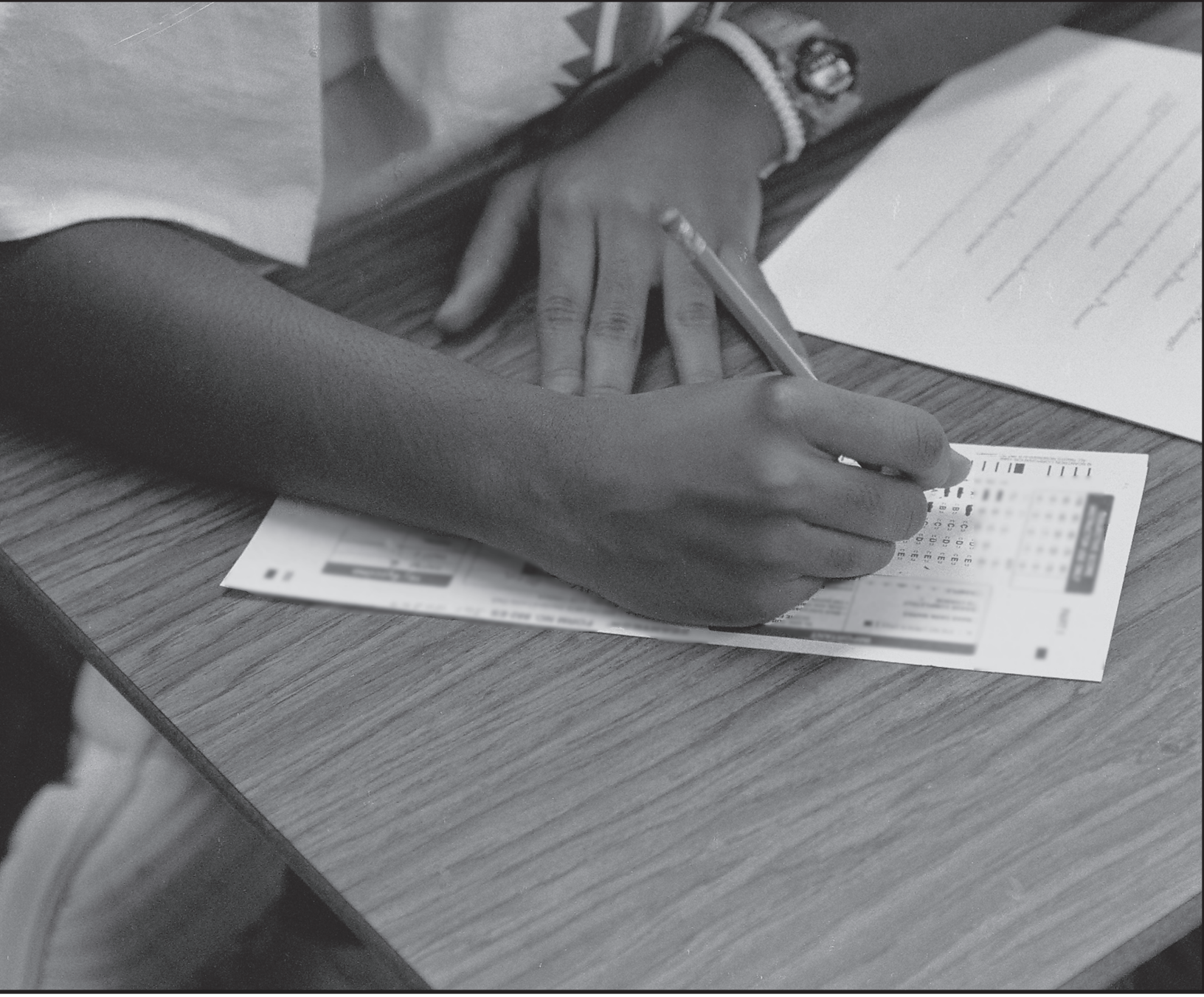


Understanding the Prairie State Achievement Exam

A descriptive report with analysis of student performance

September 2005

Stephen M. Ponisciak

Consortium on Chicago School Research *at the University of Chicago*

Acknowledgments

The author would like to thank the following colleagues from the Consortium for their reviews of this report: Elaine Allensworth, Tony Bryk, John Easton, Holly Hart, Melissa Roderick, Penny Sebring, and Mark Smylie. Consortium colleagues Stuart Luppescu, Todd Rosenkranz, Macarena Correa, Vanessa Coca, Susan Sporte, David Stevens, and David Kerbow also provided feedback on the report. Consortium Steering Committee Members Donald Fraynd, Elizabeth Hawthorne, Arie van der Ploeg, and Connie Wise offered helpful comments regarding this report as well.

This study was made possible by grants supporting the Consortium's core research from the Joyce Foundation, the John D. and Catherine T. MacArthur Foundation, and the Spencer Foundation.

When this report went to press, Chicago Public Schools had recently released preliminary results from the 2005 PSAE. The results show substantial gains in high school reading scores; however, math and science scores appear to have decreased very slightly. The Consortium does not currently have access to this data, thus we have not performed any analyses to include in this report.

Understanding the Prairie State Achievement Exam

Stephen M. Ponisciak
Consortium on Chicago School Research

September 2005

TABLE OF CONTENTS

INTRODUCTION 1

CHARACTERISTICS OF THE PSAE 2

ACCOUNTABILITY REQUIREMENTS 6

CPS STUDENTS' PERFORMANCE ON THE PSAE 10

RELATIONSHIPS AMONG TESTS 14

PERFORMANCE COMPARISONS 18

CONCLUSION 24

Appendix—Supplemental Tables 25

Endnotes 29

Works Cited 30

INTRODUCTION

In 2001, the Illinois State Board of Education (ISBE) instituted a new standardized test for high school students, the Prairie State Achievement Examination (PSAE).[1] This report gives an overview of key elements related to the PSAE and is intended to provide a foundation for the Consortium's future reports on the PSAE. First we examine the characteristics of the PSAE, including its origination, subject areas and component tests, scoring, and score reporting. We then explore how the PSAE is used in local and national accountability policies, addressing which students are required to take the exam and which students actually take it. We analyze student performance on various parts of the PSAE, as well as the relationships among the PSAE's component tests. We conclude with performance comparisons across various groups of students and look at how student attrition over time affects PSAE performance in the Chicago Public Schools (CPS).

CHAPTER 1

Characteristics of the PSAE

Since 2001, the PSAE has been administered to all eleventh-grade students in Illinois public high schools. It currently covers mathematics, reading, science, and English. From 2001 to 2004 social science and writing were also covered by the PSAE. This report is based on data from 2001 to 2004.

Origination

Like most standardized tests used in accordance with No Child Left Behind (NCLB) reporting requirements the PSAE is criterion-referenced, meaning it identifies how well Illinois students measure up to state standards in the relevant content areas. The Illinois Standards Achievement Test (ISAT), developed in part by Illinois teachers and administrators, is also a criterion-referenced test. It is administered in reading and math to Illinois public school students in grades three, five, and eight to determine whether they meet state learning goals. Originally state law required students to take the ISAT in grade 10 for reading, writing, and mathematics, and again in grade 11 for science and social science, as well as the PSAE in grade 12 for reading, writing, mathematics, science, and social science. However, before the law could be implemented, legislation was enacted in 1999 to eliminate the ISAT from grades 10 and 11 and establish the PSAE in grade 11 as the only required statewide test after grade eight, in order to reduce the testing burden.[2] In July 2004, the law was amended again to remove the writing and social science sections from the PSAE. Since spring 2005, the PSAE has included only math, reading, and science tests, as well as the ACT English test.[3]

Eleventh-grade students take the PSAE over two days in April each year. The administration and scoring of the PSAE cost ISBE about $50 per student in 2001.[4] The schedule for the 2004 PSAE administration is shown in Appendix Table 1; it was similar in the years prior to 2005.[5] Changes in state law that reinstate the ISBE writing test will make the testing schedule somewhat different in the 2006–07 school year.[6]

According to Illinois school officials, a passing score on the PSAE is not a graduation requirement, because state law does not permit it and because ACT, Inc. does not allow the ACT to be used for this purpose.[7] CPS does not use the PSAE for student-level accountability either. In fact, many students have low PSAE

CPS does not use the PSAE for student-level accountability either. In fact, many students have low PSAE scores but still graduate.

scores but still graduate. Of the CPS students who did not meet or exceed standards on any part of the 2001 PSAE, almost 75 percent graduated from CPS by fall 2002. Furthermore, of the students who received the academic warning designation (the lowest performance level) in all five subjects of the 2001 PSAE, 66 percent graduated by fall 2002. In comparison, of the students who met or exceeded standards on all five parts of the 2001 PSAE, about 92 percent graduated by fall 2002.

Subject Areas and Component Tests

From 2001 to 2004, the PSAE included three tests: the ACT (covering reading, math, English and science), the WorkKeys (reading and math), and the ISBE-developed tests (writing, science, and social science). During this time, four of the five subject areas tested by the PSAE were assessed by two different tests; social science consisted of a single test.[8] The subject areas and component tests of the PSAE from 2001 to 2004 are illustrated in Appendix Table 2.

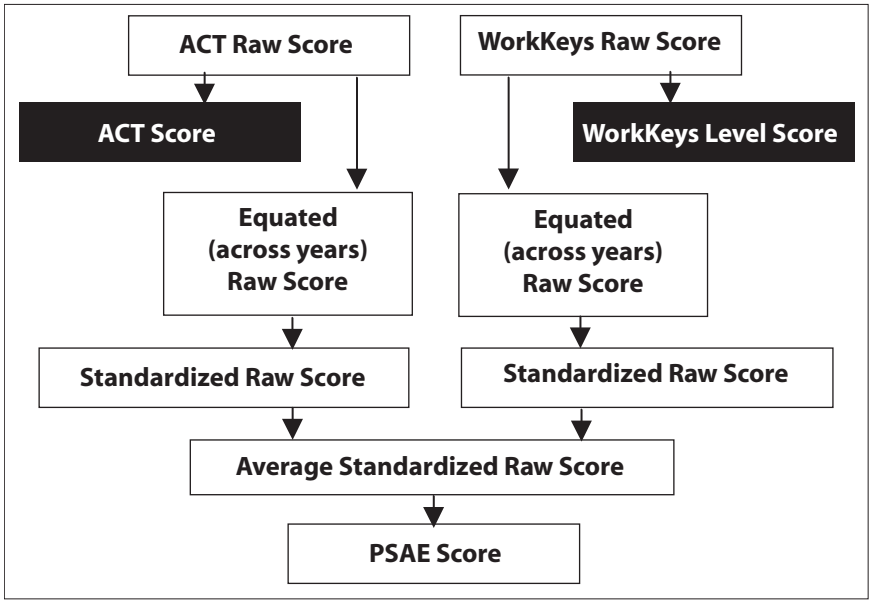
In creating the PSAE, ISBE opted to administer the ACT as a component test of the PSAE, a decision that has been somewhat controversial. One argument in favor of this decision is that the ACT is intended to measure readiness for college, thus taking it focuses all students' attention on postsecondary education. Critics counter that the ACT measures only a limited range of skills and therefore should not be part of the PSAE. ISBE argues that including the ACT is appropriate because ACT scores give students information about their strengths in the academic areas tested, which will help students select courses during their senior year of high school and during their postsecondary education. In addition, students' ACT results from the PSAE may be used for college admissions (although the scores are not admissible for the purposes of National Collegiate Athletic Association eligibility or admission to the U.S. Air Force Academy).

The ACT was designed primarily to assist selective-admission colleges in evaluating the educational development of students who attended high schools with varying grading standards; therefore the PSAE also includes the WorkKeys tests in math and reading in order to measure standards not fully assessed by the ACT. The WorkKeys tests were developed by ACT, Inc. to measure how well students apply reading and math skills to situations they might encounter in the workplace. ISBE states that employers can use WorkKeys scores to assess an applicant's suitability for a job. The WorkKeys tests have been criticized for including what some label lower-level material; however, ACT, Inc. argues that the tests measure skills needed for "professional, technical, and managerial jobs requiring higher-than-average educational levels" as well as skilled trades jobs.[9]

Scoring

Students receive a score on a scale of 120 to 200 for each of the subjects tested by the PSAE. A student's raw score on any subject test is equal to the number of questions answered correctly. For subjects covered by two tests, each raw score is converted to an equated raw score. This process controls for changes in test difficulty over time. For example, correctly answering 75

Figure 1: Combination of Scores in PSAE Mathematics and Reading



Note: ACT and WorkKeys raw scores influence a student's PSAE score; they also determine the student's ACT and WorkKeys test scores. However, it is important to note that the ACT and WorkKeys scores are determined independently of the PSAE score.

