

WORKING P A P E R

Student Achievement in the Pittsburgh Public Schools

BRIAN GILL, JOHN ENGBERG

WR-204-EDU

Prepared for A+ Schools: Pittsburgh's Community Alliance for Public Education

This product is part of the RAND Education working paper series. RAND working papers are intended to share researchers' latest findings and to solicit informal peer review. They have been approved for circulation by RAND Education but have not been formally edited or peer reviewed. Unless otherwise indicated, working papers can be quoted and cited without permission of the author, provided the source is clearly referred to as a working paper. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors. RAND® is a registered trademark.

Introduction

Although publicly available school- and district-level test results in reading and mathematics cannot provide a definitive picture of the performance of the Pittsburgh Public Schools (PPS), they can, when analyzed and interpreted with appropriate care, give policymakers and the public several useful reference points for beginning an assessment of student achievement in the district. In previous research, RAND analyzed student achievement results in the district through 2003.¹ Following the state's release of spring 2004 results on the Pennsylvania System of School Assessment (PSSA), A+ Schools asked RAND to analyze and interpret those results, with the aim of helping the citizens of Pittsburgh understand what the state test scores mean—and what they do not mean.

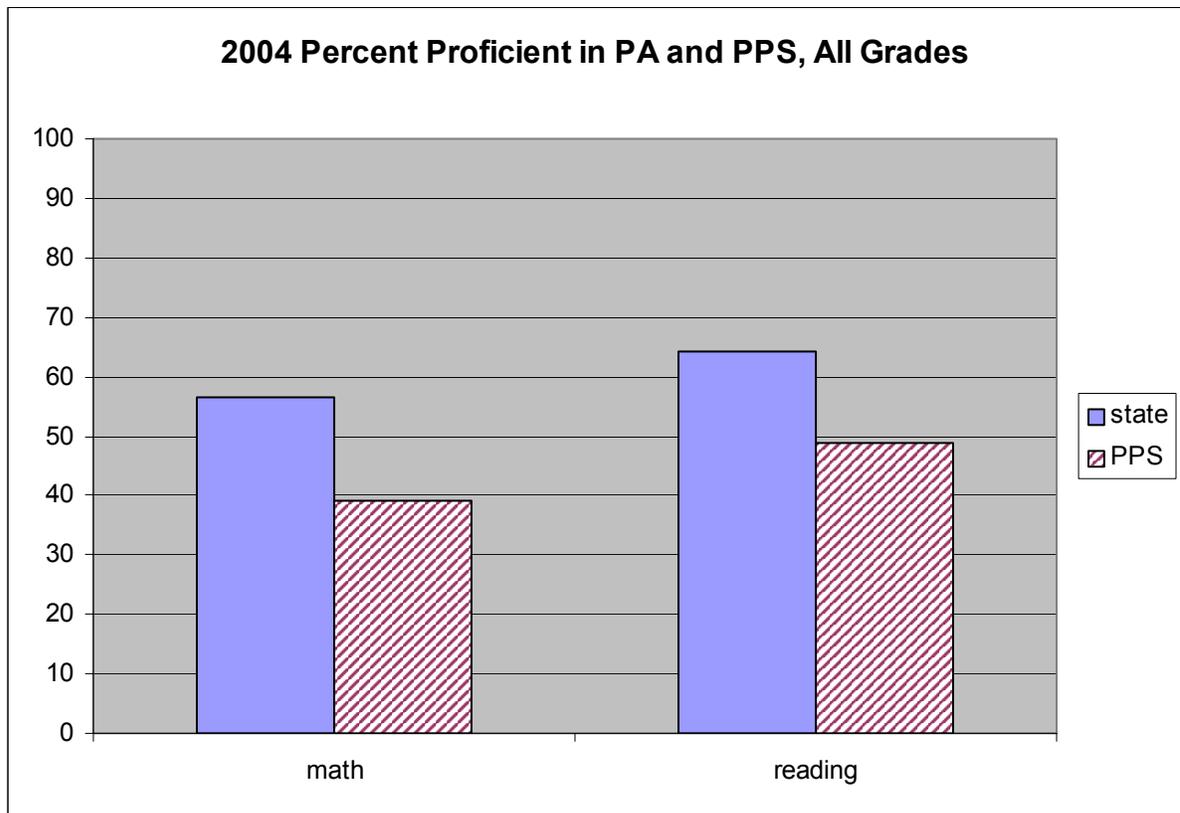
To that end, RAND examined these results and we report our analysis here. We first report current levels of achievement in the Pittsburgh Public Schools relative to the state's objective proficiency standards, to historical trends for the district and the state, and to urban schools serving similar student populations elsewhere in Pennsylvania. We then examine the racial gap in achievement in elementary, middle, and high-school grades in reading and math. Next, we incorporate publicly available district-wide data on other standardized tests in other grade levels to confirm the consistency of patterns found in PSSA data. And finally, we discuss the appropriate uses and limitations of publicly available PSSA results for individual schools in the district.

