



E&R Report No. 10.12

June 2010

END-OF-GRADE (EOG) MULTIPLE-CHOICE TEST RESULTS, 2008-09

Author: Brad McMillen

ABSTRACT

In 2008-09, results from End-of-Grade (EOG) reading and mathematics tests in WCPSS continued to demonstrate an upward trend across grade levels and student subgroups. Disaggregation of results by ethnicity, income level, disability status, and English proficiency status showed that achievement gaps between historically underperforming subgroups and their peers are still significant but are closing. Results of the new EOG science test in grades 5 and 8 showed that proficiency rates in science are below those for reading and mathematics. The effect of retesting for students who fail to score proficient on the first administration of EOG tests indicated that retesting improved proficiency rates by 4-9 percentage points, and that some groups of students benefited more than others from retesting.

BACKGROUND

End-of-Grade (EOG) tests are designed to measure student mastery of the knowledge and skills of the North Carolina Standard Course of Study (NCSCOS) for students in grades 3 through 8 in both reading and mathematics, and in grades 5 and 8 in science. The tests are administered during the final weeks of the school year, and are a vital part of the North Carolina accountability program referred to as the “ABCs”. EOG test results are also used to determine whether schools make Adequate Yearly Progress (AYP), a standard required by the federal Elementary and Secondary Education Act of 1965, as amended by the No Child Left Behind Act of 2001 (NCLB).

Changes in 2008-09. For several years, state and local student accountability standards required students who failed to score at or above Level III on their EOG to be retested with an alternate form of the test. This retesting gave individual students a second opportunity to demonstrate that their learning has prepared them for success at the next grade level. Students who did not pass on the second opportunity could be considered for retention.

While these retest scores had previously been important for decision making about individual students, they had not been used in the past for the public reporting of testing results. That is to say, a school's publicly-reported testing and accountability results (under the ABCs and AYP) were based solely on the first administration of the test for each student. Whether a student did better on a retest may have had implications for the individual student, but had no effect on the school's official results.

For the first time in 2008-09, the North Carolina Department of Public Instruction (NCDPI) chose to include retest results into each school's official accountability measures. Specifically, if a student did not pass the test on the first administration and was subsequently retested, the higher of the two scores became the official score of record for the school's publicly-reported proficiency rates under the ABCs and AYP. For that reason, statewide testing results in 2008-09 appeared to improve due in part to this change in policy.

In this report, WCPSS EOG results for the 2008-09 school year will be reported both with and without the inclusion of retest results. Although NCDPI did not report results based only on initial tests in 2008-09, this additional information will allow for more direct comparison of 2008-09 results in WCPSS with results from previous years.

MEANING OF EOG SCORES

EOG scores are generally reported in two ways: as a scale score, and an achievement level. EOG scale scores have a range of around 80-85 points, depending on the subject and grade level, with that range of scores then divided into four achievement levels. The achievement level score categorizes performance on EOG tests according to four broad categories, defined by the North Carolina Department of Public Instruction (Table 1).

Table 1
Achievement Levels for the North Carolina Testing Program

Level I: Students performing at this level do not have sufficient mastery of knowledge and skills in this subject area to be successful at the next grade level.	Level III: Students performing at this level consistently demonstrate mastery of grade level subject matter and skills and are well prepared for the next grade level.
Level II: Students performing at this level demonstrate inconsistent mastery of knowledge and skills in this subject area, and are minimally prepared to be successful at the next grade level.	Level IV: Students performing at this level consistently perform in a superior manner clearly beyond that required to be proficient at grade-level work.

Note: Official descriptions vary by grade level and are listed in NC State Board of Education Policy HSP-C-018 (<http://sbepolicy.dpi.state.nc.us/policies/HSP-C-018.asp?pri=01&cat=C&pol=018&acr=HSP>).

Achievement results are typically reported based on the achievement levels, specifically as the percentage of students scoring at or above Level III, which is the state's official "cut point" for determining which students are proficient (i.e., are ready to move on to the next grade level). In addition to reporting results in this manner, this report will also analyze results based on average scale scores, which provide a more sensitive measure of achievement that reflects student

performance across the entire continuum rather than simply focusing on how many students fall on either side of the Level III cut point.

DISTRICTWIDE RESULTS

Proficiency Rates

Two major changes have taken place in the state's EOG testing program in recent years. First, in 2007-08, the State Board of Education introduced new and more difficult reading tests in Grades 3-8. These changes were intended to better align North Carolina's EOG outcomes with the National Assessment of Educational Progress (NAEP) outcomes. The second major change was the addition of an EOG science test in grades 5 and 8 in 2008-09.

Reading. On the previous edition of the reading test given prior to 2007-08, WCPSS proficiency rates were 91.1% (Holdzkom, 2008). In the first year of implementation of the new test, the WCPSS proficiency rate fell substantially (66.2%), but still remained ahead of the statewide rate of 56.8% (Stevens, 2009)¹. In 2008-09, reading proficiency rates in WCPSS bounced back slightly, increasing to 68.2% (Figures 1 & 2). When factoring in retest results, this rate increased further to 74.7%, as compared to a statewide rate of 68.7%.

Mathematics. In mathematics, the WCPSS proficiency rate increased overall from 78% in 2007-08 to 79.5% in 2008-09 on the initial test. When factoring in retests, the 2008-09 proficiency rate was 84.5%² (Figures 1 & 2). This compares favorably with the statewide rate of 80.9% in 2008-09, including retests.

Science. The 2008-09 school year was the first year of testing in Grades 5 and 8³. The proficiency rate based on initial tests was 63.8%, increasing to 70.8% after retests were considered (Figures 1 & 2). The statewide proficiency rate for science in 2008-09, including retesting, was 64.9%.

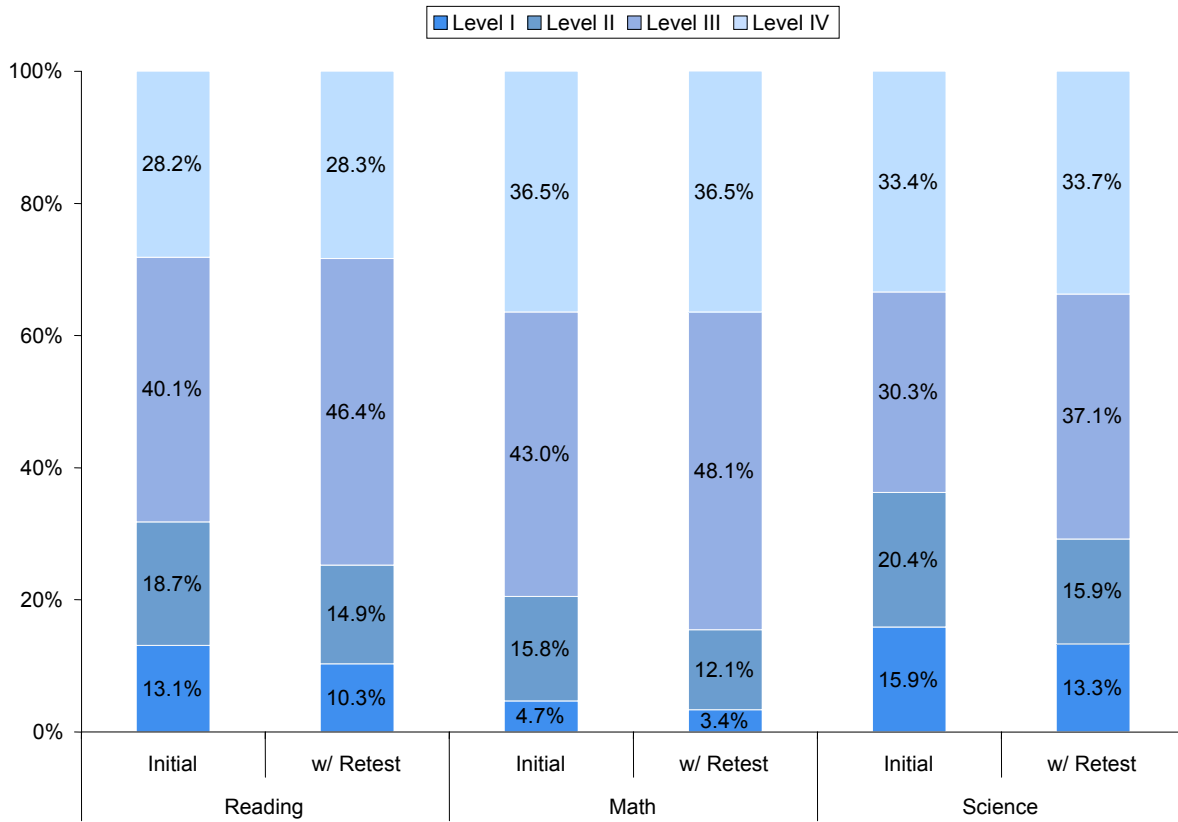
Effects of Retesting. The effects of counting retesting in 2008-09 testing results yielded approximately a 5-7 percentage point increase in proficiency rates in WCPSS, depending on the subject and grade level. As seen in Figure 1, the majority of this increase is due to students scoring in Level II on the first administration then scoring in Level III after retesting. It is also important to note that this increase due to retesting was in addition to general increases at all grade levels and subjects even when only the first administration of the test is considered (Figure 2).

¹ Additional school-by-school as well as statewide EOG performance data for 2008-09 can be found at <http://disag.ncpublicschools.org/2009/>.

² Mathematics and Science proficiency rates reported here do not include results for limited English proficient students who are in their first year in a U. S. school, as per NCDPI reporting guidelines.