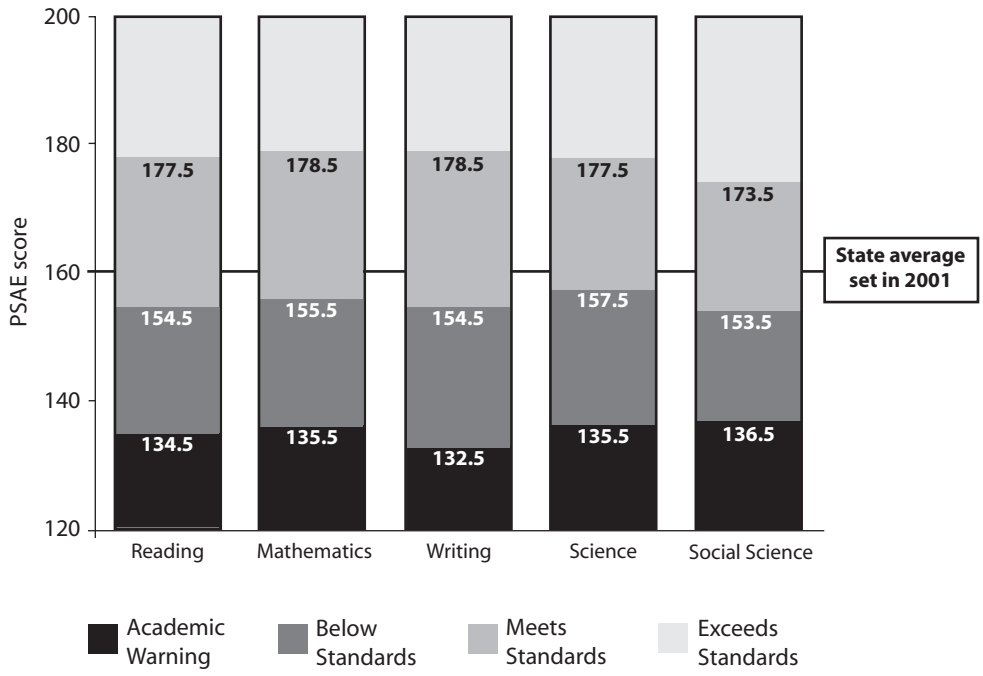
percent of the questions on an easier test might be equivalent to answering 60 percent correctly on a more difficult test. The equated scores are converted to standardized scores using the same mean and standard deviation each year, so that each student's performance is placed in the context of all Illinois public school students over multiple years. The two standardized scores for each subject are averaged, and the average standardized score is converted to a PSAE scale score between 120 and 200. When this scale was established after the first PSAE administration in 2001, the state average score was set at 160, with a standard deviation of 15.[10] In Figure 1, we show how a reading or math score is produced by combining scores from the ACT and WorkKeys tests. For science and writing scores, the WorkKeys raw score is replaced by raw scores on the ISBE-developed tests. The social science score is derived from a single test.[11]

PSAE scores in all subject areas are associated with one of four performance levels: academic warning, below standards, meets standards, and exceeds standards. After the first administration of the PSAE in 2001, ISBE convened a panel of high school teachers, curriculum directors, postsecondary faculty, and members of the business community to define the performance levels. This process, called a modified Angoff procedure, asks each panelist to estimate the percentage of students at the lower boundary of each performance level who would answer each question correctly. The ratings from this process were transformed to determine the cutoff points between the performance levels shown in Figure 2.[12]

Students' scores on the ACT, WorkKeys, and ISBE-developed tests are determined separately from their PSAE scores. Using item response theory (IRT), a student's raw score on the ACT is converted to a scale score that is assigned a value from 1 to 36.[13] Scale scores have had the same value each testing year since the current version of the ACT was introduced in 1989, allowing for the comparison of scores over time.

WorkKeys scores are reported in terms of skill levels, which indicate the level of mastery demonstrated by the student. The possible scores are less than 3, 3, 4, 5, 6, and 7. Mastery of a skill level means that a student could be expected to correctly answer at least 80 percent of the questions at that level—it also implies

Figure 2: PSAE Performance Categories



mastery of all lower skill levels. The WorkKeys tests contain six questions at each skill level and three additional items for pilot-testing purposes. IRT was used to determine the minimum number of correct answers necessary to assume mastery of a level on a given form of the WorkKeys test. ACT, Inc. defines levels 3 through 5 as the levels of skill necessary for 90 percent of jobs in the United States. Level 3 jobs require simple math skills and include service and factory jobs. Level 5 jobs require more complex ability; these positions include certified trades, such as plumbers and medical lab technicians.

Score Reporting

Students and schools receive score reports in the fall following administration of the PSAE. Students receive individual score reports, while schools receive a performance profile and school roster. These reports are summarized in Appendix Table 3.

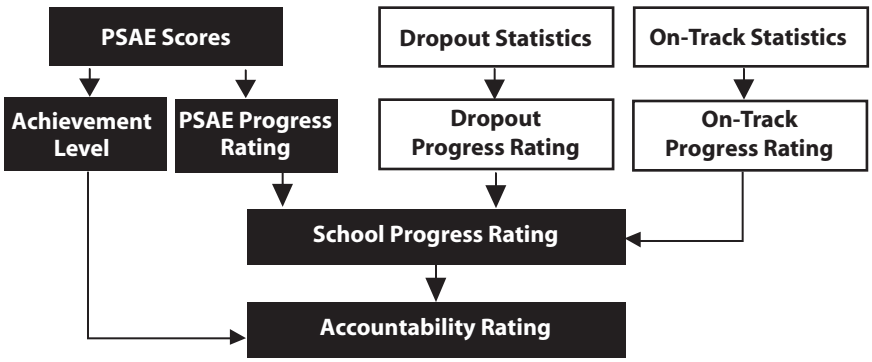
Illinois law requires that PSAE scores appear on students' transcripts. ACT scores may not be included without the student's permission. Students who want to improve their scores may retake the entire PSAE during their senior year, but these scores are not used for school and state accountability purposes.

CHAPTER 2

Accountability Requirements

Every year, CPS assigns each high school an accountability rating, based on PSAE scores and other factors, which is used to determine the schools most in need of improvement. The rating system used PSAE scores in two ways until 2004, when it also began to calculate adequate yearly progress. First, PSAE scores for the current school year determine a school's achievement level. Second, the change in PSAE performance from the previous year determines the school's PSAE progress rating. Changes in the school's dropout rate determine the school's dropout progress rating, and changes in the school's on-track rate (the percentage of students on schedule to graduate, determined by the number of core courses passed and the number of F's received) determine the school's on-track progress rating. PSAE, dropout, and on-track progress ratings are combined to yield a school progress rating. This rating is combined with the achievement level to determine the accountability rating.[14] We illustrate this process in Figure 3. As of 2004, adequate yearly progress is calculated with regard to the percentage of reading and math scores that meet or exceed standards, the participation rate of students taking the state tests, and the graduation rate.

Figure 3: PSAE and Accountability in CPS (prior to 2004)



Note: Black boxes show elements of the accountability rating that are affected by PSAE scores. Beginning in 2004 as part of the accountability policy the state calculates whether a school has made adequate yearly progress with regard to the percentage of reading and math scores that meet or exceed standards, the participation rate of students taking the state tests, and the graduation rate.

In addition to meeting accountability standards within CPS, schools and districts must meet national accountability standards in accordance with NCLB. In order to make adequate yearly progress, schools and districts must have at least 95 percent of their eligible students (overall and within subgroups defined by gender, race/ethnicity, income, migrant status, bilingual education participation, and special educa-

tion identification) participate in the PSAE or an alternative test.[15] Both schools and districts must reach specified performance targets in reading and math. Also, schools and districts must have a specific percentage of the students who were enrolled by October 1 of the current school year meet or exceed standards in reading and math. Elementary schools must meet an attendance rate goal, high schools must meet a graduation rate goal, and the districts must meet both. [16]

A school's performance on the PSAE (and therefore its status in an accountability program) is affected by the group of students who take the test each year. Knowing how many of a school's students took the test and whether those students are representative of the school's population of students is important for interpreting test scores. Subsequent analyses in this report are restricted to students who were active in CPS during the spring semester of the testing year, and who had enrolled before March 1 of the testing year. Because the PSAE is administered in late April (with make-up testing in May), the students included in this analysis are those who were enrolled in a Chicago public high school for two months or more during the relevant school year.

Exemptions and Modifications

Almost all students, including those with disabilities, are required to take the PSAE. The only exceptions are for students with significant cognitive disabilities whose individualized education programs (IEPs) indicate that state testing is inappropriate for them. Those students may take the Illinois Alternative Assessment (IAA) instead. If state testing with accommodations is appropriate for the student, ACT, Inc. must approve the accommodations for the ACT portion of the PSAE. The approved procedures can be used as a guide for accommodations on WorkKeys and ISBE-developed tests, but final decisions about these arrangements are left to the local school leadership. Accommodations may include extra time, testing over more than one day, alternate test formats, or additional breaks. In order to receive accommodations from ACT, a student's disability must be professionally diagnosed, the student's IEP or 504 plan must require extended time, and the student must receive these accommodations for regular school work and testing.[17]

Students in bilingual education are exempt from the PSAE if they have been in bilingual education for less than five years and their lack of understanding of English would prevent them from comprehending the questions. Until spring 2004, students had to take the PSAE unless they had been in bilingual education for fewer than three years. The alternative test for students with limited English proficiency (LEP) in a state-approved bilingual education program is the Illinois Measure of Individual Growth in English (IMAGE). The Consortium does not have test results for CPS students who took the IAA or IMAGE—we only have information about the number of students tested.

Who Actually Takes the PSAE?

There are two important considerations when examining PSAE participation. The first issue is whether enough eleventh-grade students take the PSAE to meet NCLB requirements for participation and adequate yearly progress. The second issue is to what extent PSAE scores reflect the performance of CPS high schools. The district's high dropout and failure rates mean that many CPS students do not remain in school long enough to take the PSAE.

Under NCLB, 95 percent of eleventh-grade students must participate in assessment tests. About 85 percent of the CPS students enrolled in eleventh grade in 2004 took at least part of the PSAE. Students taking alternative assessments are added to those taking the PSAE to determine compliance with NCLB. Adding the students who took the IMAGE or IAA would raise the percentage tested in 2004 to 91 percent. However, this is still below the 95 percent participation requirement for NCLB.

Figure 4: Changes in 11th-Grade Test Inclusion among CPS Students, 2001–2004

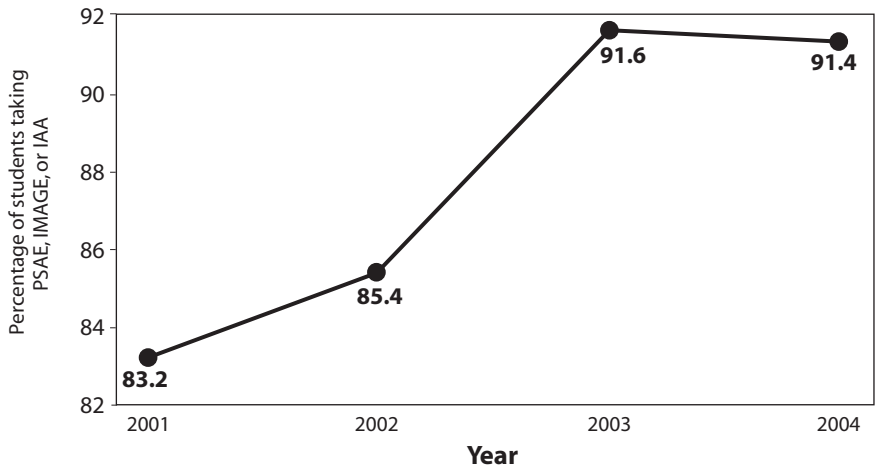


Figure 4 shows the change in the percentage of students enrolled in CPS for at least two months who took the PSAE, IMAGE, or IAA from 2001 to 2004. The large increase from 2002 to 2003 is primarily the result of an improvement in record keeping, as the number of students with incorrect identification codes in the CPS data files was considerably smaller in 2003 and 2004 than it was in 2001 and 2002. If all the codes were repaired, the proportion of active eleventh-grade students who were enrolled in CPS for at least two months before PSAE administration, and who actually took the PSAE, could be about 5 percentage points higher in 2001 and 2002.