¹ Gill, Brian P., Rachel Christina, Rebecca Clothey, and Deanna Hill (2003). "Student Performance in the Pittsburgh Public Schools." Santa Monica, CA: RAND (DRU-3149-EDU).

Proficiency rates in PPS and statewide

PSSA tests are given to students in reading and math in grades 5, 8, and 11.² For each grade and subject, the state has set a standard for determining whether a student has achieved proficiency. According to state standards, proficiency rates in the Pittsburgh Public Schools in 2004 were similar to what we observed in 2003. Averaging across grade levels, 49% of tested PPS students achieved proficiency in reading, and 39% achieved proficiency in math. By comparison, 64% of Pennsylvania’s students achieved proficiency in reading, and 56% achieved proficiency in math (see Figure 1).³

Figure 1



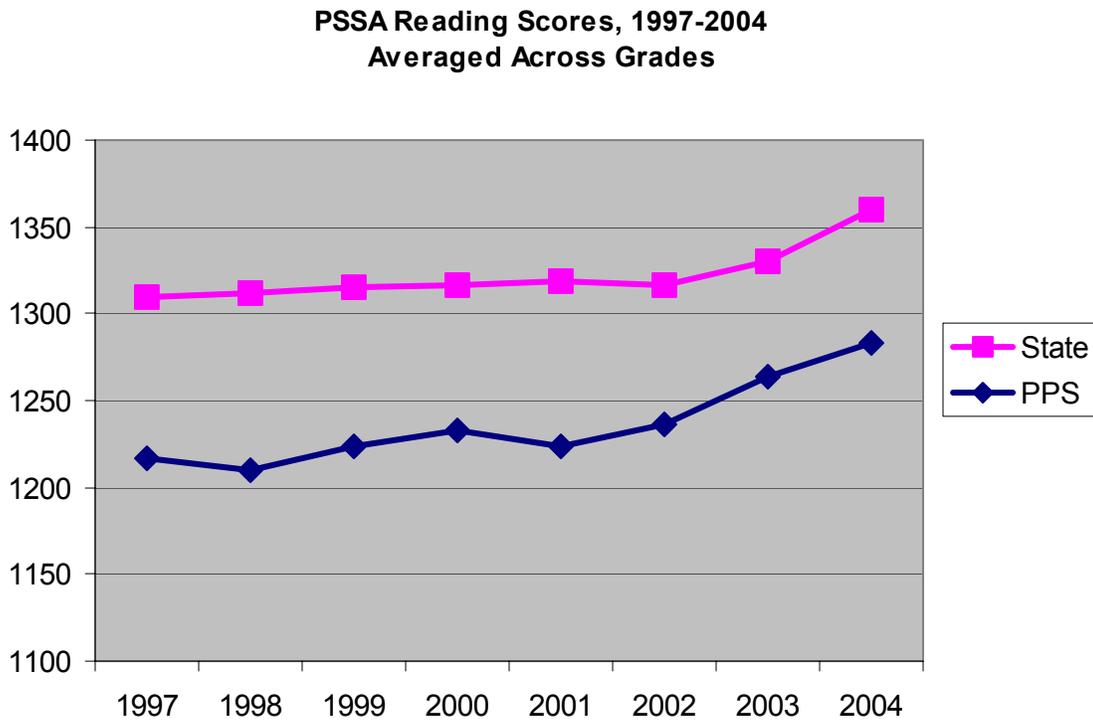
² Tests are also given in grade 3, but those results are not yet included in the state’s accountability calculations, and 2004 results on those tests have not yet been made available by the Pennsylvania Department of Education.