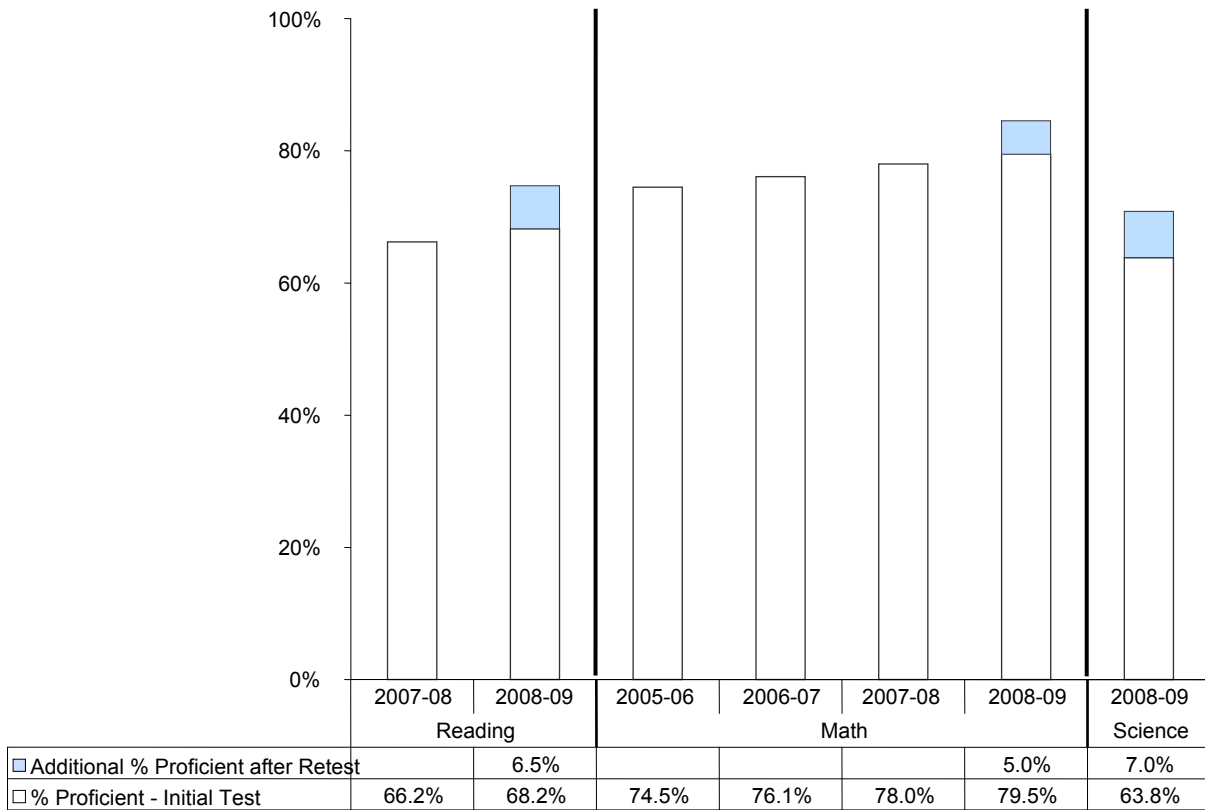
³ Although science EOG results for 2007-08 are available on NCDPI's Web site at <http://disag.ncpublicschools.org/2009/>, the 2007-08 administration of that test was not used as part of any official testing and accountability calculations that year.

Figure 1
Percentage of EOG Scores by Achievement Level, 2008-09



Note: Figures in chart may not match text exactly due to rounding.

Figure 2
EOG Proficiency Rates by Subject Area, Grades 3-8 Combined



Note: Figures in chart may not match text exactly due to rounding.

Figures 3 and 4 show the two-year change in proficiency rates for reading for elementary and middle grades, respectively. Proficiency rates based on initial testing showed increases in all grade levels, ranging from 0.2% in grade 8 to 3.2% in grade 7. The addition of retest scores resulted in a further increase of between 5.7 and 7.5 percentage points in each grade level.

Figure 3
Reading EOG Proficiency Rates, Grades 3-5

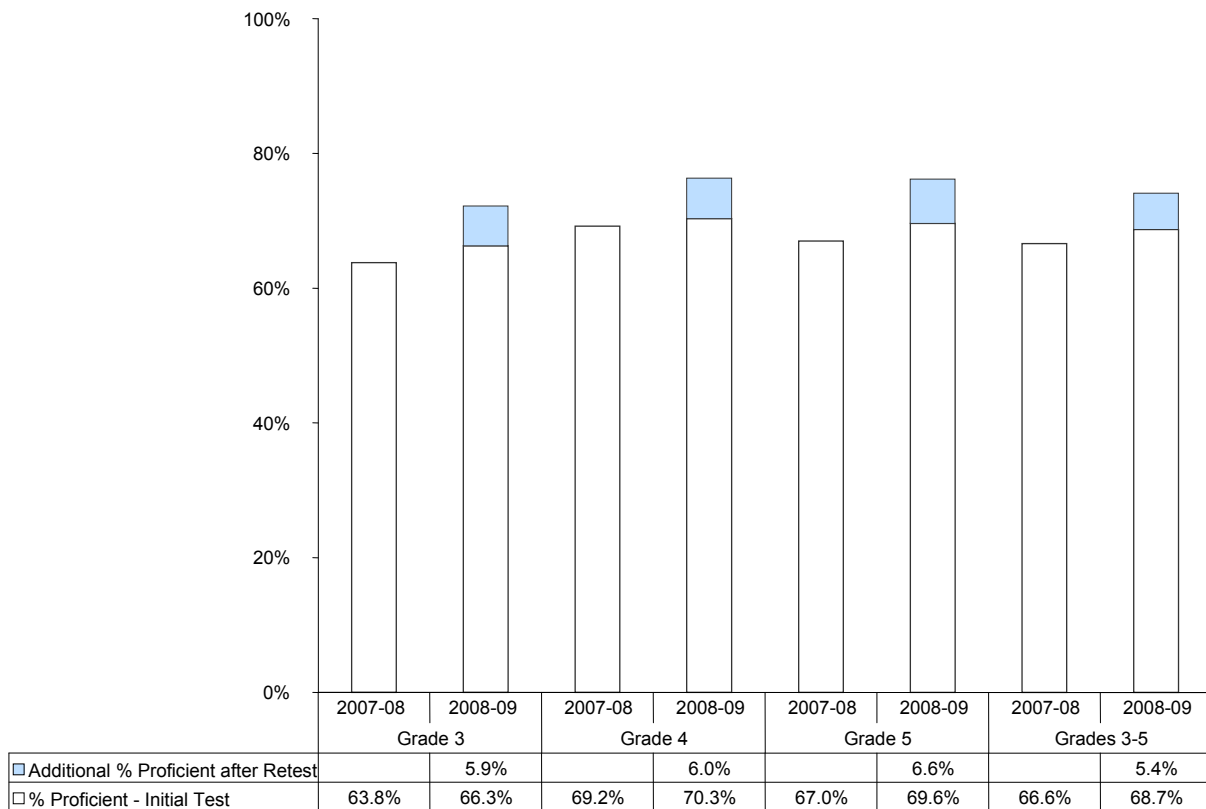
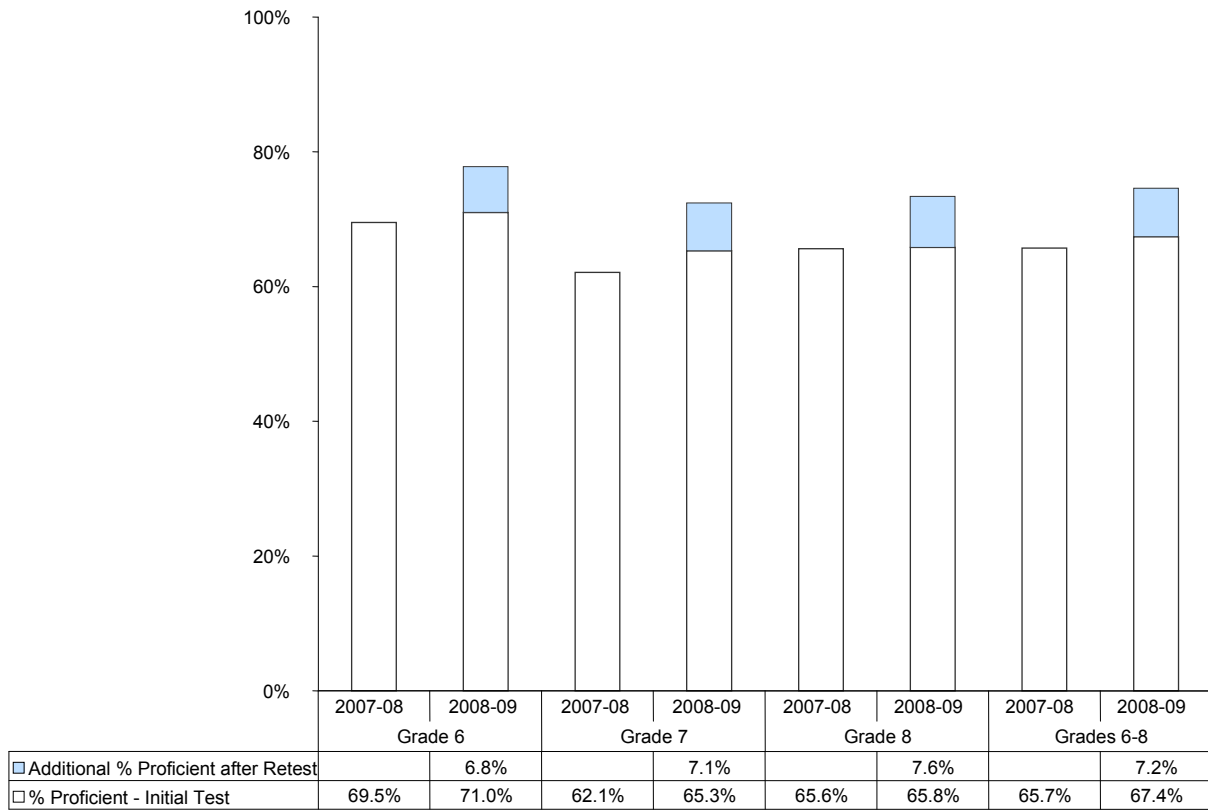


Figure 4
Reading EOG Proficiency Rates, Grades 6-8



Figures 5 and 6 show the percentage of all students who scored at or above grade level in mathematics between 2005-06 and 2008-09. Systemwide, more than three-quarters of elementary and middle school students scored proficient based on initial testing in 2008-09. Since the current edition of the mathematics EOG test was introduced in 2005-06, every grade 3 through 8 has shown improvement every year. Beyond that, the inclusion of retest results for the first time in 2008-09 added approximately 4-6 additional percentage points to the proficiency rates for each grade level.