Figure 5 shows the percentage of CPS students in eleventh grade who took all, part, or none of the PSAE in 2004. The proportion of students tested varies among racial/ethnic groups, with Asian students participating at the highest rate. Within each racial/ethnic group, female students were more likely than male students to take the PSAE in each year.

The second important question regarding PSAE participation is whether the test is a complete depiction of student performance in Chicago public high schools. Many students have left CPS before the administration of the PSAE in spring semester of eleventh grade. PSAE scores are likely higher than they would be if the dropout rates were lower, because the students who remain in school and take the PSAE probably have experienced more academic success than the students who dropped out of school.

As shown in Figure 6, about 51 percent of the students who were freshmen for the first time during the 2000–01 school year remained in CPS and took the PSAE in 2003, and about 2 percent took the test a year later, for a total of 53 percent. Out of this group of freshmen, 4 percent were in eleventh grade in CPS during 2003 or 2004 but did not take the PSAE. Many of these students probably took the IMAGE or IAA. Another 11 percent were in CPS in 2003 or 2004 but had not yet progressed to eleventh grade. The remaining 32 percent were not enrolled in a CPS high school during either 2003 or 2004 because they had either dropped or transferred out of CPS. These figures are similar to previous years' figures.

When we examine attrition by race in Figure 6 we find that 66 percent of Asian students who were freshmen in fall 2000 took the 2003 or 2004 PSAE. Fifty percent of African-American students, 55 percent of Latino students, and 55 percent of white students who were first-time freshmen in CPS high schools in fall 2000 took the 2003 or 2004 PSAE. These rates also vary by gender within racial groups, with female students participating at a higher rate than male students within each racial group. Asian students took the test at a higher rate than students from any other racial or ethnic group.

Figure 5: Percentage of 11th-Grade CPS Students Who Took the PSAE in 2004

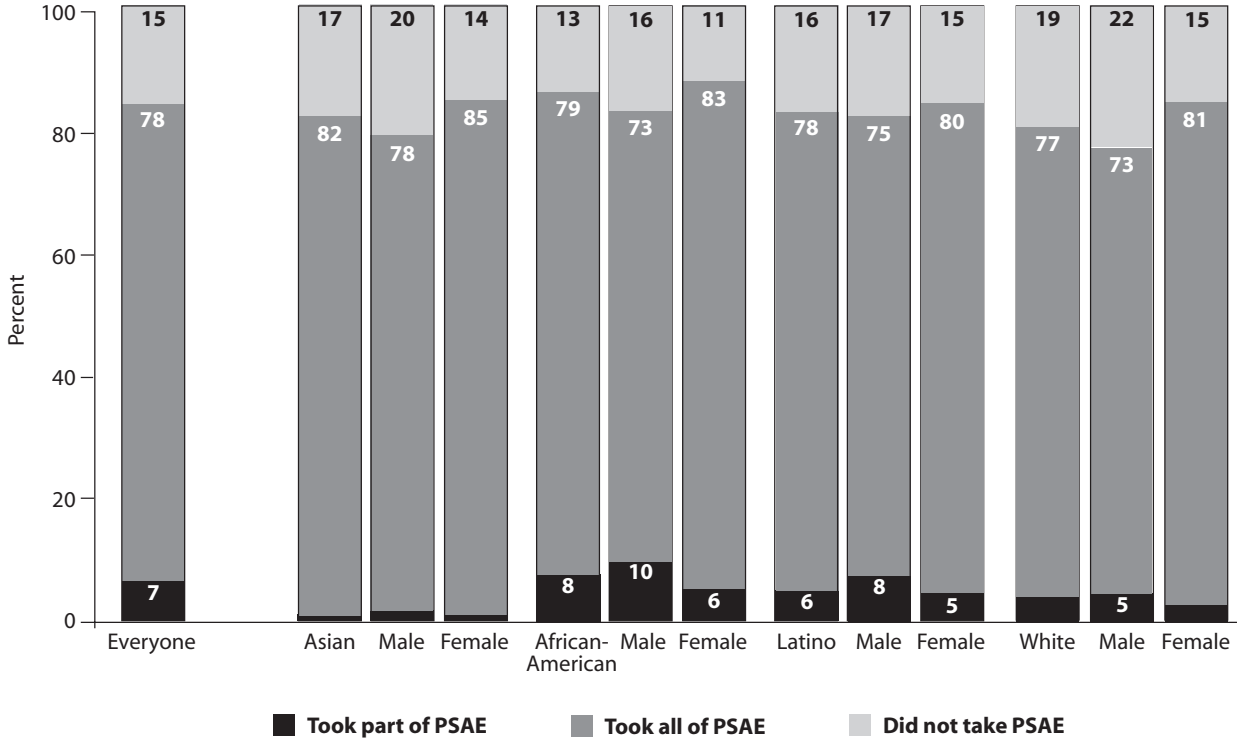
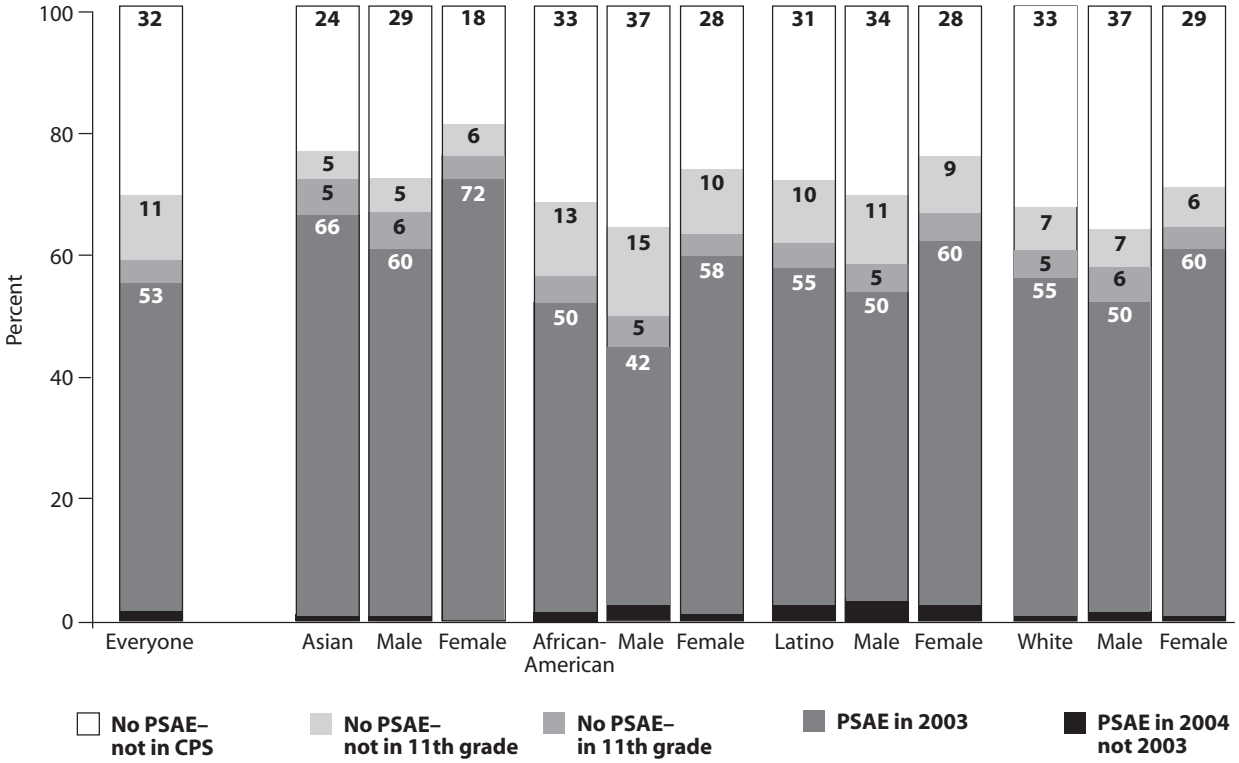


Figure 6: Percentage of 2000–01 CPS Freshmen Who Took the PSAE in 2003 and 2004



CHAPTER 3

CPS Students' Performance on the PSAE

Now we examine the performance of those CPS students who took the PSAE. In this section, we compare the scores of CPS students with students from the state as a whole. It would be preferable to compare CPS to the rest of Illinois, but data for the state includes Chicago, and it is not currently possible for us to remove CPS data from Illinois data. This limits the quality of the comparisons we can make.

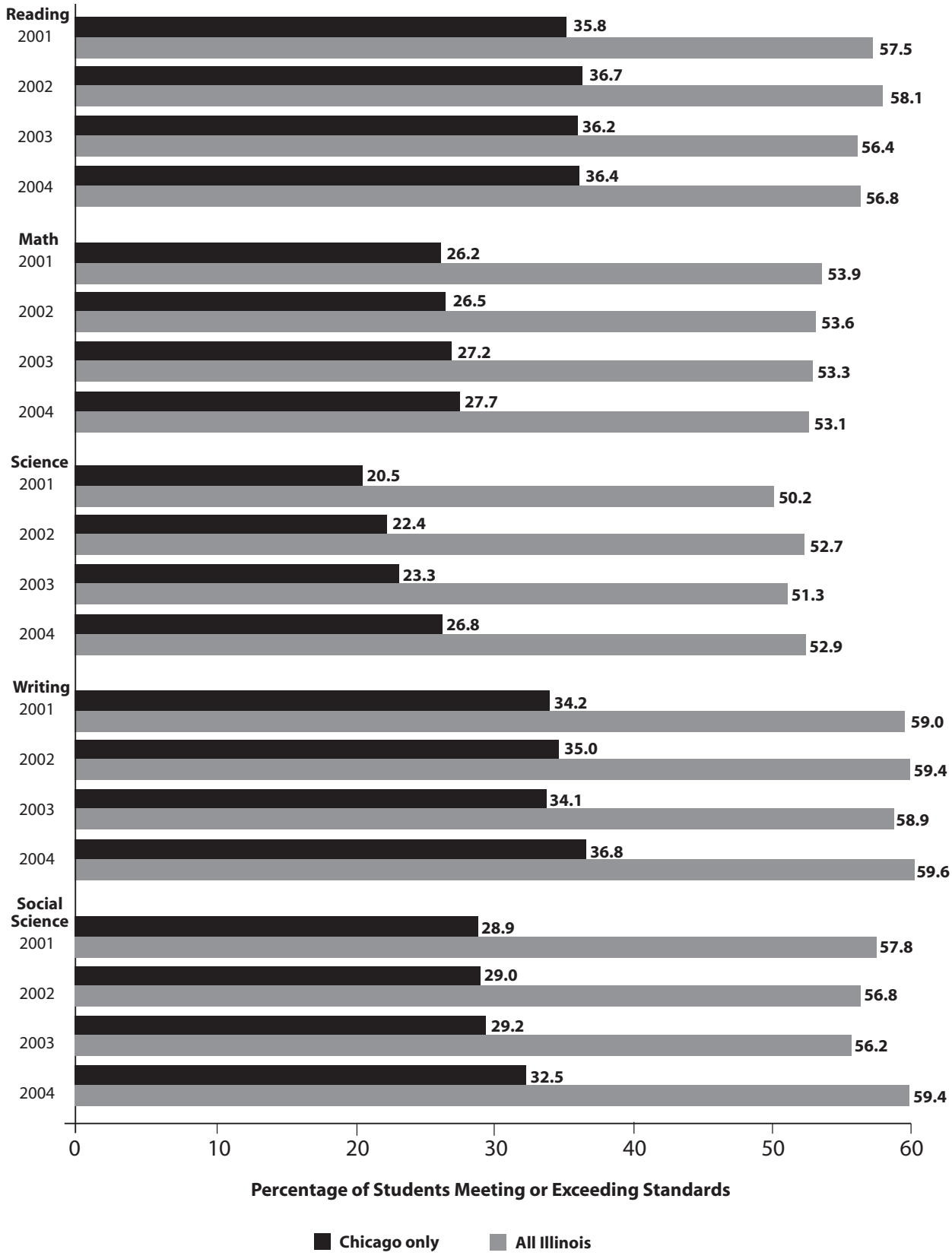
After the first PSAE administration in 2001, the statewide average score for each subject was set at 160 as a part of the process of creating the scale of scores (with a range from 120 to 200). In the first four years of testing, statewide average scores have remained near 160, but the CPS average has been much lower. CPS performance relative to the performance standards has remained relatively constant and low as well.