³ These figures constitute averages that are weighted by the number of students tested in each grade.

Achievement trends in PPS and statewide

For most of the last seven years, Pittsburgh’s average PSSA scores have been gradually gaining on the state’s scores, as Figures 2 and 3 indicate. Between 2003 and 2004, however, the state saw relatively large gains in both reading and math, exceeding the gains posted by the district.⁴

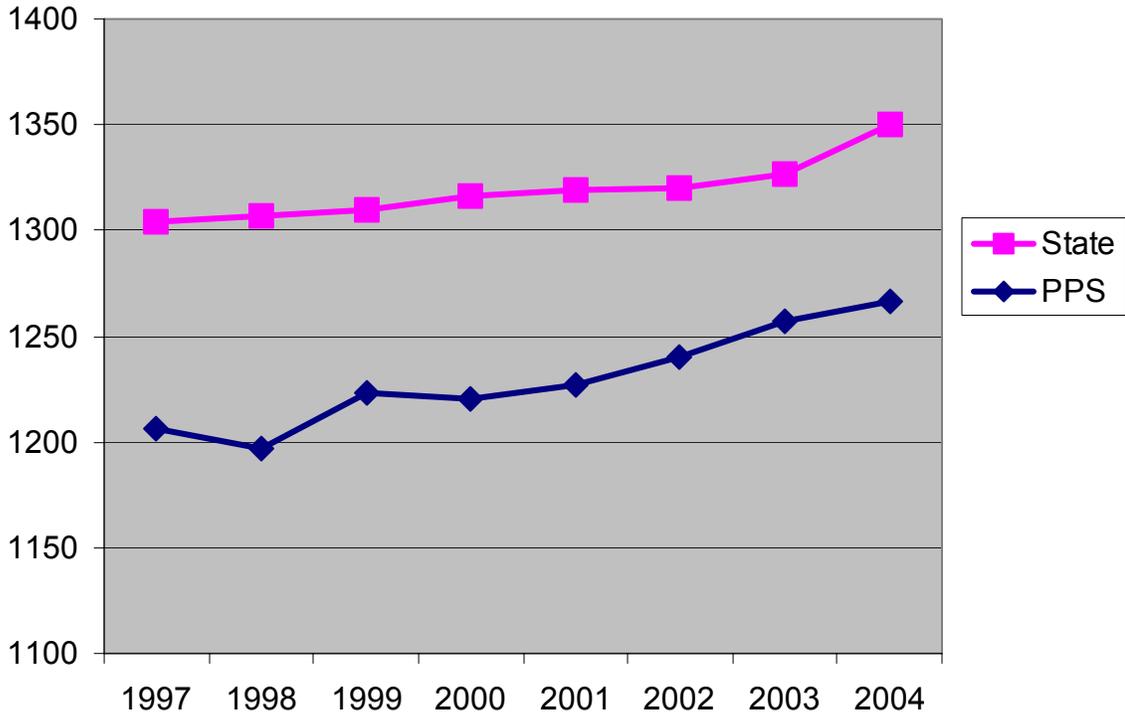
Figure 2



⁴ These results are unweighted averages across grades 5, 8, and 11—each grade level is weighted equally.

Figure 3

**PSSA Math Scores, 1997-2004
Averaged Across Grades**



The only other standardized test on which the district has trend data back to the 1990s is the New Standards Mathematics Reference Exam that it gives to students in grade 4. As on the 5th-grade PSSA math exam and the other PSSA indicators, the district’s results on the 4th-grade New Standards test improved in the late 1990s. Results on the New Standards exam have plateaued over the last four years, however.