Figure 5
Mathematics EOG Proficiency Rates, Grades 3-5

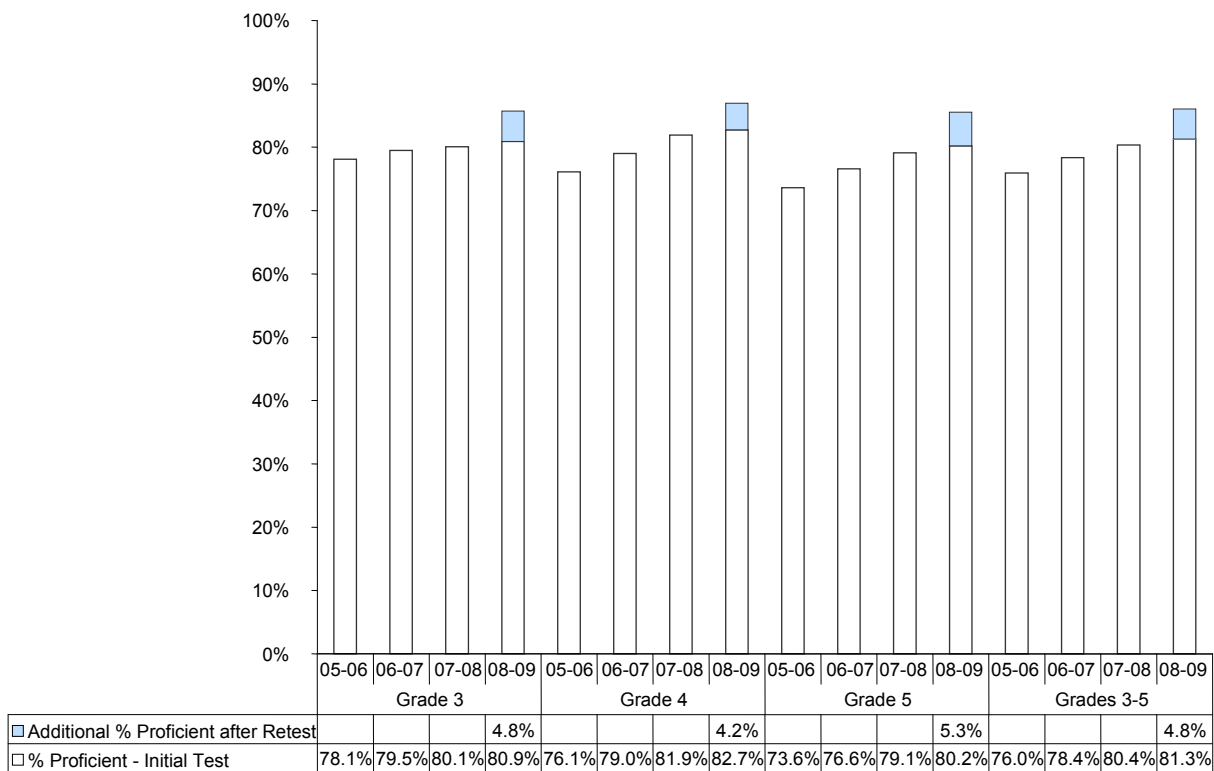


Figure 6
Mathematics EOG Proficiency Rates, Grades 6-8

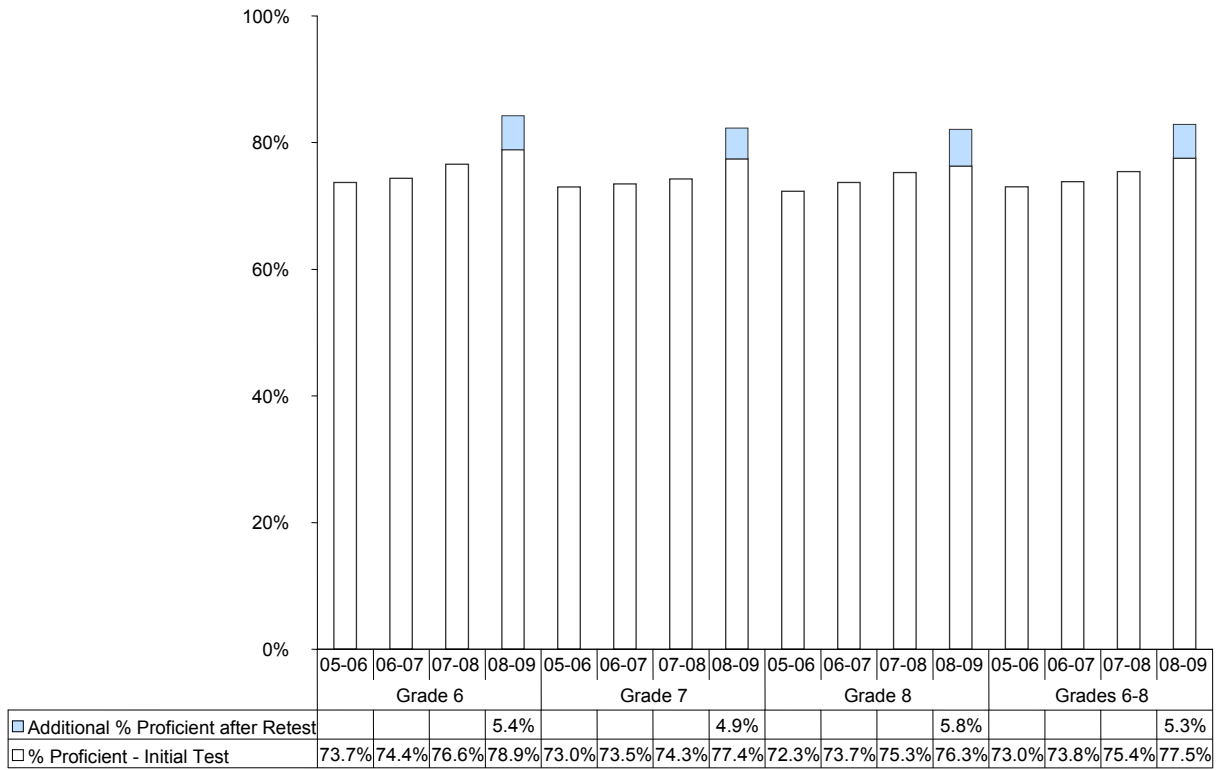
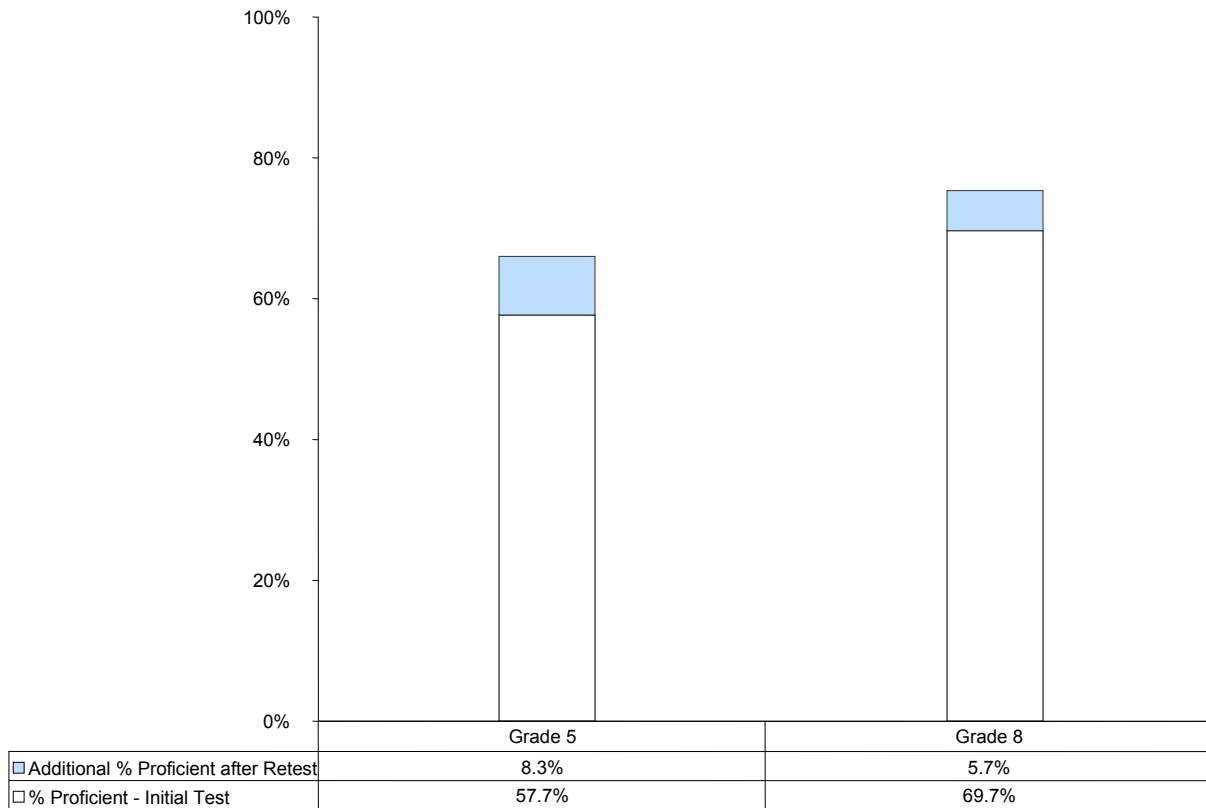


Figure 7 shows results for science. Proficiency rates based on initial testing were 12 percentage points higher in grade 8 than in Grade 5. Retesting helped to close this gap somewhat, but not entirely. Trend data are not shown for science EOG results, as 2008-09 was the first official year of testing.