CPS students' scores look especially low given that only about 51 percent of first-time freshmen in 2000–01 remained in CPS long enough and completed one grade every year in order to take the PSAE during their junior year. CPS students perform at a much lower level on the PSAE than their peers in the rest of Illinois, as judged by the percentage of students meeting or exceeding standards shown in Figure 7. This figure shows that about half as many students met or exceeded standards in CPS as in all of Illinois. However, from 2001 to 2004 the percentage of students meeting or exceeding standards in math appears to have increased in CPS but decreased across Illinois.

There were small improvements in average PSAE scores in CPS from 2001 to 2004 (as shown in Figure 8) in all five subject areas, yet the changes in math and writing over this period were not statistically significant. The differences in science, social science, and reading were statistically significant and indicate a small improvement from 2001 to 2004. Because the mix of students changes each year, these changes in average scores could be the result of differences in the sample of students tested, rather than differences in student learning.

Another way to examine student performance is to examine the distribution of PSAE scores in each year. In Figures 9 and 10, vertical lines separate the performance levels, and each bar has a width of two scale score points, so that the bar furthest to the right includes PSAE scores of 199 and 200. Figure 9 shows that only about 36 percent of CPS students

Figure 7: PSAE Proficiency Rates, 2001–2004



met or exceeded standards in reading in 2004. Most CPS students received either the academic warning or below standards designation in reading in 2004. A floor effect (meaning that many students earned the lowest-possible score) is apparent in Figure 9, which shows that a large group of students received a score of 120 to 122. This effect seems to be caused by the scaling method used by ISBE, whereby a large group of average standardized scores are grouped into the PSAE score of 120. The floor effect also occurred in math and writing (not shown), but not in science or social science. The existence of a floor effect demonstrates that the PSAE does not measure low-performing students well in these subjects. Even the inclusion of WorkKeys scores does not provide much additional information about the performance of low-scoring students.

CPS students have performed most poorly in the science portion of the PSAE each year. As shown in Figure 10, more than 22 percent of CPS students scored at the academic warning level in 2004. Only 27 percent of CPS students met or exceeded standards in science, which is the lowest percentage of students to meet or exceed standards among the five subject tests. In addition, the average CPS score of 148.4 in science in 2004 is nearly one standard deviation below the statewide average of 160. The gap between CPS and Illinois average scores is largest in science, although this gap has shrunk somewhat since 2001.

Figure 8: Changes in Average PSAE Scores in CPS, 2001–2004

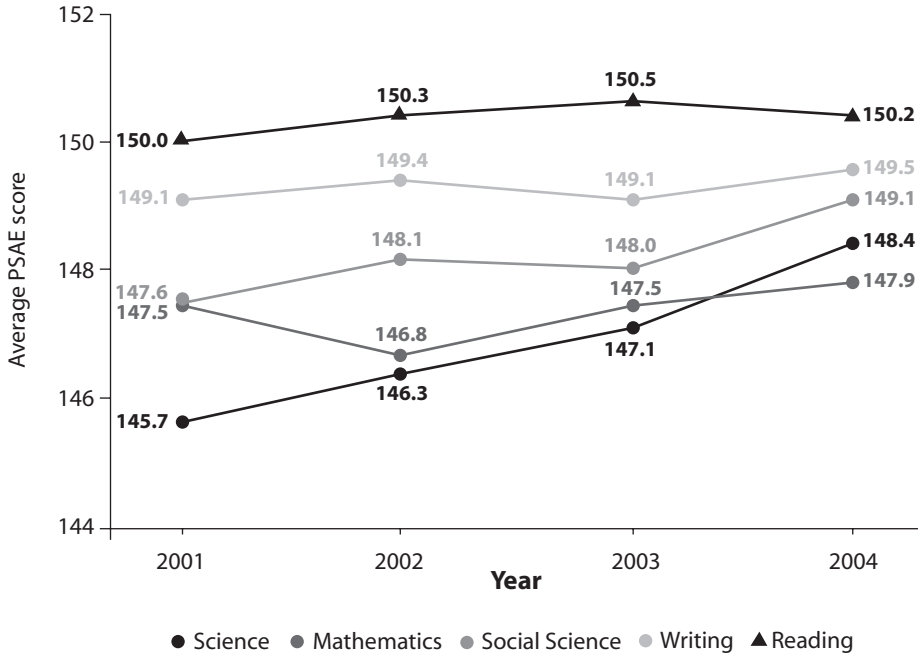


Figure 9: 2004 PSAE Reading Scores in CPS

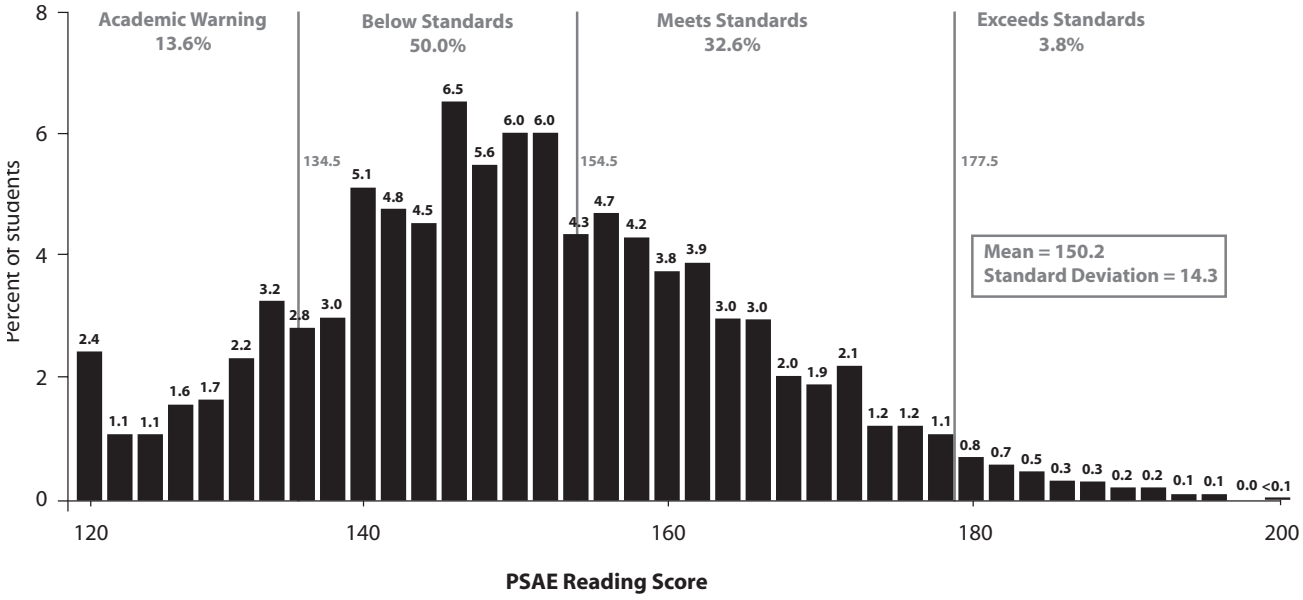
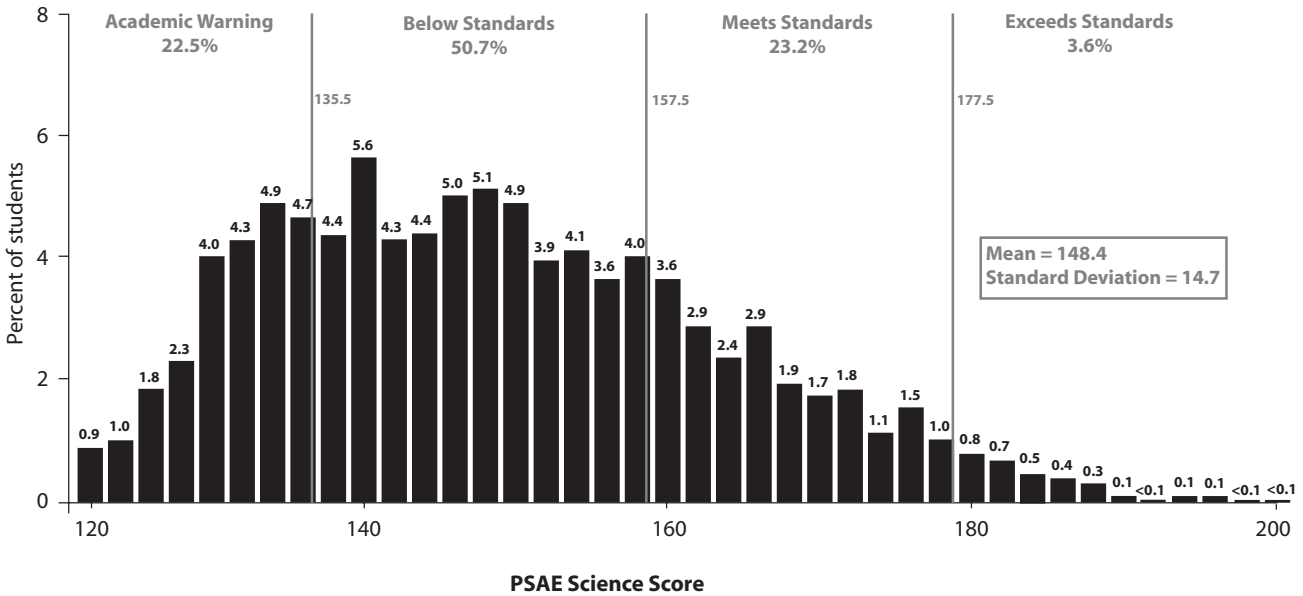


Figure 10: 2004 PSAE Science Scores in CPS



Note: In Figures 9 and 10, each bar represents two scale score points, so that the bar furthest to the right represents PSAE scores of 199 and 200. The bar furthest to the left is an exception—it represents scores 120, 121, and 122.

CHAPTER 4

Relationships among Tests

In addition to the performance of CPS students on various parts of the PSAE, we are also interested in the extent to which scores on the different PSAE component tests are related—for example, the relationship between scores on the ISBE writing test and the ACT English test. In earlier studies of the ISAT and ITBS, the Consortium found strong relationships between tests on the same subject, and to a lesser extent, between subjects on the same test.[18] We are examining these relationships to see how student performance varies across different tests.

How Are PSAE Tests of the Same Subject Related to Each Other?

Since four of the five subject areas tested by the PSAE are covered by a combination of two tests, we will examine the relationships among these tests. The PSAE combines two tests in a subject in order to increase the breadth of information covered in that area and ensure that progress on state learning standards is measured accurately. Correlation between two test scores is limited by the reliability of each test; therefore, the actual correlation should be compared to the maximum possible correlation that could be achieved.[19] One interpretation of test reliability is the likelihood that an individual receives the same score on different samples of items.

The correlation of students' scores on the ACT science test with scores on the ISBE-developed science test was 0.74 in 2004.[20] Therefore, 55 percent (that is, the square of the correlation, 0.74 x 0.74) of the variability in 2004 ACT science scores can be explained by ISBE-developed science scores (and vice versa). Given the reliability of each test, the maximum value of this correlation is 0.87,[21] and the maximum amount of variability in one test that could be explained by the other test is 76 percent (0.87 x 0.87). The relationship between these tests was similar in earlier years. The actual correlation is close to the maximum possible correlation. Therefore, a student's performance on one of these tests will likely be very similar to her performance on the other test, especially when the tests are administered consecutively (as they are in the PSAE).

The correlation of students' ACT English scores with ISBE-developed writing scores was 0.63 in 2004, with similar correlations from 2001 to 2003. Therefore, only 40 percent (0.63 x 0.63) of the variability in ACT English scores in 2004 can be explained by ISBE-developed writing scores, although this value could have a maximum of 77 percent.

Although they purport to measure different skills, there is a relationship between students' ACT and WorkKeys scores. We cannot present correlations of ACT scores with WorkKeys scores because the WorkKeys scores (though presented as numbers) are not on an internal scale. That is, the difference between WorkKeys scores of 3 and 4 is not necessarily the same as the difference between 4 and 5 and the difference between 5 and 6. As the score on one test increases, the other score also tends to increase. This relationship appears to be stronger in the math portions of the tests, because the variability of the ACT scores associated with each WorkKeys score is smaller for the math test than it is for the reading test. However, as Figure 11 shows, the range (and median) of ACT scores does not vary much for WorkKeys scores of less than 3, 3, and 4. This suggests that the ACT does not distinguish well among students at the lower end of the achievement spectrum, while the WorkKeys does differentiate achievement levels among these students.