Between 2003 and 2004, the proportion of the district’s schools that achieved “adequate yearly progress” (AYP) under state standards established in accordance with the federal

No Child Left Behind Act increased from 45% to 72%. But a direct comparison of AYP results for the two years is confounded by the fact that the state relaxed many of the rules for meeting AYP between 2003 and 2004. The fact that the district's overall average proficiency rate changed little suggests that most of the improvement in AYP results was due to the changes in state standards, rather than proficiency improvements. Indeed, the number of PPS schools where schoolwide proficiency rates exceeded the state's AYP benchmarks (35% proficient in math and 45% in reading) increased only slightly between 2003 and 2004. In reading, the number of schools in the district exceeding 45% proficiency overall increased from 42 in 2003 to 45 in 2004, while in math, 48 PPS schools exceeded 35% proficiency in 2003, versus 51 schools in 2004. In sum, the improvements in AYP status among the district's schools from 2003 to 2004 are not explained primarily by improvements in overall proficiency rates.

Proficiency rates in PPS and similar schools

Although proficiency rates, state averages, and historical trends provide three useful benchmarks for assessing student achievement in the district, it is also useful to examine how achievement in PPS compares to achievement in schools that serve similar populations of students elsewhere. Student achievement is a product not only of the performance of schools, but also of families, peers, and communities. PPS serves a population of students that is disproportionately low-income, African-American, and enrolled in special education programs, compared to state averages. By conducting a statistical analysis that implicitly compares schools serving similar populations of students, we can take a step toward determining the effectiveness of the Pittsburgh Public

Schools in promoting student achievement, relative to other schools and districts. To be sure, the results should be interpreted with caution, because there may be other, unmeasured characteristics of students that affect achievement.

Although Pennsylvania has only a small number of school districts with student demographics similar to Pittsburgh's, it has substantial numbers of individual schools that serve student populations that appear similar to those of individual Pittsburgh schools. Our statistical analysis therefore is conducted with school-level PSSA data, examining the percent of students achieving proficiency in reading and math in grades five, eight, and eleven. We use data from schools across the entire state of Pennsylvania in spring 2004, controlling for students' poverty (the proportion of low-income students in the school), race and ethnicity (percent African-American, Hispanic, Asian, Native American), and special needs (percent of students with special education services, and percent of students classified as English-language learners), as well as the urbanicity of the school. Accounting for all of those factors in a statistical regression, we can estimate how proficiency rates in Pittsburgh compare to those of schools serving similar populations.

In 2004, the regression analysis suggests that Pittsburgh Public Schools are doing neither better nor worse than similar schools elsewhere in the state. In both reading and math, in elementary school, middle school, and high school, proficiency levels in Pittsburgh

schools are statistically indistinguishable from those of schools serving similar populations.⁵

The findings of the regression analysis are less favorable to PPS than in prior years. Gill et al (2003) conducted a similar analysis of 2001 data, finding statistically significant advantages for Pittsburgh schools relative to similar schools in reading and math in both fifth grade and eighth grade (while there was no difference in eleventh grade in 2001, as in 2004). Similarly, in 2003, Pittsburgh's fifth-graders continued to demonstrate higher proficiency on PSSA exams in both reading and math than did fifth-graders in similar schools elsewhere. By 2003, Pittsburgh's advantage in eighth-grade math had eroded, but it maintained an advantage in eighth-grade reading. In sum, until 2004, PPS had a fairly consistent advantage on PSSA exams in elementary and middle-school grades over schools serving similar populations elsewhere in the state, but that advantage has largely disappeared in 2004, leaving the district with proficiency levels that are comparable to those of similar schools. Given that proficiency levels in Pittsburgh have not declined in absolute terms, the elimination of the positive results in Pittsburgh is probably due to improved achievement among comparison schools elsewhere in the state.