Figure 7
Science EOG Proficiency Rates, Grades 5 and 8



Average Scale Scores

In addition to examining how many students score above the proficiency cut point, it is also important to examine the change in the average EOG scale score from year to year. Looking at changes in average scores provides a more sensitive measure of change, as the average takes into consideration movements along the entire performance continuum, rather than simply counting scores on either side of the proficiency cut point. It is possible for proficiency rates to increase (or decrease) regardless of whether the average score is changing in a similar direction. Therefore, looking at average scale scores can provide additional confirmation that changes in proficiency rates are in fact related to a general upward movement of scores overall.

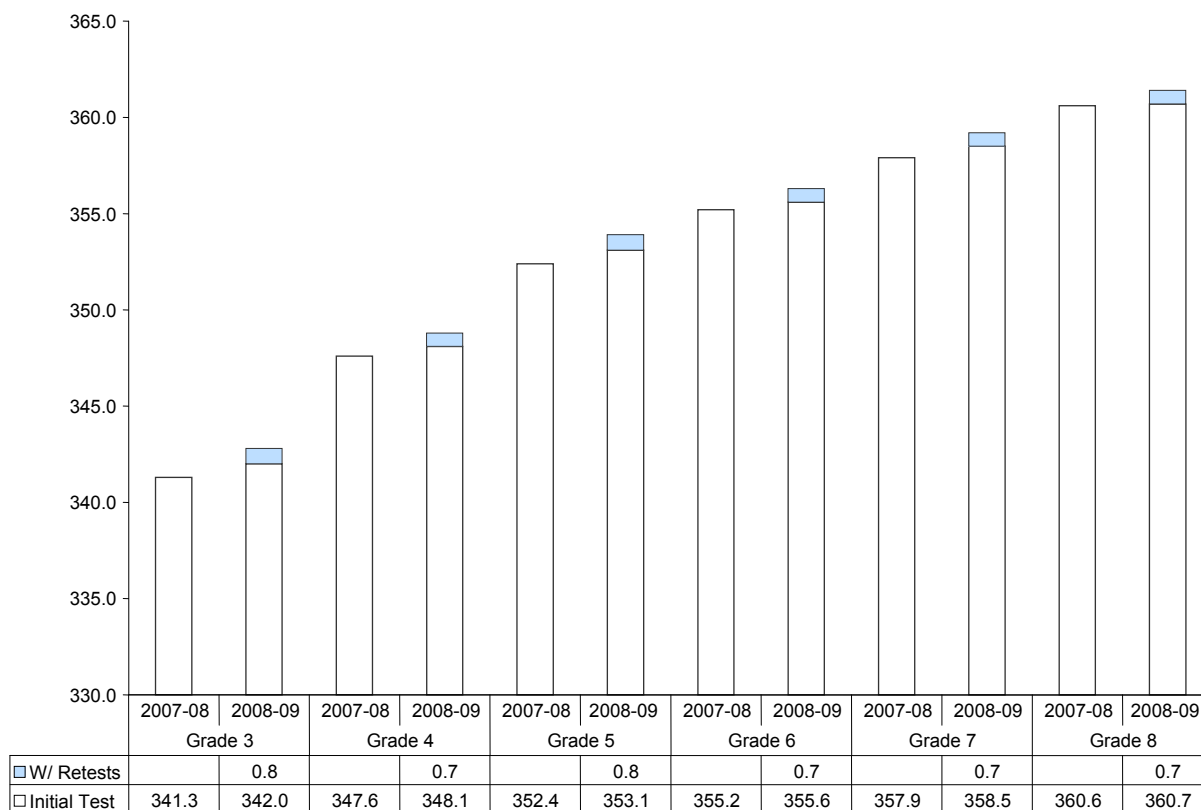
Since the scale of possible scores on EOG tests is different for every grade level, the following charts are grade and subject-specific. It is important to note that it is not legitimate to directly compare the average scale score in one grade level with that in another, or to compare average

scale scores between two different subjects, as the underlying scales on which the tests are measured are unique to each grade level and subject.

The average scale score on the reading EOG test for the past two years is displayed in Figure 8, while the results for mathematics are displayed in Figure 9. In reading, average scale scores increased between .1 and .7 scale score points at each grade level in 2008-09, based on the first administration of the test. The inclusion of retest results adds an additional .7 and .8 scale score points to the average for each grade level. In mathematics, the increase in average scale scores based on the first administration was between .2 and .7 points in 2008-09, depending on the grade level. The inclusion of retest results added an additional .4 to .6 scale score points to those increases. As with the proficiency results, mathematics average scale scores have increased every year at every grade level since 2005-06 with the exception of grade 6 in 2006-07, when it remained level from the previous year.

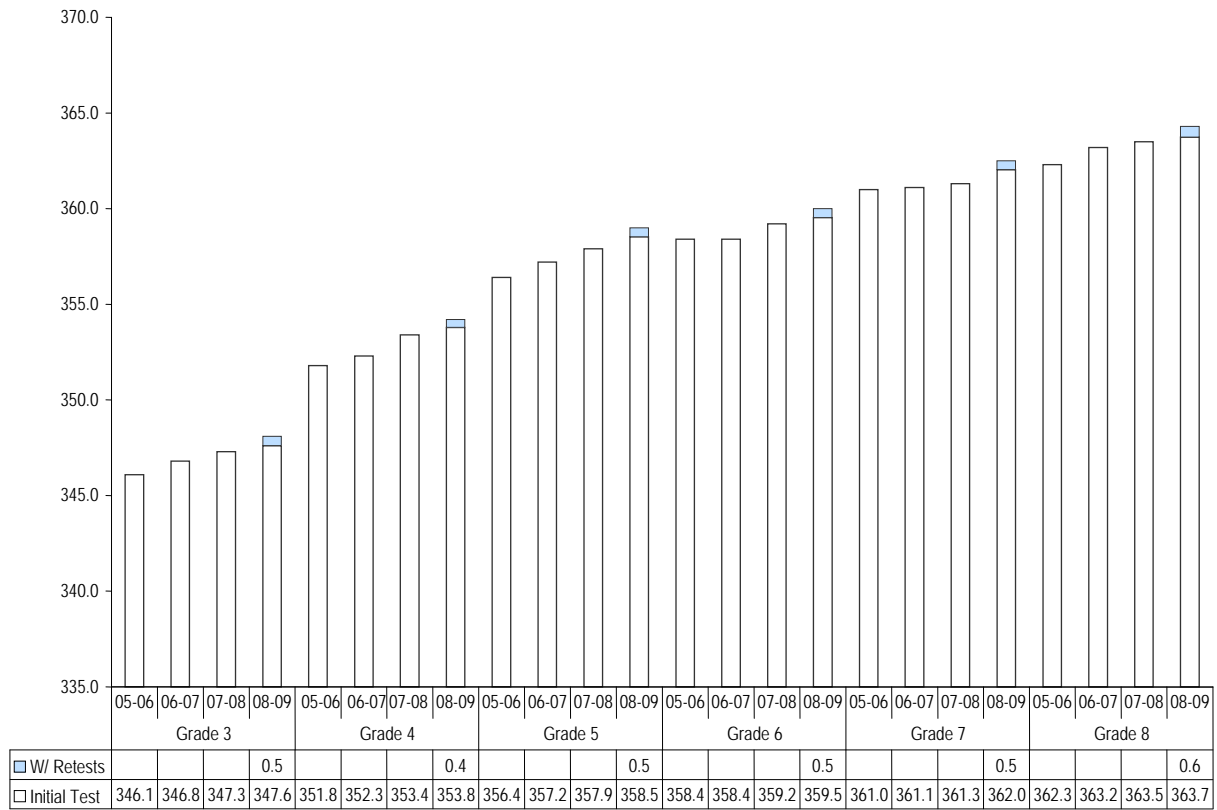
In science (Figure 10), scale scores are shown for grade 5 and 8 for 2008-09 only, as this was the first year that the science EOG test was officially part of the statewide testing and accountability reporting system. The addition of retest results added .9 points to the average in grade 5, and .5 points in grade 8.