How Are PSAE Tests of Different Subjects Related to Each Other?

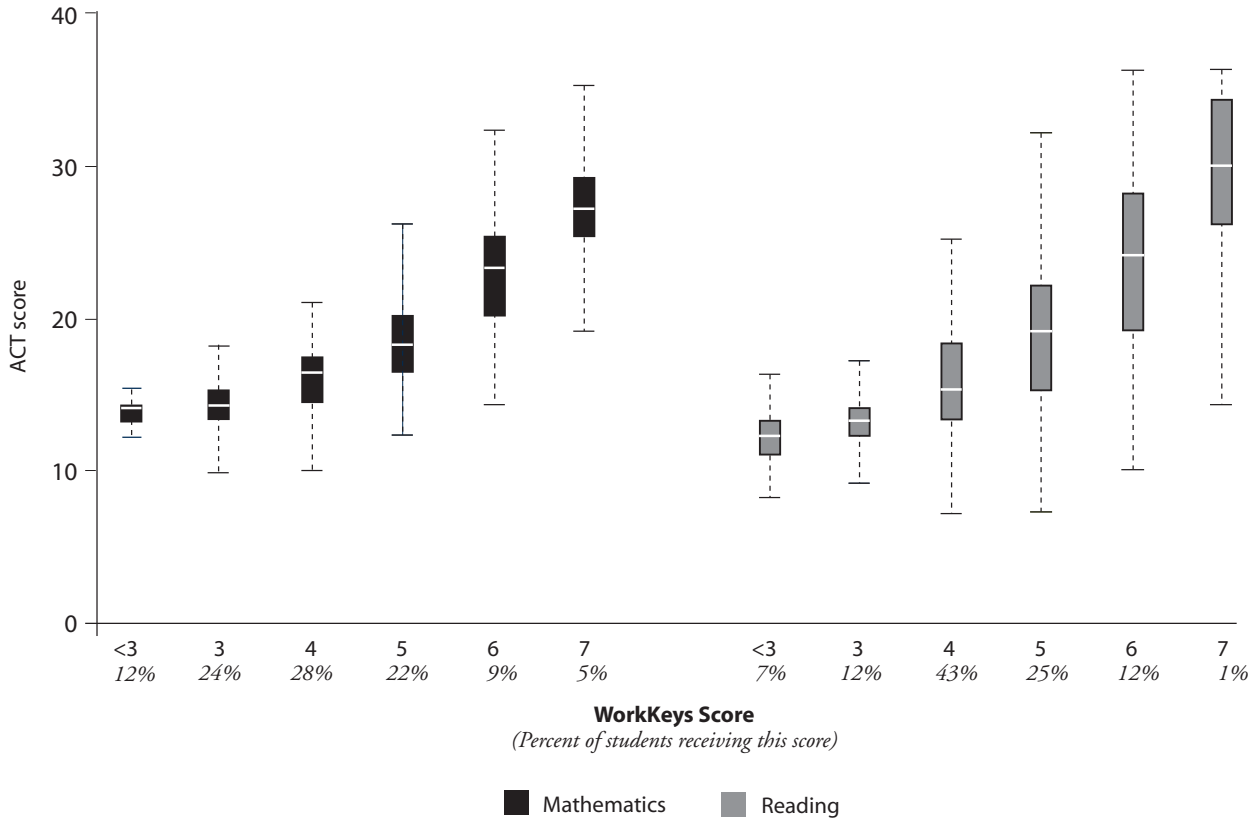
As is the case when comparing scores for two tests of the same subject matter, a low score on one subject test is usually associated with low scores on tests of other subjects, and a high score on one subject test is usually associated with high scores in other subjects. This relationship holds true for the scale scores and for the performance levels. Appendix Table 4 shows the disattenuated correlations for 2004. Disattenuated correlation is the value of the correlation adjusted for the reliability of each test. It is notable that all of the scores are highly correlated with each other.

We can see the strength of the relationship between PSAE reading and math scores in Figure 12, where we compare student performance levels on the two tests. In 2004, 20 percent of all students who took the math portion of the PSAE received the academic warning designation. Of these students, 44 percent were in the academic warning category and 53 percent were below standards on the PSAE reading test. Only 3 percent earned a meets standards designation and less than 1 percent earned the exceeds standards designation on the reading test. Of the students who took the math portion of the 2004 PSAE, 52 percent were below standards. Of these students, 64 percent were below standards in reading, 27 percent earned a meets standards designation, 9 percent received the academic warning designation, and less than 1 percent were in the exceeds standards category. The higher achievement levels showed similar consistency, as 67 percent of the students who earned a meets standards designation in math earned the same designation in reading, and 53 percent of the students who exceeded standards in math also exceeded standards in reading. This pattern of consistent performance generally holds across pairs of subjects and years. Of particular concern is the fact, illustrated in Figure 12, that 5,321 students, the single largest group, earned below standards designations in both reading and math. Furthermore, 9,165 students did not meet standards on either test. That is, they were at the academic warning or below standards levels on both the math and reading tests.

Several factors contribute to the high correlations discussed in this section. First and foremost, the content is similar (though not identical) in different tests of the same subject, so one would expect students to display the same level of performance when tests purport to measure the same material. Even when the subjects tested are different, there is often a common thread throughout the tests. For example, reading ability is important on nearly all tests, so to a certain extent the correlations are driven by this common factor.

It is important to note, however, that high correlations do not necessarily mean that tests are measuring the same knowledge and skills. A number of other factors play a role here. For example, in general, students who do well in one subject in school tend to do well in others.

Figure 11: ACT Versus WorkKeys Scores in CPS Mathematics and Reading, 2004



How to Read a Box Plot and Why We Use Them

The box plot details the relative frequency of positive and negative school reports. Each box encloses the middle 50 percent of the schools. The lines, called "whiskers," extending up and down from the box, show the range of scores for the top and bottom quartile schools. These are the highest and lowest performing school on each particular scale. Within each profile, the scales are centered on the systemwide average for the schools that participated in the survey.

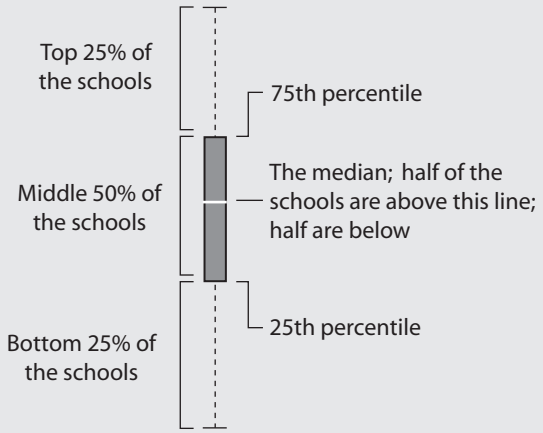
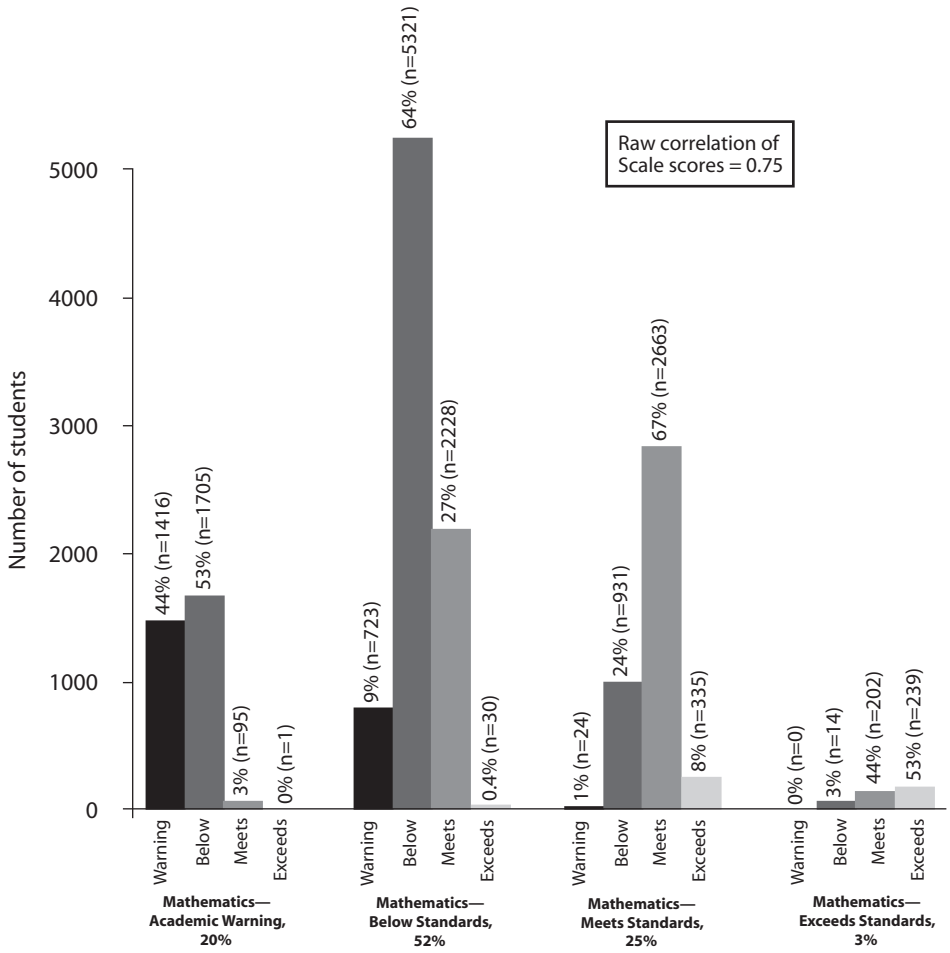


Figure 12: 2004 PSAE Reading Proficiency, Broken Down by Mathematics Proficiency (for CPS Students)



PSAE Reading and Proficiency, Broken Down by Mathematics Proficiency

Many students who know a lot of math also know a lot of science, not because the content is the same, but because they have learned both. A related factor is students' opportunity to learn. Schools with talented teachers and high-quality curriculum and instruction are likely to have strengths across subject matters; therefore students who are exposed to more rigorous math instruction also are exposed to more rigorous science instruction. Other factors, including students' motivation and test-taking skills, contribute to the high correlations.

Test-score correlations are frequently examined in the education literature to assess the validity of a test. Some have argued that high correlations between tests of the same subject mean that the second test does not provide much additional information about students' knowledge and skills and may be unnecessary. However, the utility of a given test should be judged based on the importance of its content and its alignment with learning standards, in addition to its correlations with other tests.

CHAPTER 5

Performance Comparisons

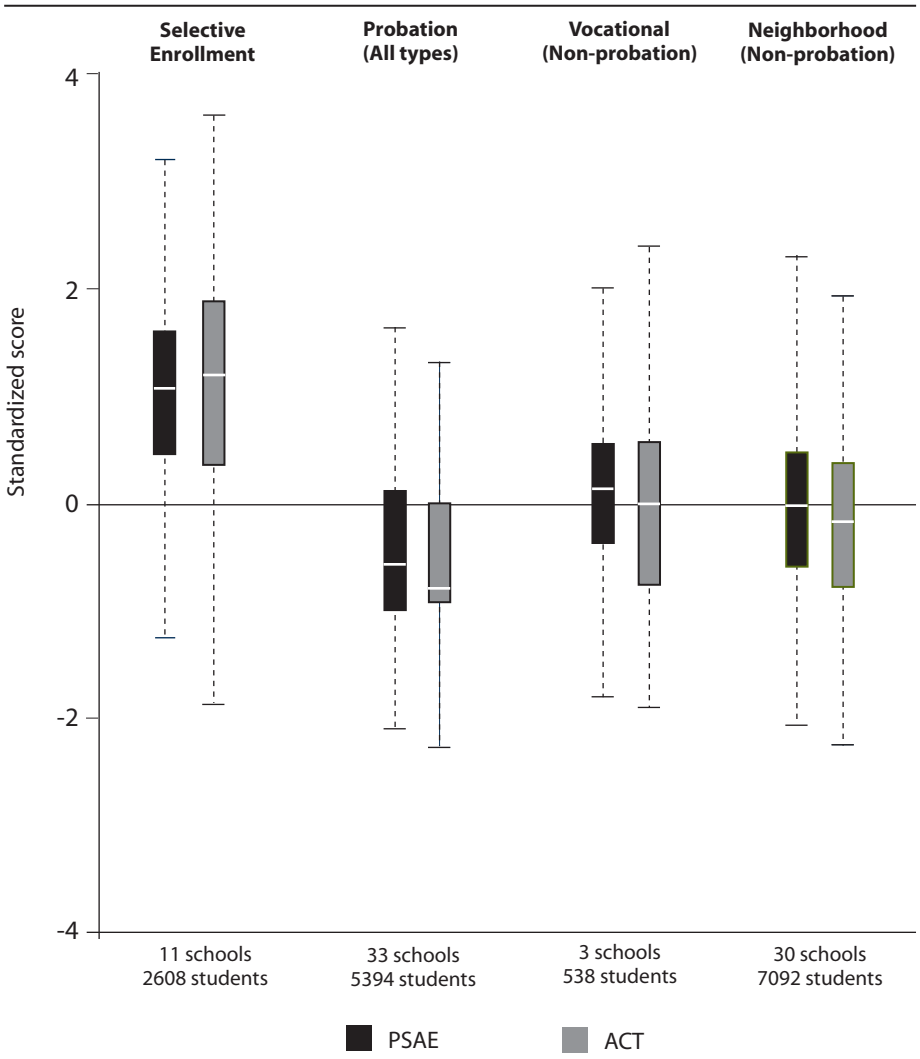
We now move to comparisons of PSAE scores for different types of schools. We also discuss the scores of students grouped by race/ethnicity, gender, and income both in CPS and for Illinois as a whole. We conclude by illustrating the performance of students who remained in CPS from first grade until taking the PSAE.