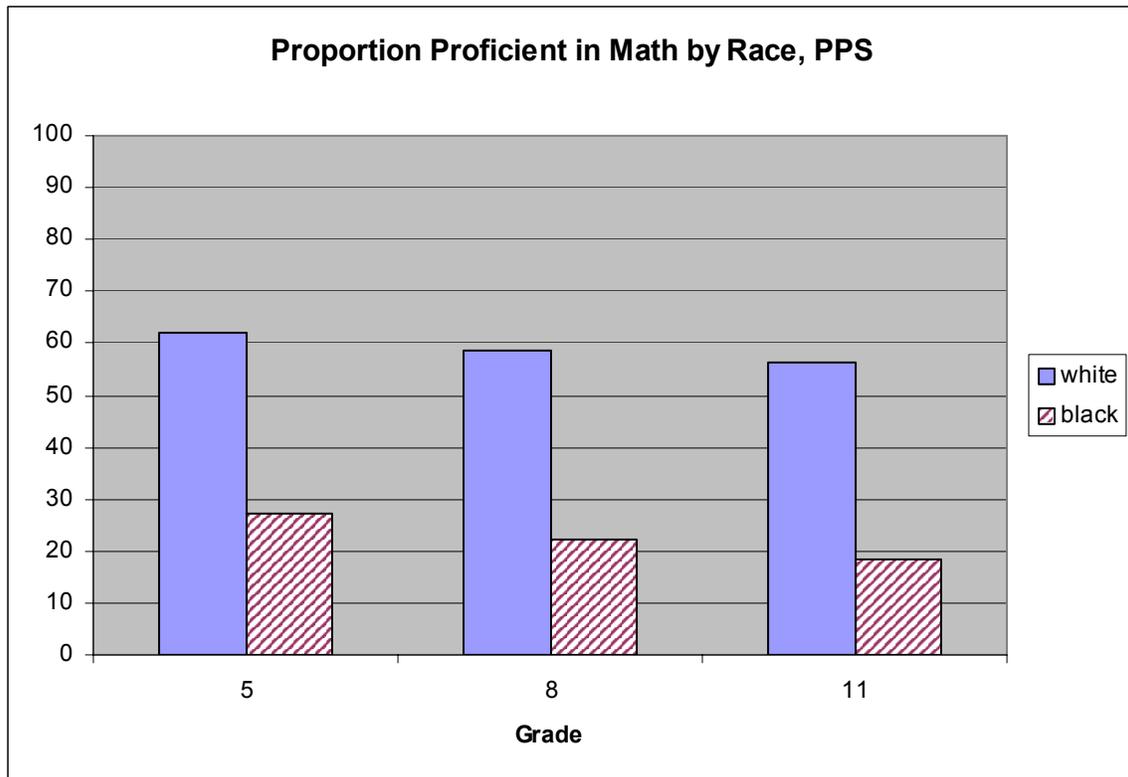
Pittsburgh's racial achievement gap

Nationwide, the gap between the achievement of black students and white students is widely recognized as one of the most serious and long-standing problems facing public

⁵ Note that the regression does not account for differences in the resources available to different schools in the state. Pittsburgh's per-pupil expenditures are well above the state average, and it has a highly experienced teaching force. In 2002, according to Standard & Poor's School Evaluation Services, operating expenditures in PPS were \$11,282 per pupil, versus a Pennsylvania average of \$8295 per pupil (see www.ses.standardandpoors.com).