Figure 8
Average Reading EOG Scale Score by Grade Level



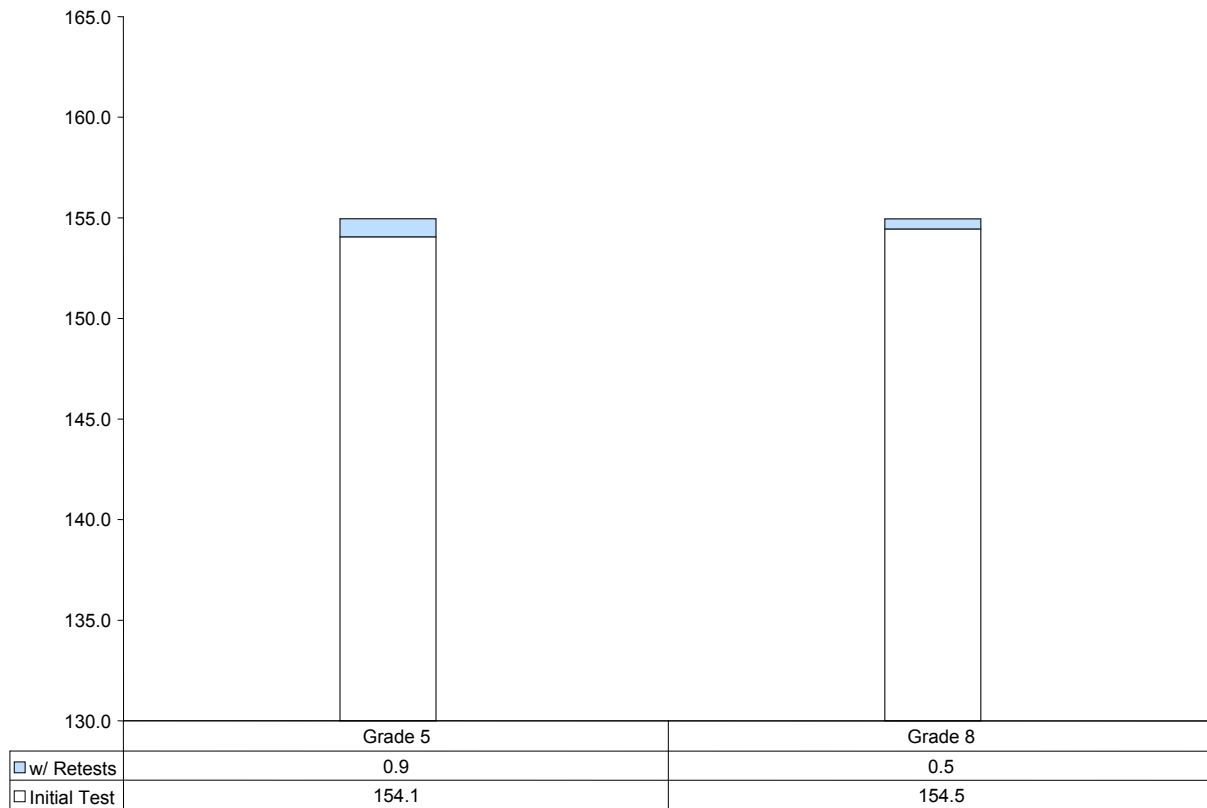
Interpretation example: The average scale score in Grade 3 increased from 341.3 in 2007-08 to 342.0 in 2008-09 (0.7 points). When retests were included, the average scale score in Grade 3 in 2008-09 increased further to 342.8.

Figure 9
Average Mathematics EOG Scale Score by Grade Level



Interpretation example: The average scale score in Grade 5 increased from 357.9 in 2007-08 to 358.5 in 2008-09 (0.6 points). When retests were included, the average scale score in Grade 5 in 2008-09 increased further to 359.0.

Figure 10
Average Science EOG Scale Score by Grade Level



Interpretation example: The average scale score in Grade 5 was 154.1 in 2008-09. When retests were included, the average scale score in Grade 5 in 2008-09 increased to 155.0.

DISAGGREGATED PROFICIENCY RESULTS FOR 2008-09

This section of the report presents EOG proficiency results for different subgroups of students within the grade levels of 3-5 and 6-8. Groupings are based on gender, race and ethnicity, eligibility for free or reduced-price lunch (FRL), students with disabilities (SWD), or limited English proficient (LEP) status as set forth by NCLB.

Although results for American Indian students are reported in the charts, results from that subgroup are not discussed specifically in the text due to very small numbers of students in that group, which causes their results to be relatively unstable from one year to the next.

Results by Ethnicity

Figures 11 and 12 show disaggregated results for reading proficiency by ethnic group for elementary and middle grades, respectively. Of all ethnic groups in the elementary grades,

proficiency rates with retests ranged from a high of 89.3% (White) to a low of 49.8% (Hispanic/Latino). In grades 6-8, the highest proficiency rate was found in the Asian subgroup (89.6%) and the lowest was among Black/African American students (51.5%).

Looking at year-to-year changes, results based on initial testing only at the elementary level showed that students in all ethnic groups made proficiency gains in 2008-09 compared to 2007-08. Asian (3.3 percentage points), Multiracial (2.8), and White students (2.3) demonstrated the largest gains in 2008-09. When retests were factored in, however, Black/African-American and Hispanic/Latino students showed the greatest gains in 2008-09, with both groups increasing by more than 10 percentage points (Figure 11).

Figure 11
Reading EOG Proficiency Rates by Ethnic Group, Grades 3-5

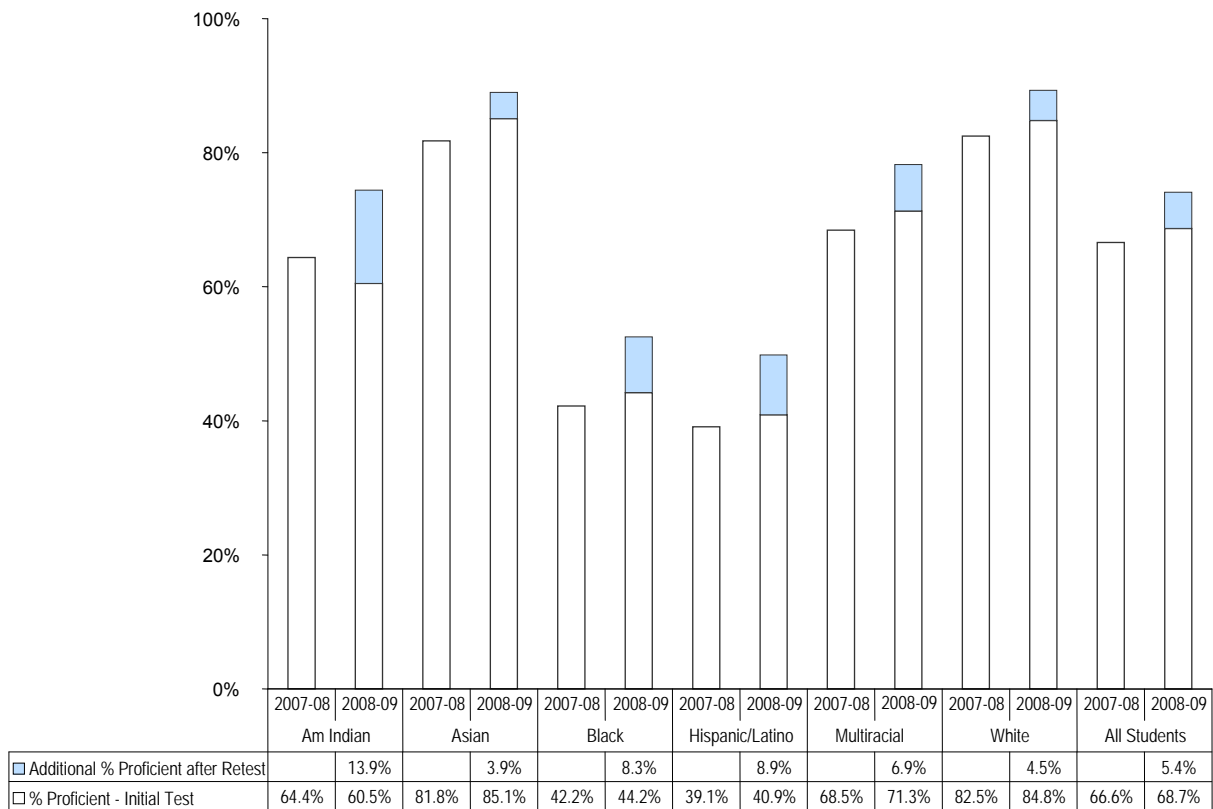
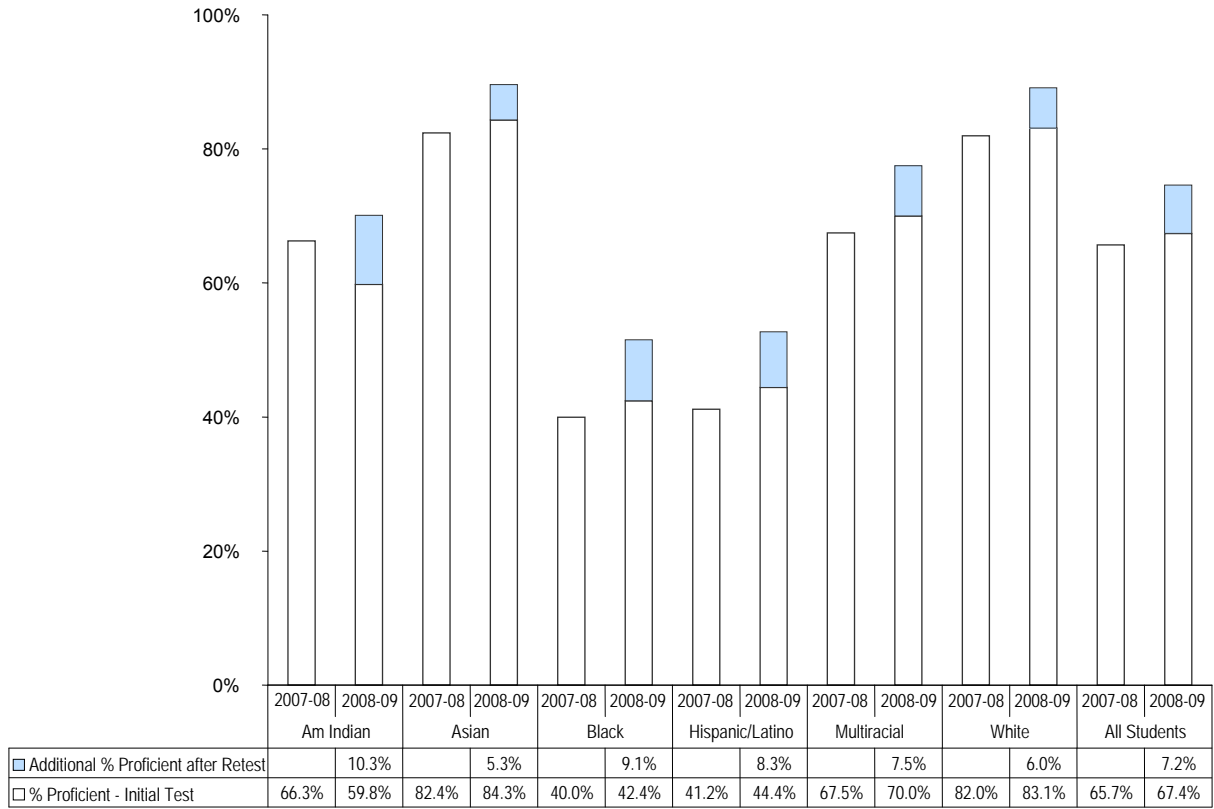