Differences between Types of Schools

Earlier we demonstrated that the WorkKeys tend to provide more information about lower-performing students and that the ACT is more descriptive of the achievement of higher-performing students. Students' scores on the ACT and WorkKeys cannot be compared directly because WorkKeys scores are categorical, meaning that they are expressed as a category that is equivalent to a range of scores rather than as a particular number. Therefore, we examine standardized PSAE and ACT scores in order to determine the relative impact of the ACT and WorkKeys on schools' PSAE performance. If a school's standardized ACT score is higher than its standardized PSAE score, it means the school performed better on the ACT than on the WorkKeys. If the two standardized scores are similar, then the school performed equally well on both the ACT and WorkKeys (as is the case for selective enrollment schools). If the standardized PSAE score is higher than the standardized ACT score (as it is, albeit slightly,

African-American, Latino, and low-income students tend to meet or exceed state standards at about the same rate as their counterparts in the rest of the state. White students and non-low-income students tended to perform at a lower level in Chicago than in the state as a whole.

Figure 13: Standardized PSAE and ACT Reading Scores by Type of School, 2004



for schools on probation, vocational schools, and neighborhood schools), then the school's performance on the WorkKeys was better than its performance on the ACT.

Figure 13 shows that average standardized PSAE scores are higher than ACT scores for schools on probation, which demonstrates that the ACT does not measure performance at the low end of the achievement spectrum as well as the WorkKeys. Selective-enrollment schools' scores are similar on the two tests. This pattern is expected, since students in selective-enrollment high schools are likely to be college oriented and the ACT was developed with college-bound students in mind. Finally, students in non-probation vocational and neighborhood schools appear to perform somewhat better on the WorkKeys reading test than on the ACT reading test, since their standardized PSAE scores are higher than their standardized ACT scores.

Group differences

Significant performance differences exist between students grouped according to race/ethnicity, gender, and income. These group differences among CPS students are generally similar to differences observed among groups of students across Illinois. African-American, Latino, and low-income students tend to meet or exceed state standards at about the same rate as their counterparts in the rest of the state. White students and non-low-income students tended to perform at a lower level in Chicago than in the state as a whole. These trends are shown in Figures 14, 15, and 16. Such comparisons cannot be made for 2004 because the state did not provide PSAE scores disaggregated by race/ethnicity or income level. Chicago cannot be disaggregated from the state as a whole due to the way figures are reported by ISBE. Therefore information about Illinois includes data from Chicago.

Figure 14: PSAE Reading Proficiency, 2001–2003

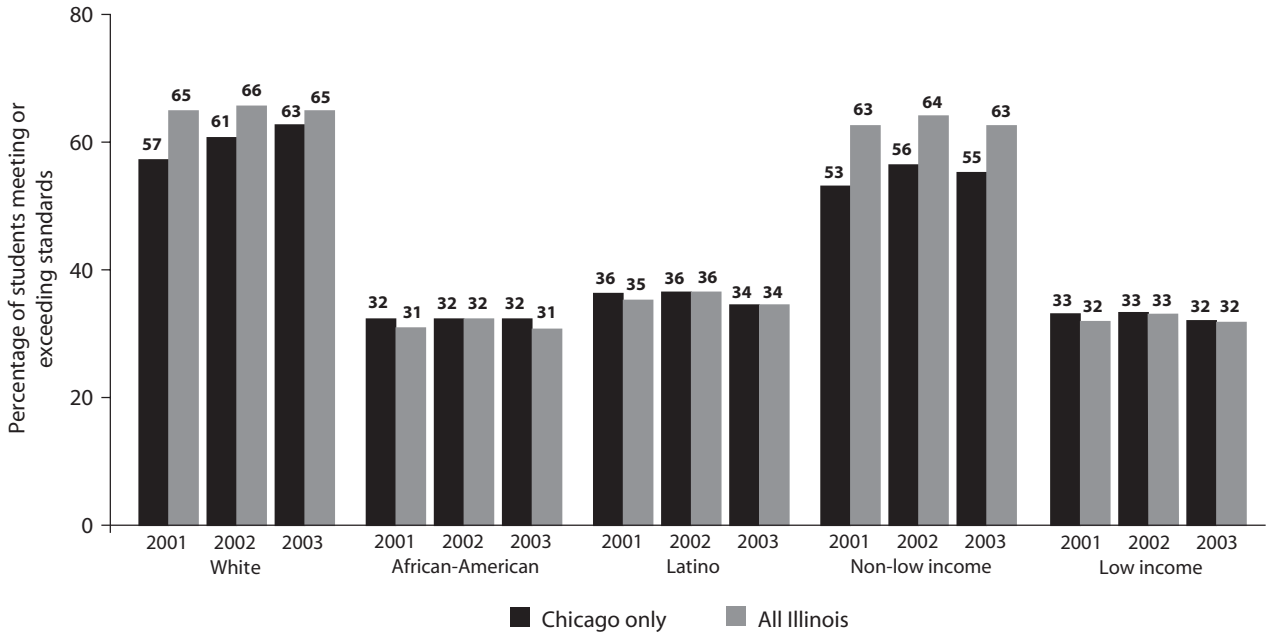


Figure 15: PSAE Mathematics Proficiency, 2001–2003

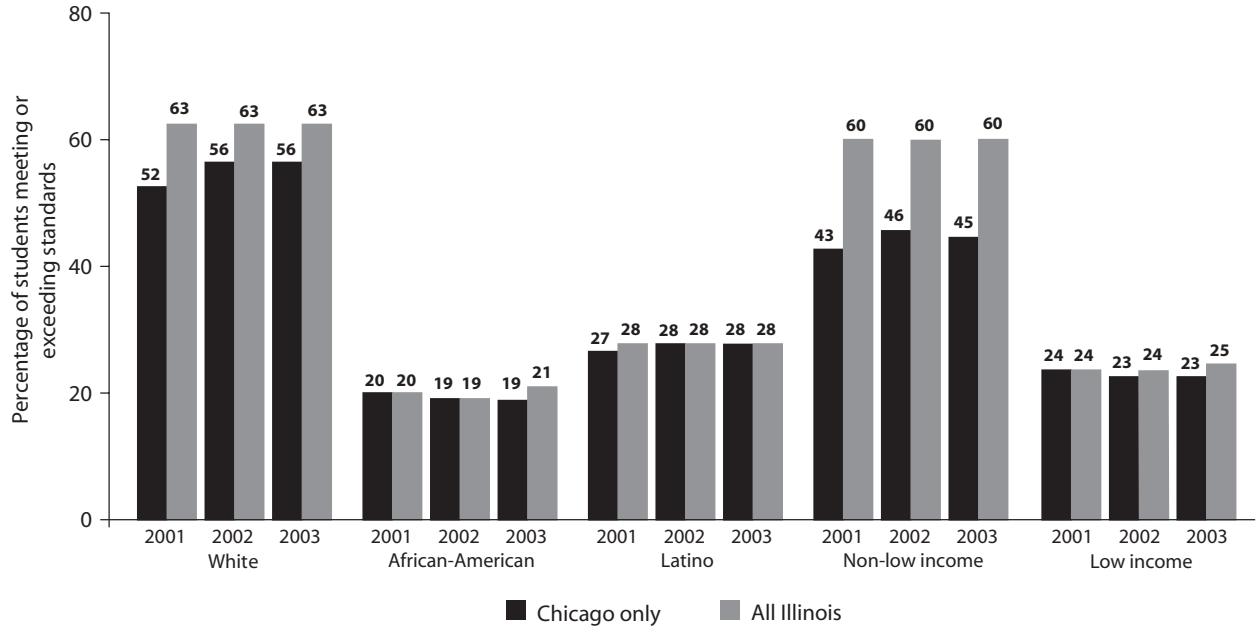
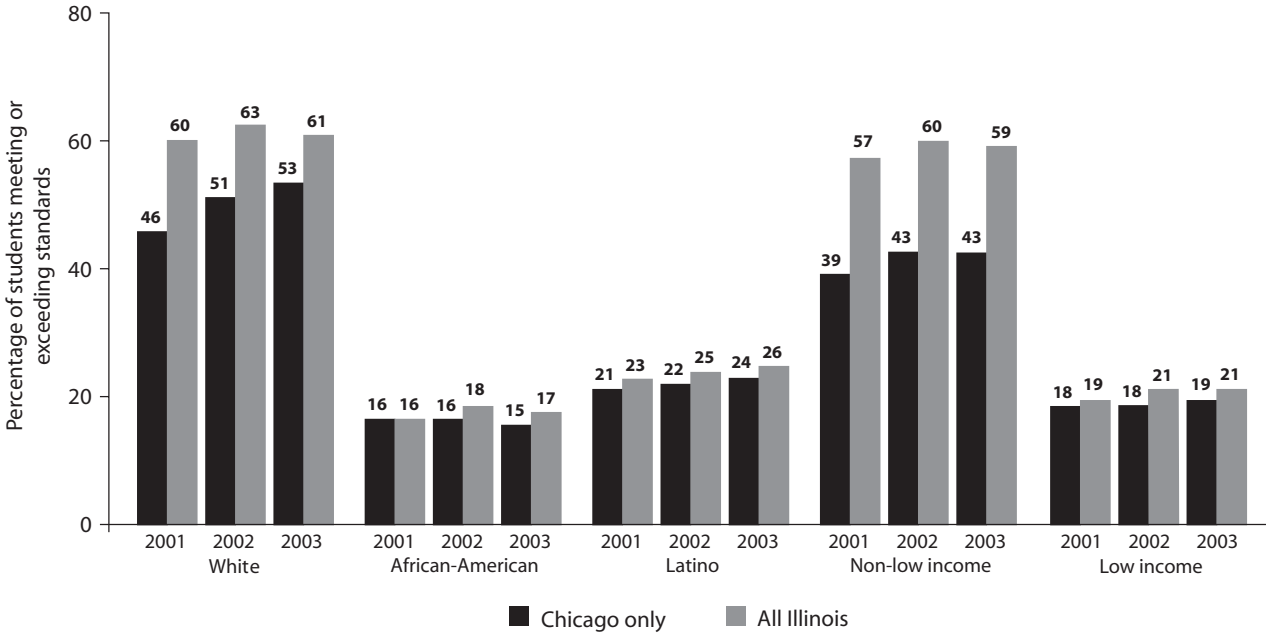