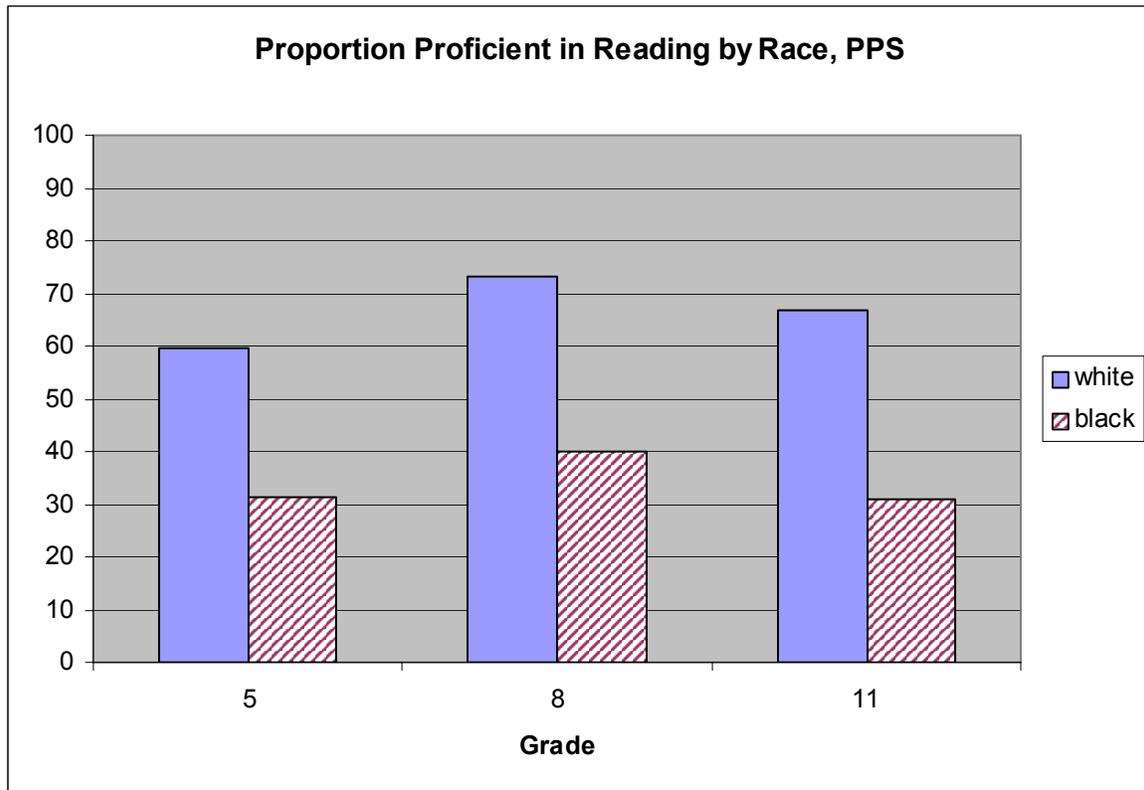
schools and American society. In Pittsburgh, as elsewhere, all evidence suggests that the racial gap is distressingly large. As Figures 4 and 5 indicate, 2004 PSSA results find white students in the district outscoring black students by large margins; in some cases (especially in math), the proportion of white students achieving proficiency is more than double the proportion of black students achieving proficiency. By nearly all available measures (including TerraNova and New Standards exams as well as PSSA), the district’s gap between the scores of white and black students is even larger than the gap between poor and non-poor students.⁶

Figure 4



⁶ Prior analyses conducted by the district, using student-level data, have demonstrated that the racial achievement gap exists even after accounting for low-income status (see Gill et al, 2003).