Figure 12
Reading EOG Proficiency Rates by Ethnic Group, Grades 6-8



Figures 13 and 14 show disaggregated results for mathematics proficiency by ethnicity. In grades 3-5, Asian students had the highest proficiency rate after retests at 97.0%, and the Black/African American students had the lowest with 69.5% proficient based on retests. In grades 6-8 after retest, Asian students again had the highest proficiency rate at 95.7%, and Black/African American students had the lowest with 63.7%.

Since 2005-06, most subgroups in both grade spans have been making steady progress in mathematics proficiency. Proficiency rates among Black/African American and Hispanic/Latino students, however, have been increasing at a higher rate than is true for other groups. For example, the elementary mathematics gap between the highest (Asian) and lowest-performing (Black/African-American) ethnic groups in 2005-06 was 40.3 percentage points. By 2008-09, that gap had decreased to 33.4 percentage points. The inclusion of retest results in 2008-09 further accelerated this trend, as Black/African American and Hispanic/Latino students were more likely to benefit from retesting than were other ethnic groups (Figures 13 and 14).

Figure 13
Mathematics EOG Proficiency Rates by Ethnic Group, Grades 3-5

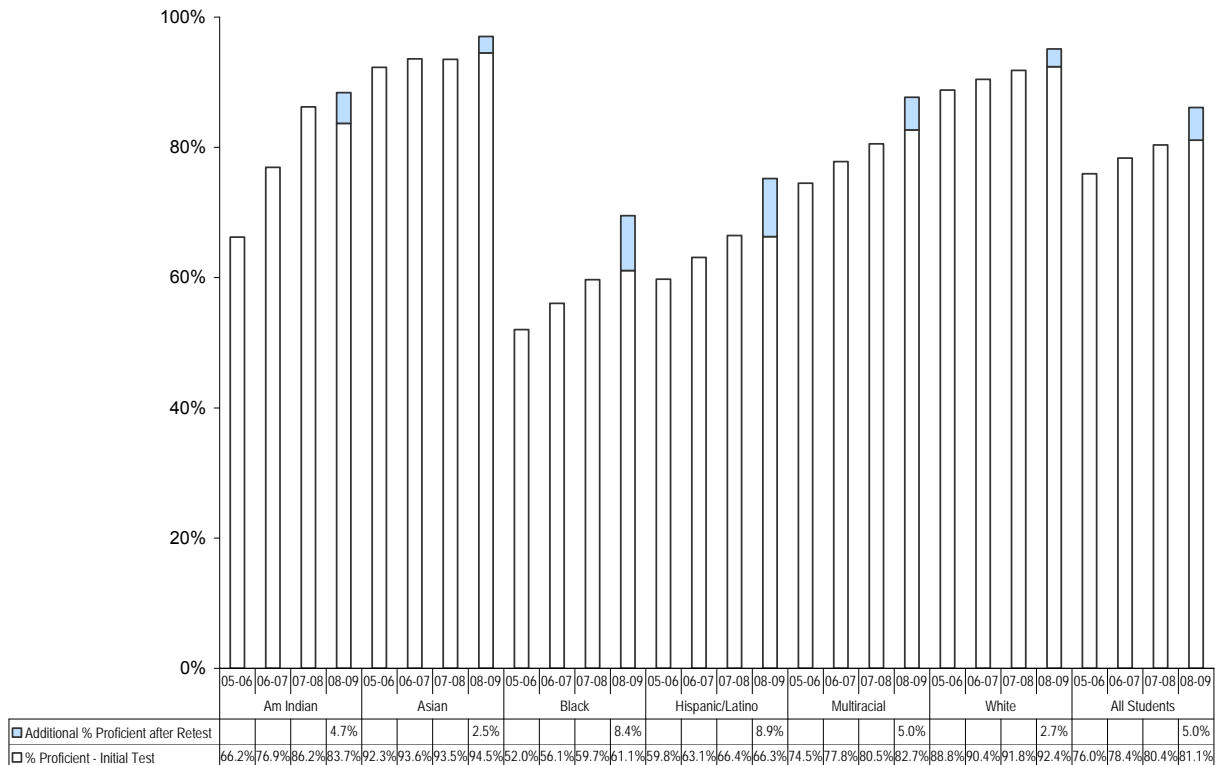
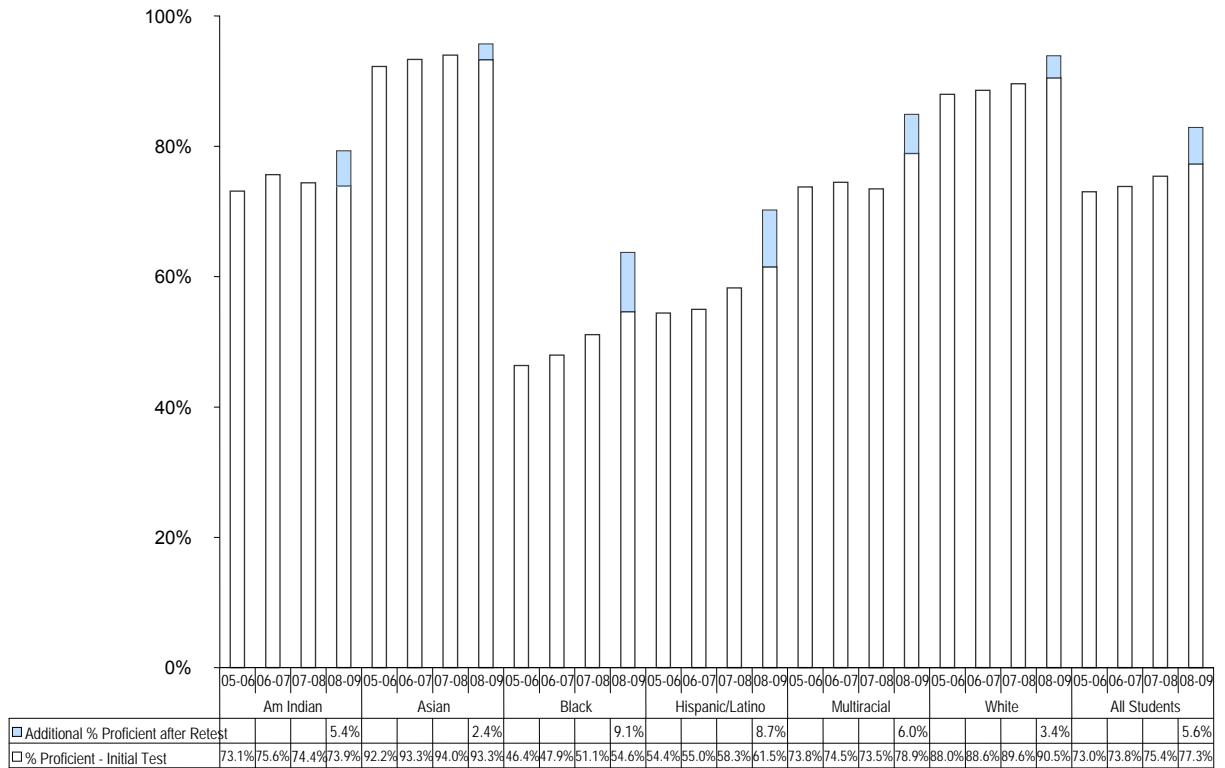


Figure 14
Mathematics EOG Proficiency Rates by Ethnic Group, Grades 6-8



Science proficiency rates for each grade span are displayed in Figures 15 and 16. While the same pattern of performance between ethnic groups found in reading and math also exists in science, the gaps between groups in science are somewhat larger. This is particularly evident in grade 5, where fewer than 30% of Black/African-American and Hispanic/Latino students scored proficient on the initial test. While grade 5 science had the lowest initial proficiency rate across all grades and subjects in 2008-09, that test also saw the largest benefits of retesting, with an additional 8.3% of students scoring proficient on their second attempt at the test.