Figure 16: PSAE Science Proficiency, 2001–2003



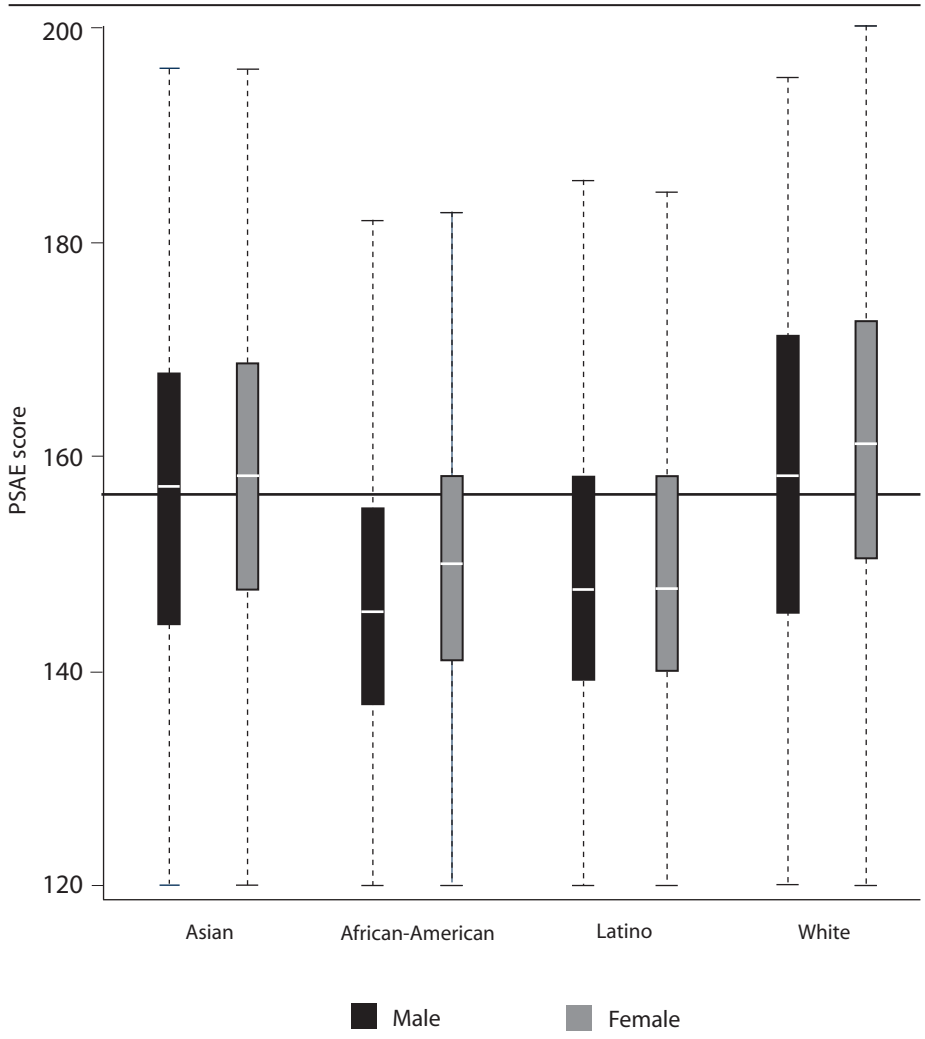
Across all racial and ethnic groups, there are differences in average PSAE scores by gender in 2004, as shown in Figures 17 and 18.

The extent of these differences varies across years. Female students had higher scores on the writing section of the PSAE than male students of the same racial group from 2001 to 2004, as indicated in Appendix Table 5. Female students had higher scores on the reading section of the PSAE than male students of the same racial/ethnic group each year, with the exception of Asian students in 2002 and 2003 and white students in 2002, when the scores were similar. Male students had higher PSAE social science scores than female students each year in all racial/ethnic groups except African-American students, among whom scores were similar for male and female students. Male students of all races tended to have higher scores in math and science than female students, with a few exceptions where scores were similar.

Performance of Students Who Were in First Grade in Spring 1994

Earlier sections of this report questioned whether the PSAE is really a good indicator of achievement in CPS high schools given the system's high attrition rates. In other words, school-by-school differences in PSAE scores may be the result of a selection effect, by which some students remove themselves from the tested group. In some respects, the students who take the PSAE in CPS are survivors who may have remained in school longer due to greater academic success. To illustrate this problem, consider the progress of a group of student beginning with first grade and continuing through high school. The majority of students who began first grade in a Chicago public elementary school did not appear in a Chicago high school 10 years later when they would have been in eleventh grade. This is due primarily to families leaving Chicago and CPS,

Figure 17: 2004 PSAE Reading Scores by Race and Gender
(for CPS Students)

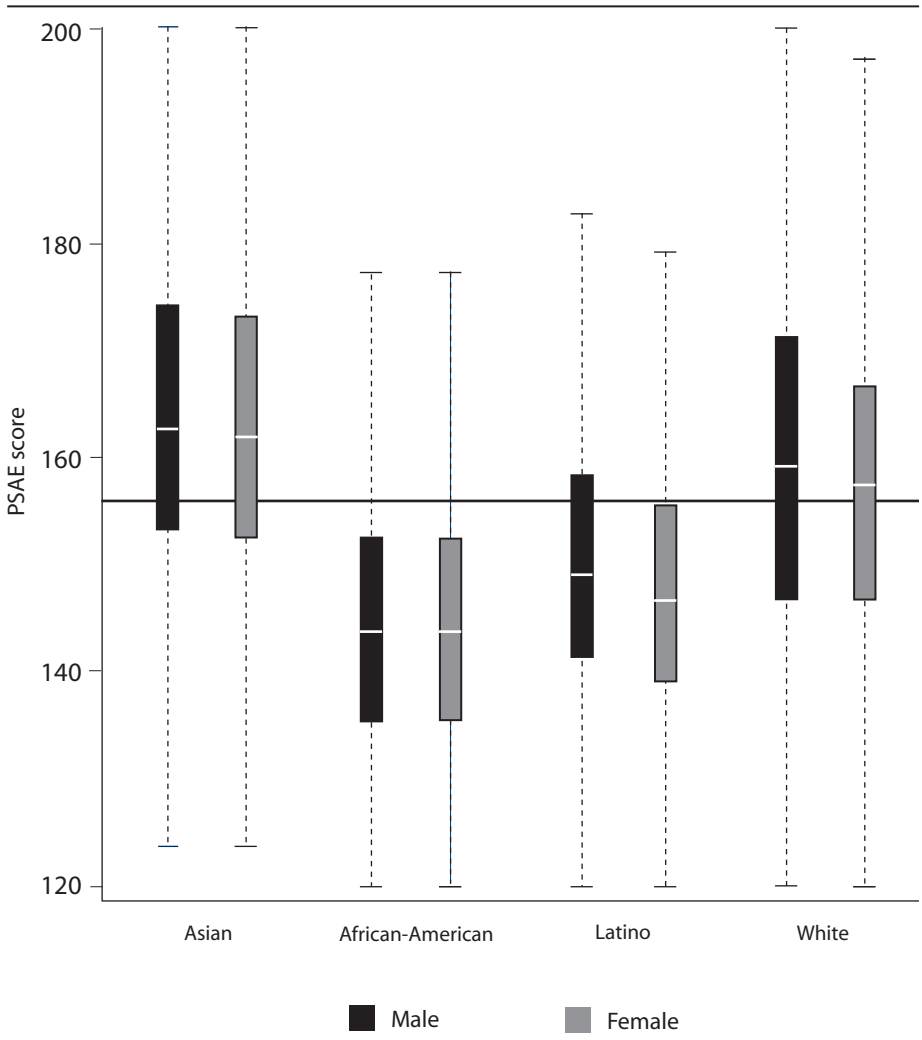


but other factors, such as students remaining in CPS but falling off track and students dropping out of school, also play a large role.

In spring 1994, 35,370 students were enrolled as first-graders in CPS. If we exclude students who transferred out of CPS from our calculations, we find that just 42 percent of the remaining group of students took the PSAE ten years later in 2004. Of these students, only 18 percent met or exceeded standards in reading or math, and just 9 percent met or exceeded standards in both subjects. The 1992 and 1993 first-grade cohorts have similar results. Those cohorts reached grade 11 in 2002 and 2003, respectively. See Appendix Tables 6 and 7 for more details.

The group of 8,698 students who progressed continuously through CPS (regardless of the number of schools they attended) from first grade in 1994 to eleventh grade in 2004 performed better on the PSAE than those who were not continuously enrolled in CPS from first through eleventh grade. The continuously enrolled students were less likely to have PSAE scores at the academic warning level in reading (9 versus 18 percent) and in math (16 versus 25 percent), and were more likely to meet or exceed standards on either test (40 versus 32 percent in reading, 30 versus 25 percent in math). However, continuously enrolled students were still more likely to earn scores that were in the below standards category than in any other level on the reading and math portions of the PSAE. That is, the performance of consistently enrolled students was better than that of students who left and returned at least once between first and eleventh grade; however, the majority of the continuously enrolled students received scores that were below standards in reading and math.

Figure 18: 2004 PSAE Mathematics Scores by Race and Gender
(for CPS Students)



Very few students who start school in CPS—about 9 percent of those who do not transfer out—are able to meet the Illinois Learning Standards in reading and math by the time they reach eleventh grade. This result is a consequence of a combination of factors: high rates of attrition from CPS, high rates of grade repetition within CPS, and poor performance among the students who remain in CPS until the eleventh grade. The high rates of attrition and grade repetition make it difficult to judge trends in test scores, since the average score of each group of eleventh-graders is affected by who has reached eleventh grade in that year. The limited progress in PSAE scores over the last several years may be the result of decreasing dropout rates rather than lack of progress in student learning. These factors limit our ability to assess student learning progress in Chicago public high schools.

CONCLUSION

Student performance on the PSAE, both in CPS and in Illinois as a whole, has remained at about the same level since the test's first administration in 2001. Average PSAE performance in CPS is far below state standards in all subjects. Performance in science is weakest, although it is one of the two subjects where improvements are evident. The poor performance of CPS students is of even greater concern when we consider that the students who take the PSAE are the "survivors." Their presumably weaker peers have already dropped out.

About 80 percent of CPS eleventh-graders took the PSAE in 2001 (83 percent were tested including students who took alternative assessments); this value rose to about 85 percent in 2004 (91 percent including students who took alternative assessments). The percentage of CPS students that took the PSAE varies across race and gender combinations.

In CPS, about 53 percent of students who entered ninth grade in fall 2000 took the PSAE within four years. This percentage varies by race and gender, with a larger proportion of female students taking the test. The low rate of PSAE participation is affected by both retention policies and dropout patterns.

Students' test scores are highly correlated across subject areas. CPS students' scores were slightly higher in 2004 than in 2001 in all subjects. Scores vary by race and gender, with female students tending to have higher scores in reading and writing, while male students tend to score higher in math, science, and social science. Minority students and students from low-income backgrounds have met or exceeded standards at about the same rate in CPS as in Illinois as a whole, but these groups' scores are much lower than those of non-minority students or those who are not from low-income backgrounds. Science and math should be areas of specific concern, due to the relatively smaller percentages of students who meet or exceed standards in CPS in these subjects.

In a forthcoming report, the Consortium will analyze the relationships among students' course-taking, grades, and PSAE scores in math and science to determine whether students are adequately prepared for this test and whether this preparation differs across schools. There are many students for whom good performance in the classes required for graduation did not yield good PSAE results, so the forthcoming report will also examine the reasons for poor PSAE performance despite apparently satisfactory classroom performance. The report will use a hierarchical model for this data, controlling for a student's school, courses passed, grades, bilingual and special education status, and scores on the ISAT in eighth grade.