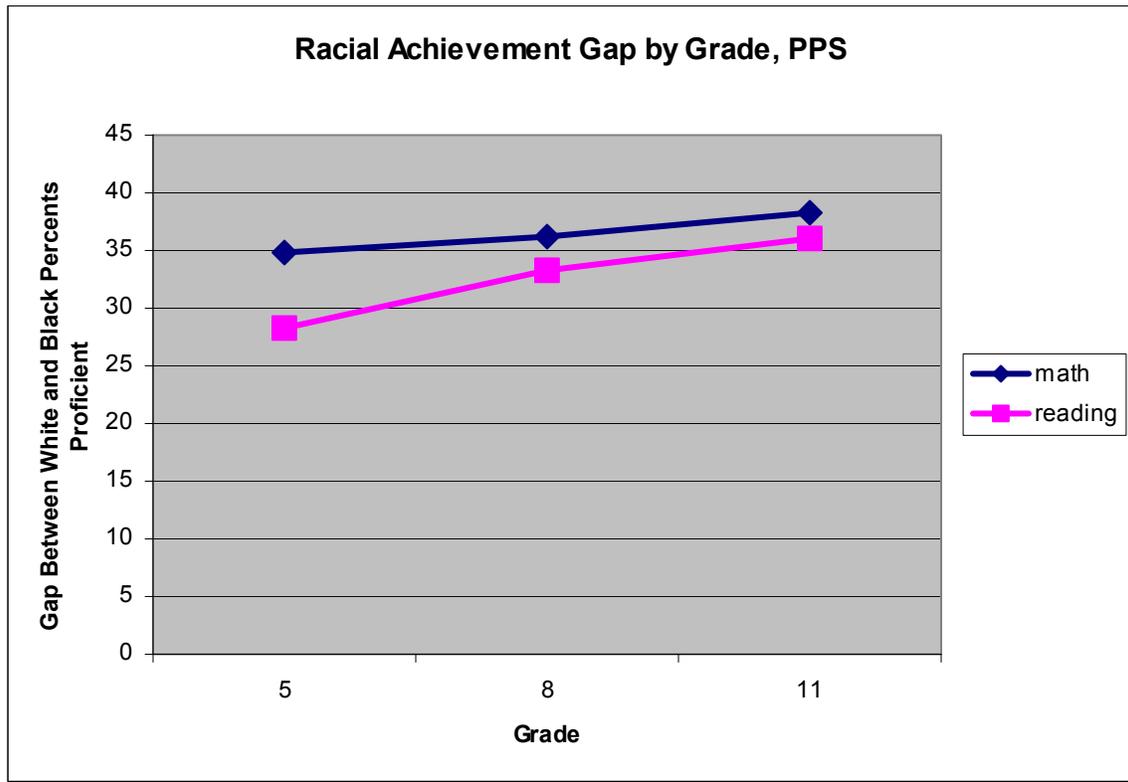
Figure 5



In both reading and math, Pittsburgh’s racial achievement gap is larger in grade 8 than grade 5, and is larger still in grade 11, as shown in Figure 6. A larger racial achievement gap in higher grades can also be seen in much of the district’s data on tests other than the PSSA. On the New Standards Mathematics Reference Exams, now given to Pittsburgh students in grades 4 and 10, the racial achievement gaps in proficiency are substantially larger in grade 10 than in grade 4, on all three dimensions measured on the exams (skills, concepts, and problem-solving). Similar results can be seen on the TerraNova exams taken by Pittsburgh students in reading and math in a variety of grades.⁷

⁷ Racial achievement gaps on the TerraNova exams are slightly smaller in grade 9 than in grade 7, but in all other pairwise comparisons, the racial achievement gaps are larger in the higher grade.

Figure 6



A larger apparent racial achievement gap in higher grades could result merely from differences in the ways that tests are scaled at different grades. But there are indications that the larger gaps in PPS at higher grades are more than just a scaling phenomenon, and represent a Pittsburgh-specific issue. We conducted statistical regressions not only for the district as a whole, but also separately for black students and white students in the district, using school-level scores disaggregated by race, and controlling for the same factors used in the analysis of overall results. The disaggregated results by race tell a different story than the overall results.

As reported above, the overall results, including students of all races, showed no “Pittsburgh effect”—PPS students are doing neither better nor worse than students in similar schools elsewhere in the state, in both subjects at all three tested grade levels. The regression analyses of disaggregated scores show similar results in grade 5, with some divergence of black and white results in grade 8 and further divergence in grade 11. In 8th grade, Pittsburgh’s black students show no difference in achievement compared to black students in similar schools elsewhere, while Pittsburgh’s white students have a 6-percentage point advantage in reading over white students in similar schools (and no difference in math). In 11th grade, Pittsburgh’s white students have a statistically significant advantage over white students in similar schools elsewhere in Pennsylvania, by 10 percentage points in reading and 14 points in math. Black 11th-graders in Pittsburgh, by contrast, achieve proficiency at rates no better than black students in similar schools in the state.