Figure 15
Science EOG Proficiency Rates by Ethnic Group, Grade 5

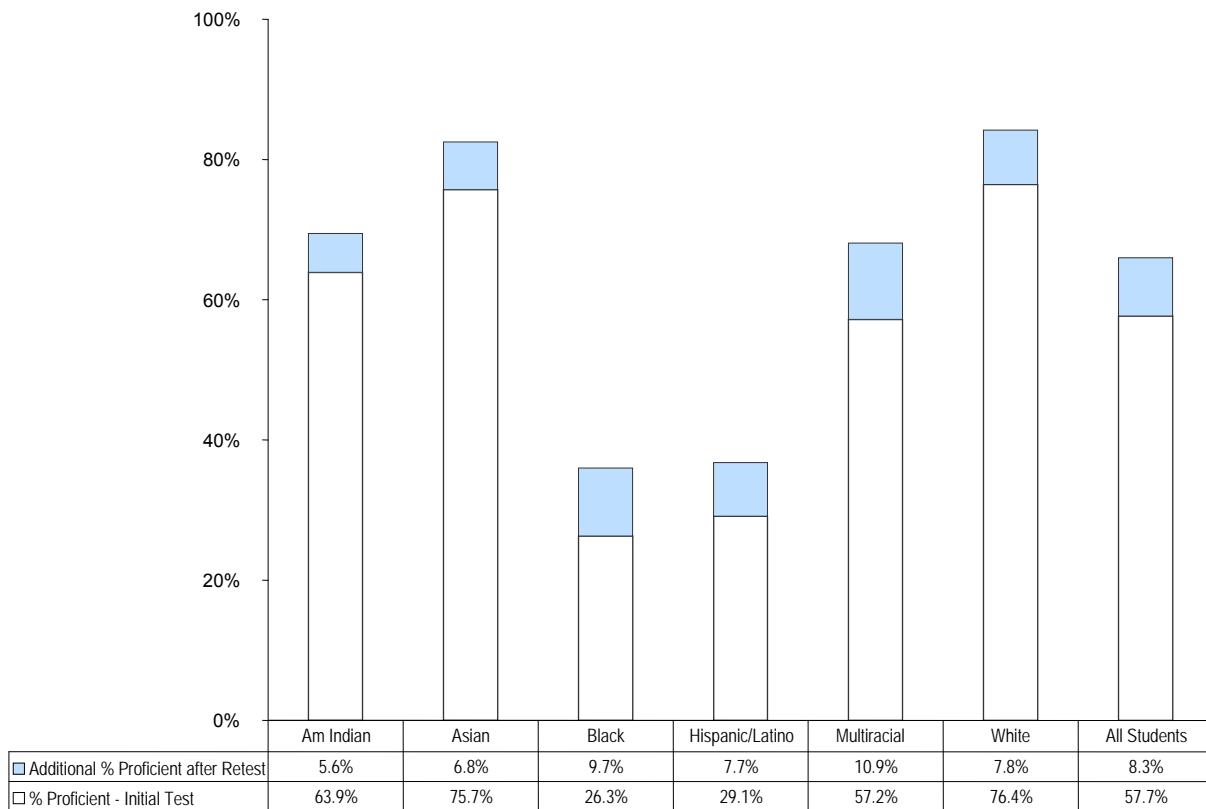
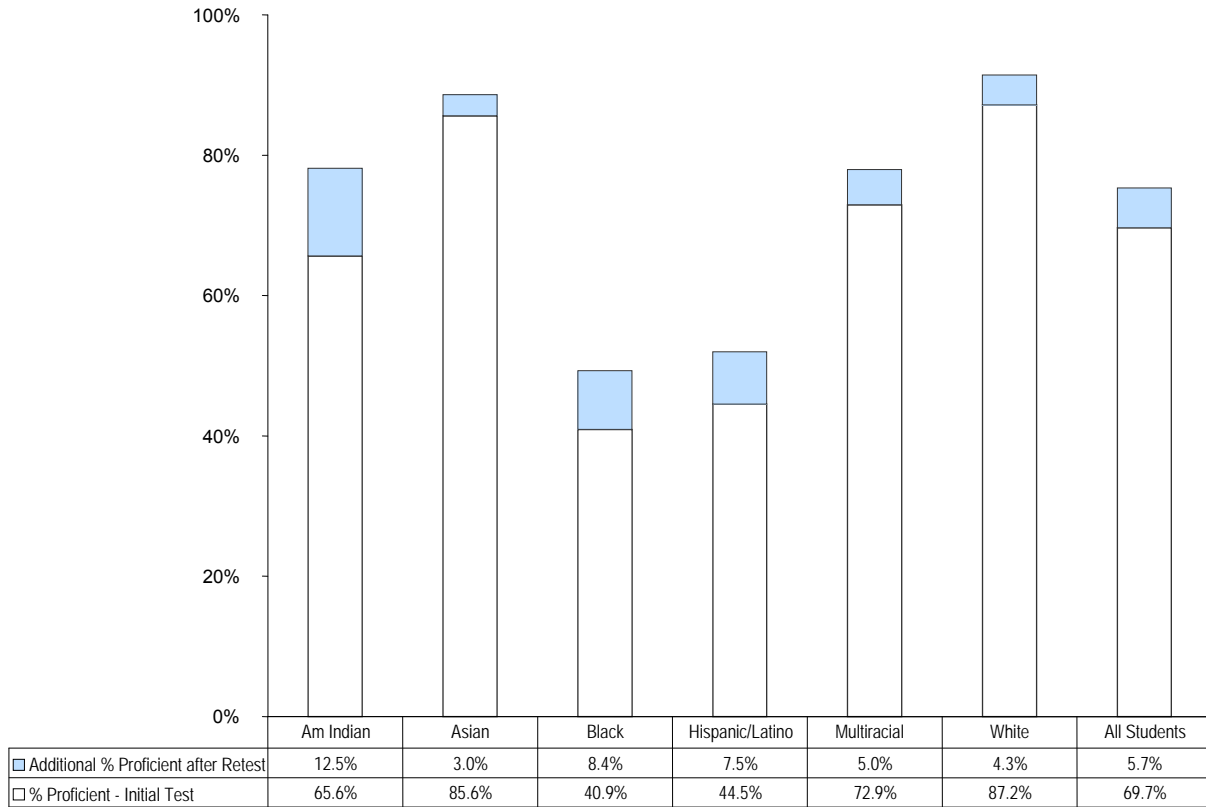


Figure 16
Science EOG Proficiency Rates by Ethnic Group, Grade 8



Results by Ethnicity and Gender

Figures 17 through 19 show disaggregated results for proficiency for male and female students within ethnic groups for each subject area. Overall, female students outscored their male counterparts in reading and math in 2008-09, although those margins were smaller for mathematics than for reading. This is consistent with results from prior years (Holdzkom, 2008; Stevens, 2009). In science, the reverse was true, with male students outscoring female students (Figure 19).