APPENDIX—SUPPLEMENTAL TABLES

Table 1: Schedule for 2004 PSAE

Day 1 *Wednesday, April 28, 2004*	ACT English—45 minutes ACT Mathematics—60 minutes *[required 15-minute break]* ACT Reading—35 minutes ACT Science—35 minutes
Day 2 *Thursday, April 29, 2004*	ISBE-Developed Writing—40 minutes ISBE-Developed Science—40 minutes *[required 15-minute break]* WorkKeys *Applied Mathematics*—45 minutes WorkKeys *Reading for Information*—45 minutes *[required 15- to 60-minute break]* ISBE-Developed Social Science—60 minutes

Table 2: Subjects Covered by the PSAE from 2001 to 2004

Test	Test	Number of Questions	Type of Questions
Reading	ACT Reading	40 multiple choice	10 questions for each of four 750-word passages: fiction, non-fiction, science, social science
	WorkKeys Reading	33 multiple choice	15 reading samples (memos, policy statements, procedures, regulations, business texts), length from 50 to 300 words
Math	ACT Math	60 multiple choice	Six subject areas: pre-algebra, elementary algebra, intermediate algebra, coordinate geometry, plane geometry, trigonometry
	WorkKeys Math	33 multiple choice	Workplace-related problems: read the problem, gather information, make calculations
Writing	ACT English	75 multiple choice	40 questions on conventions, 35 on rhetorical skills, from five passages
	ISBE Writing	1 essay	Expository or persuasive prompt, scored on focus, support, organization, conventions, integration
Science	ACT Science	40 multiple choice	Seven sets of scientific information: reading passages and data summaries from biology, chemistry, physics, and earth and space sciences
	ISBE Science	45 multiple choice	Data summaries and short prompts
Social Science	ISBE Social Science	65 multiple choice	Disciplines (political, economic, and social systems), geography, history

Table 3: Score Reporting

Test	Student receives . . .	School receives . . .
PSAE	PSAE scores, with school, district, and state percentiles and average scores	School, district, and state average scores
PSAE	PSAE performance categories, with percentage of students in school, district, state at each performance category	PSAE performance categories for all students, and for each subgroup of students
ACT	ACT scores, with school, district, and state averages	School, district, and state average scores
WorkKeys	WorkKeys scores, with school, district, and state averages	School, district, and state average scores
ISBE-Developed	ISBE-developed test scores and subscores, with school, district, and state averages	School, district, and state average scores and subscores

Table 4: Disattenuated Correlations Among PSAE Subjects in 2004
(for CPS Students)

Disattenuated Correlation (2004)	**Reading**	**Mathematics**	**Writing**	**Science**	**Social Science**
Reading	1	0.87	0.89	0.93	0.88
Mathematics		1	0.80	0.89	0.81
Writing*			1	0.83	0.78
Science				1	0.89
Social Science					1

* Because the ISBE-developed part of the writing test is an essay, and is scored by readers, there can be no calculation of its reliability, which would be necessary to determine the disattenuated correlation. Therefore, we must assume that this part of the test is perfectly reliable or completely unreliable, because the reliability of an essay question cannot be calculated. We assume perfect reliability here.

Table 5: PSAE Scores by Race and Gender

	Average PSAE Reading Score											
	White			African-American			Latino			Asian		
	All	Male	Female	All	Male	Female	All	Male	Female	All	Male	Female
2001	157.3	156.3	158.2	147.9	145.6	149.4	148.9	148.1	149.5	158.0	156.4	159.5
2002	159.3	159.1	159.4	148.0	145.7	149.5	149.4	148.5	150.2	157.3	157.1	157.4
2003	159.9	158.5	161.1	148.3	146.2	149.7	149.3	148.0	150.5	158.4	158.1	158.8
2004	159.4	157.9	160.8	148.2	146.0	149.7	149.1	148.8	149.3	157.8	157.0	158.5

	Average PSAE Mathematics Score											
	White			African-American			Latino			Asian		
	All	Male	Female	All	Male	Female	All	Male	Female	All	Male	Female
2001	155.7	156.7	154.6	143.9	143.7	144.0	147.4	148.3	146.6	161.2	161.0	161.4
2002	156.6	158.7	154.7	142.8	142.8	142.8	147.1	148.2	146.3	159.1	160.7	157.5
2003	157.1	157.9	156.4	144.0	144.4	143.6	147.5	148.1	146.9	160.0	161.1	159.1
2004	158.1	158.9	157.5	144.0	143.9	144.0	148.2	149.2	147.2	162.6	163.0	162.2

	Average PSAE Writing Score											
	White			African-American			Latino			Asian		
	All	Male	Female	All	Male	Female	All	Male	Female	All	Male	Female
2001	157.1	154.9	159.4	146.4	143.8	148.1	148.6	146.9	150.0	158.0	155.6	160.1
2002	159.0	157.4	160.5	146.6	143.3	148.8	148.7	146.5	150.5	158.3	156.9	159.6
2003	159.6	156.5	162.2	146.4	144.0	148.2	148.0	145.7	150.0	159.0	156.6	161.2
2004	160.0	157.2	162.7	146.6	143.7	148.6	148.6	146.8	150.1	160.5	157.5	163.3

	Average PSAE Science Score											
	White			African-American			Latino			Asian		
	All	Male	Female	All	Male	Female	All	Male	Female	All	Male	Female
2001	155.7	157.1	154.4	141.9	142.2	141.7	145.4	146.9	144.2	158.2	158.8	157.7
2002	157.4	159.3	155.7	142.6	142.5	142.6	146.2	147.1	145.4	157.1	159.1	155.3
2003	158.4	159.1	157.8	143.1	143.2	143.0	147.2	147.9	146.6	159.5	161.0	158.1
2004	159.6	160.0	159.1	144.5	144.0	144.8	148.5	149.7	147.4	161.4	161.5	161.3

	Average PSAE Social Science Score											
	White			African-American			Latino			Asian		
	All	Male	Female	All	Male	Female	All	Male	Female	All	Male	Female
2001	156.5	158.6	154.4	144.0	144.6	143.6	147.8	149.7	146.1	158.2	159.7	156.7
2002	158.0	160.4	155.8	144.3	144.5	144.1	148.9	150.2	147.8	157.0	158.9	155.2
2003	157.9	159.3	156.7	144.4	144.8	144.2	148.4	149.5	147.4	158.3	159.9	156.9
2004	159.4	160.5	158.4	145.5	145.9	145.2	149.6	151.3	148.1	158.9	159.4	158.5

Table 6: Status of Students in First Grade in 1994

	Grade 1 in 1992		Grade 1 in 1993		Grade 1 in 1994	
Outcome after 10 Years	**Students**	**Percentage**	**Students**	**Percentage**	**Students**	**Percentage**
Not in CPS—transfer	10,459	30	10,663	31	10,972	31
Not in CPS—dropout	3,934	11	3,664	11	3,497	10
Not in CPS—other	381	1	338	1	254	1
No PSAE, not in 11th or 12th grades *(but still in CPS 9 or 10 years later)*	9,006	26	8,754	26	9,345	26
No PSAE—in 11th grade	1,214	4	1,015	3	1,072	3
No PSAE—in 12th grade	24	<1	16	<1	24	<1
Took PSAE, below standards or academic warning *(Reading)*	5,728	17	6,003	18	6,344	18
Took PSAE, met or exceeded standards *(Reading)*	3,665	11	3,848	11	3,862	11
Total in Grade 1	**34,411**	**100**	**34,301**	**100**	**35,370**	**100**

Table 7: Grade Progression for Students in First Grade in 1994

Grades for Students Still Enrolled in CPS											
	Year										
Grade	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**
1	35,370	1,875	43								
2		30,275	2,337	89							
3		143	27,907	3,015	714						
4			196	25,607	3,061	599	86				
5			21	236	23,627	3,593	575	58			
6				23	242	22,228	4,895	956	36		
7					16	283	19,704	4,029	341	13	
8					10	35	345	19,379	4,603	208	13
9						7	12	317	16,968	6,167	1,397
10							6	10	219	13,128	4,704
11								5	9	192	11,077
12									6	11	231
Other		183	220	288	265	140	30	26	11	6	6
Total	**35,370**	**32,536**	**30,724**	**29,258**	**27,935**	**26,885**	**25,653**	**24,780**	**22,193**	**19,725**	**17,428**

Endnotes

[1] ACT, Inc. and Illinois State Board of Education, 2003.

[2] Illinois State Board of Education, 2002.

[3] In 1996 American College Testing changed its name to ACT, Inc. Throughout this report the term "ACT" is used to refer to the test. "ACT, Inc." is used to refer to the company.

[4] Illinois State Board of Education, 2000.

[5] Illinois State Board of Education, 2005.

[6] Rado, 2005.

[7] Illinois State Board of Education, 2001.

[8] ACT, Inc. and Illinois State Board of Education, 2003.

[9] ACT, Inc., n.d.

[10] The standard deviation is a measure of the spread of the distribution of a quantity, such as test scores. It describes the average amount by which a group of scores differs from the mean of that group. Standard deviation is greater when scores differ widely.

[11] ACT, Inc. and Illinois State Board of Education, 2003.

[12] Scores were transformed to a pseudo-PSAE scale (that is, one different from the PSAE's 120 to 200 scale) for this process so that panelists could not compare their proposed PSAE cutoff scores to ISAT cutoff scores.

[13] IRT is a method of analysis that many statisticians use to determine an individual's unobserved ability based upon his or her responses to test questions.

[14] Chicago Public Schools, 2004.

[15] In order for these calculations to be valid, all subgroups must contain at least 40 students.

[16] Illinois State Board of Education, 2003.

[17] Illinois State Board of Education, 2001.

[18] Consortium on Chicago School Research, 2003.

[19] The maximum correlation of two test scores is equal to the square root of the product of the two tests' reliabilities.

[20] Correlation is a measure of the strength of the relationship between two quantities, taking values between -1 and 1. Correlations near 1 suggest a strong relationship, so that an increase in the value of one of the variables is associated with an increase in the value of the other variable. Correlations near 0 suggest a weak or nonexistent relationship, and correlations near -1 suggest a strong negative relationship, so an increase in the value of one variable is associated with a decrease in the other. We use the words "associated with" because correlation does not imply a causal relationship between the variables.

[21] The maximum correlation of two test scores is equal to the square root of the product of the two tests' reliabilities. The reliability of the ACT science test is 0.85, and the reliability of the ISBE science test is 0.89, so the maximum value of the correlation of these two tests is the square root of (0.85 x 0.89), or 0.87.

Works Cited

ACT, Inc., n.d. ACT WorkKeys: Frequently asked questions. Online at www.act.org/workkeys/overview/faq.html.

ACT, Inc. and Illinois State Board of Education. 2003. Prairie State Achievement Examination Technical Manual, 2001 Baseline Testing Cycle with 2002 Testing Cycle Supplement. Springfield: State of Illinois.

Chicago Public Schools. 2004. Accountability system for all Chicago Public Schools. In *Chicago Public Schools Policy Manual.* Online at policy.cps.k12.il.us/documents/302.6.pdf.

Consortium on Chicago School Research, 2003. Research Data Brief: How Do They Compare? ITBS and ISAT Reading and Mathematics in the Chicago Public Schools, 1999 to 2002. Chicago: Consortium on Chicago School Research. Online at www.consortium-chicago.org/publications/pdfs/p58.pdf.

Illinois State Board of Education. 2005. 2005 Prairie State Achievement Examination Test Administration Schedule. Online at www.isbe.net/assessment/PSAE2005schedule.htm.

Illinois State Board of Education. 2003. Collection and use of data for AYP calculations to fulfill requirements of the federal "No Child Left Behind" legislation. Online at www.isbe.net/assessment/PDF/DataAccuracy.pdf.

Illinois State Board of Education. 2002. Prairie State Achievement Examination (PSAE) Questions and Answers, December 2002 Update. Online at http://www.isbe.net/assessment/PSAEQandA.pdf.

Illinois State Board of Education. 2001. Prairie State Achievement Examination (PSAE) Teacher's Handbook, 2001-2002. Springfield: State of Illinois.

Illinois State Board of Education. 2000. PSAE Q&A. [Prairie State Achievement Exam Questions and Answers] Online at www.isbe.net/news/2000/psaeqa.doc.

Rado, Diane. 2005. State revives writing exam: Test was axed in '04 to balance budget. *Chicago Tribune*. June 23: p. 1.

This report was edited, designed, and produced by:
Sandra Jennings, Associate Director for Publications and Communications
Lura Forcum, Publications Editor/Research Assistant
Cover photo by John Booz

About the Author

Stephen M. Ponisciak is a Senior Research Analyst at the Consortium on Chicago School Research, where he is analyzing the PSAE, comparing the Pennsylvania System of School Assessment and TerraNova in the Philadelphia school district, and working on value-added models for school performance. Stephen earned his PhD from the Institute of Statistics and Decision Sciences at Duke University, where he focused on Bayesian analysis of teacher effectiveness.

Consortium on Chicago School Research

Mission

The Consortium on Chicago School Research aims to conduct research of high technical quality that can inform and assess policy and practice in the Chicago Public Schools. By broadly engaging local leadership in our work, and presenting our findings to diverse audiences, we seek to expand communication among researchers, policy makers, and practitioners. The Consortium encourages the use of research in policy action, but does not argue for particular policies or programs. Rather, we believe that good policy is most likely to result from a genuine competition of ideas informed by the best evidence that can be obtained.

Founded in 1990, the Consortium is located at the University of Chicago.

Directors

$6.00
ISBN 0-9726035-7-3
51000>

Consortium on Chicago School Research
1313 East 60th Street, Chicago, IL 60637
773-702-3364 fax -773-702-2010
www.consortium-chicago.org