These disaggregated regression results suggest that Pittsburgh’s larger racial achievement gap for older students is driven by the fact that white students in the district are doing better (relative to their counterparts in similar schools in the state) in 11th grade than in earlier grades. In 5th grade, both black students and white students in Pittsburgh are doing about as well as students of their own race in similar schools in Pennsylvania. In eighth-grade, black students in PPS are doing as well as black students in comparison schools, while white students in PPS are doing a bit better in reading, compared to white students elsewhere. In 11th grade, black students in Pittsburgh maintain pace with black

students in similar schools elsewhere, while white students in Pittsburgh are doing substantially better than white students in similar schools elsewhere.

It is not clear why the results for white students in Pittsburgh are better for older students while the results for black students are not. It is possible that the white population in PPS high schools is more advantaged than the white population in lower grades—a phenomenon that could occur if PPS high schools see an influx of well-prepared white students from private and parochial schools. Unfortunately, the aggregate data that are currently available are too coarse to provide further insight into the explanation for the district's high-school racial gap. Further investigation is warranted, and the question could be explored in depth using the district's longitudinal, student-level database.

Assessing the performance of individual PPS schools

Parents, the public, and district officials are appropriately interested not only in the average performance of the district as a whole, but also in the performance of individual schools. Proficiency rates in individual schools across the district vary widely. In mathematics in 2004, individual PPS schools achieved proficiency rates varying from 6% on the low end to 87% on the high end. Similarly, reading proficiency rates in the district varied from 11% to 87%.

An examination of statistical regression results at the school level permits a preliminary, tentative assessment of whether individual schools in the district are doing better or worse than individual schools serving similar student populations elsewhere in

Pennsylvania. Using the same method used to compare average results for schools districtwide to similar schools elsewhere, we can attempt to account for differences among individual Pittsburgh schools as measured by poverty, race, limited English proficiency, and special education status. Regression results for individual schools are less reliable than results aggregated to the district level, for two major reasons. First, they involve small numbers of students—typically including only a single grade level at each school—which substantially reduces the reliability of the estimates. (Similarly, annual fluctuations in the proficiency rate for individual schools are often meaningless, driven more by random changes in cohorts of tested students than real changes in school performance.) Second, many Pittsburgh schools house magnet programs that students choose to attend, making it likely that there are systematic differences in school populations that are not fully reflected in the observable student characteristics that are incorporated in the analysis. Simply because they have chosen to attend, the family characteristics of magnet school students are likely to differ in important ways from those of students who appear similar in terms of race, poverty, English proficiency, and special education status but who do not choose to attend these programs.

Given the limitations of the available data, we think it would not be appropriate to publish the results of the regression analysis for individual schools in the district. Nevertheless, it is useful to conduct such an analysis to examine the extent to which school-level results across the district change when they are compared to demographically similar schools across the state, rather than ranked in terms of raw

proficiency results. For the school-level analysis, we aggregate results across three years of data (2002-2004) for each school to improve the stability of the estimates.

The school-level regression results confirm that some of the differences in proficiency rates among schools within the district reflect differences in the student populations served. But they also clearly show that there are large differences in the performance of PPS schools that are not explained by observable student characteristics. Pittsburgh has some schools where proficiency rates are substantially higher than their demographics would predict, and other schools where proficiency rates are substantially lower than their demographics would predict. Proficiency rates vary substantially within the district, even among schools serving similar populations. Again, further analysis using longitudinal, student-level data available to the district would be well warranted, and could be used to provide more-reliable indicators of the performance of individual schools.

Possibilities for further analysis

The comprehensive electronic data warehouse that PPS has developed could be used for more-extensive and rigorous assessment of the performance of schools and specific educational initiatives, with sophisticated statistical methods that examine growth over time in the achievement of individual students. Publicly available schoolwide and districtwide PSSA results have limited utility, because they capture only a snapshot of student achievement at single points in time in a small number of grades. But much more might be learned about educational performance in the district through intensive analysis

of the district's impressive, student-level data warehouse. We encourage the district to take advantage of its existing electronic data system by applying some of the powerful statistical tools that are available to provide estimates of performance that come closer to measuring the "value added" provided by the schools in raising the achievement of individual students.