting of key data as well as in-depth studies of high priority issues." One needs to look at more than just "student test scores," it says. One needs to establish "suitable indicators" that "track how well the city's public schools are doing." "In-depth studies should be designed to provide deeper analysis of specific questions about high priority issues," such as "teacher recruitment and retention."

If most of this guidance consists of harmless bromides, one recommendation has an edge to it: The evaluation "must be independent of school and city leaders and responsive to the needs of all stakeholders." Read in the context of D.C. politics, this seems to say: Keep the mayor and chancellor out of any independent evaluation, but let the unions play a major role. Now that Vincent Gray is mayor, one wonders just how eager he will be to act on that recommendation!

The committee has not issued a final document, but it has put out a press release and a prepublication version of an unedited version of the report. The rush to print seems to have been necessary in order to carry out the committee's second objective: providing "feedback" on the Rhee record, which it apparently wanted to accomplish before her successor officially assumed office. The first substantive information in the committee's press release reads as follows: "Data suggest that a modest improvement in student test scores has continued...but the committee cautions that it is premature to draw general conclusions about the reforms' effectiveness at this time." Note that the press release talks about a "continuation," not an "acceleration," in "modest," not "striking," improvement in student achievement. An *Education Week* reporter explains that "the evaluators confirmed that students' NAEP scores started to improve before the overhaul law passed, as noted in a report last month by Alan Ginsburg." Clearly, the NRC committee leadership was willing to put an NRC stamp on Ginsburg's claims.

Do Teachers Need to Be at School for Students to Learn?

How did the committee cast doubt on Rhee's effectiveness? The general strategy is to admit the evidence on school improvement in D.C., but then insist that it is impossible to see any connection between that improvement and the work of the chancellor. Of course, it is, as we have said, quite impossible, without experimental evidence, to prove connections between Rhee policies and changes in student gains, but that is not the committee's agenda. Not in its executive summary, in its press release, or anywhere in the report does the committee call for the conduct of experiments that could establish causal relationships between policies and outcomes. On the contrary, the committee recommends gathering still more trend data and conducting old-fashioned case studies that in the end will prove little more than what is already known. And in the pursuit of its second objective, giving feedback on the Rhee reforms, it does not carry out even minimal case-study research to see whether a probable relationship may exist between Rhee policies and classroom outcomes.

Take, for example, the decline in student and teacher truancy. According to 8th-grade student self-reports, the rate of absenteeism declined significantly between 2007 and 2009. Teacher absenteeism also dropped noticeably over these same two years. The days on which 98 percent or more of the teachers were at school climbed from about 68 percent to approximately 85 percent.

Instead of congratulating the district on this improvement, the committee cautions: "It is important to note...that the fact that teacher absenteeism is correlated with achievement does not mean that the absenteeism causes the low achievement. There are many other factors, such as school safety, that affect both teacher absenteeism and student achievement. This is just one example of the many limitations of these data."

In this passage we see a certain bias at work. The incidence of student and teacher truancy declined, the committee admits. But that hardly proves Rhee was a success or that students, in order to learn, need the stability that comes with the presence of their regular teacher. Perhaps school safety also improved, but the committee makes no effort to gather statistics on this point or carry out a case study to see whether Rhee had worked to make schools safer. We are simply left with the caution that a drop in the rate of absenteeism might not prove anything.

Comparing D.C. to Other Big Cities

The committee also acknowledges a notable climb in test scores on the DC CAS test and says that "NAEP shows increases similar to those seen on the CAS." But, it says, "in comparison with other urban districts, the District's scores were similar; many others also showed consistently significant gains."

Really? At the 4th-grade level, D.C. students in math and reading gained 6 scale score points between 2007 and 2009, while the average gain in the other 10 cities for which comparable data are available was only 1 point and 2.2 points, respectively. In 8th-grade math, the D.C. gains were 7 points, as compared to an average of 2.9 points for the other cities. Only in 8th-grade reading does the District of Columbia lag behind, dropping a point, while the others gained 1.7 points (see Figure 3).

Do Demographics Explain Gains?

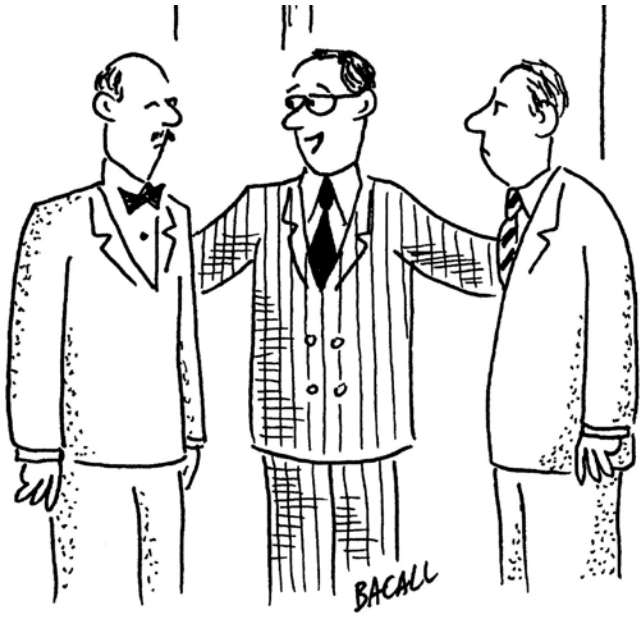
The committee next worries over whether the gains may be due to a change in the composition of the student population in D.C. "The composition of students tested in DCPS...has changed markedly since 2007," the report says. "These patterns could bias the...statistics." *Education Week*'s reporter was told that "the numbers of students with disabilities or limited English proficiency fell during that time. The district also had fewer black students and more white and Hispanic students by 2010."

But is there any reason to believe the gains on the NAEP between 2007 and 2009 were attributable to a shift in the D.C. demography? Did high-income whites and blacks bring their children into the district's public schools, while low-income blacks and Hispanics moved out? According to the committee's own report, signs point in the opposite direction. The percentage of students identified as economically disadvantaged grew from 63 percent in 2007 to 70 percent in 2009. The percentage African American slipped slightly from 85 percent to 83 percent of the total, but the percentage Hispanic increased from 9 percent to 10 percent, while the white population rose from 4 percent to 5 percent. Those needing instruction in the English language increased from 7 percent to 10 percent. It's true that the percentage identified as in need of special education budged downward by 1 percentage point, but the participation rates of special education students on the NAEP increased by 1.5 percent over the two-year period. Nothing in these data indicates that the D.C. schools had fewer challenges in 2009 than they had in 2007.

Rhee's Record

In all the numbers Rhee's critics have assembled, the two facts that stand out have nothing to do with test scores, but rather with student and teacher absenteeism. One does not know how quickly leaders can have an impact on student learning, but strong educational leaders are known for their impact on school culture. If we take Rhee at her word, changing culture was what she was trying to do, and those falling absenteeism indicators suggest that she may have had an effect, even in a short period of time. It's even possible that a change in the D.C. school climate accelerated learning gains. About that one cannot be certain when only two years of NAEP data are available. But one can be quite sure that a case against Rhee has yet to be established.

Paul E. Peterson directs Harvard's Program on Education Policy and Governance.



"Al, this is Jack. He's with the Committee to Eliminate the Board of Education. Jack, this is Al. He's with the Committee to increase Funding for the Board of Education."