Figure 17
Reading EOG Proficiency Rates by Ethnicity and Gender

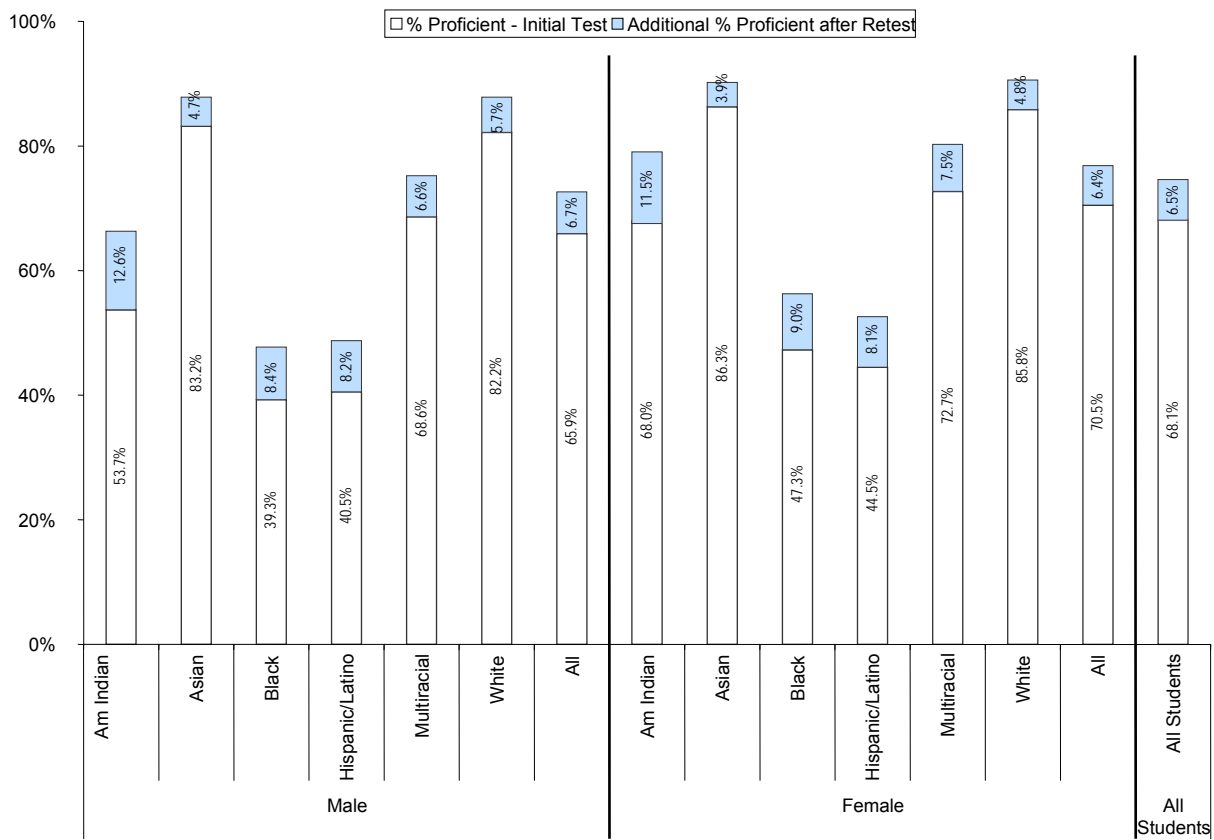


Figure 18
Mathematics EOG Proficiency Rates by Ethnicity and Gender

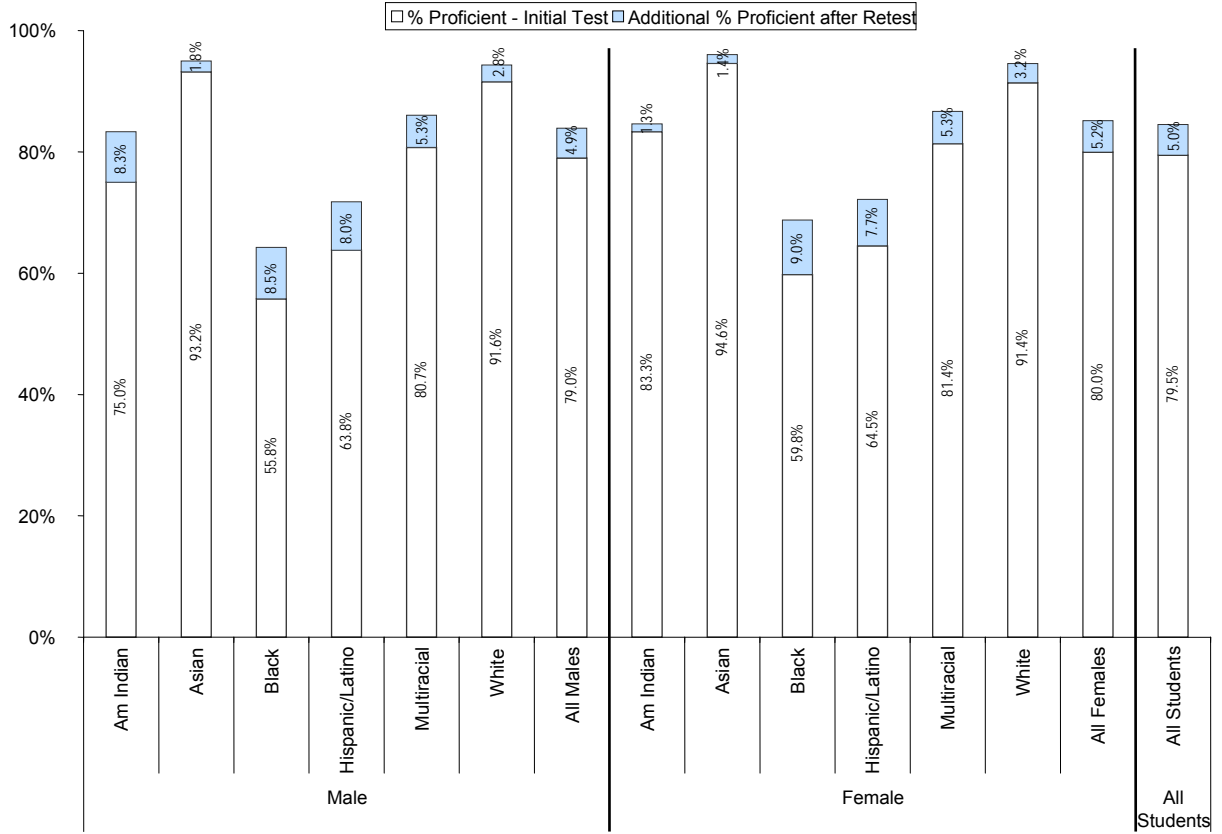
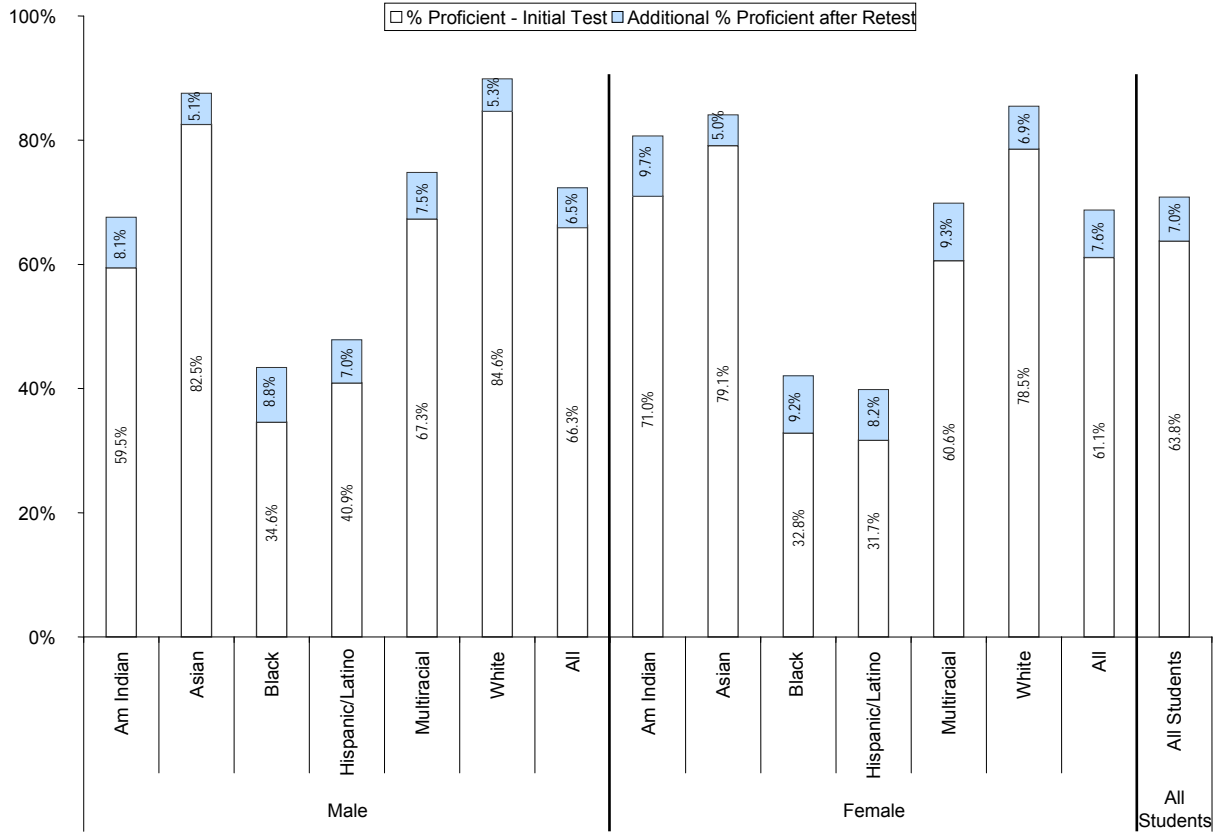


Figure 19
Science EOG Proficiency Rates by Ethnicity and Gender

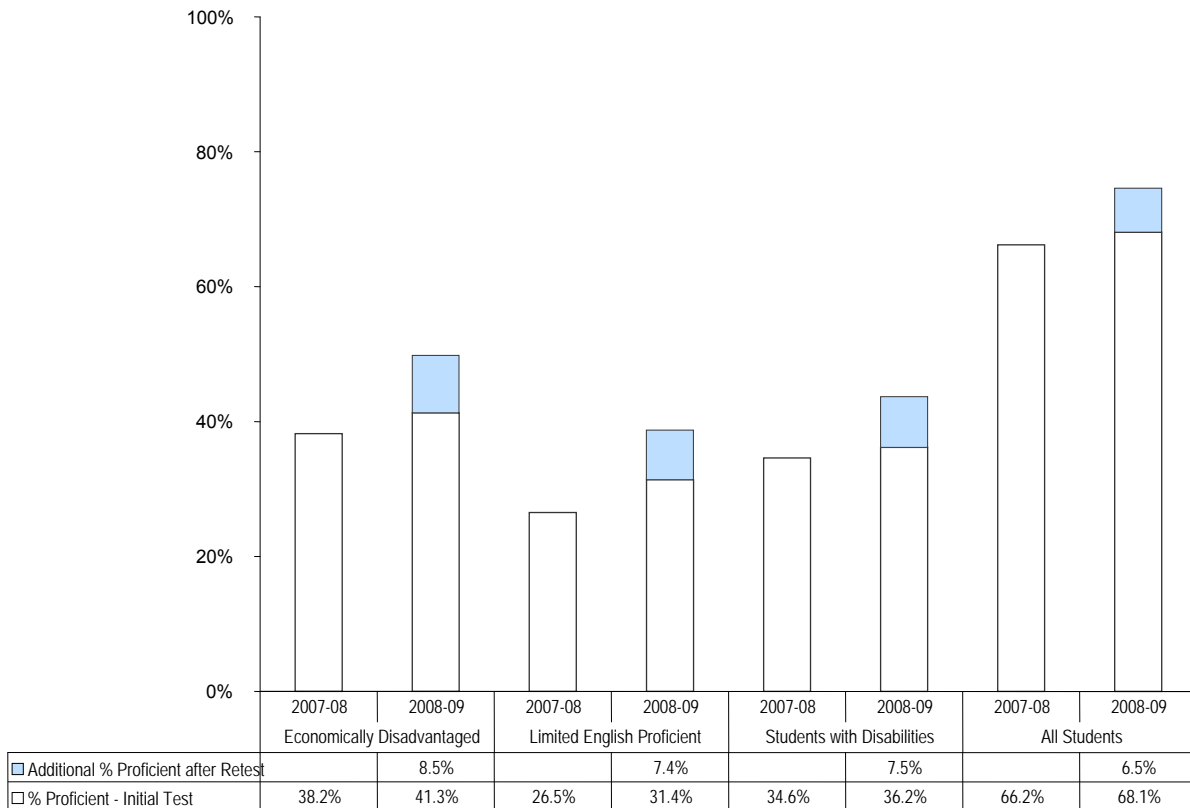


Results by NCLB Risk Factor Categories

Figures 20 through 22 show the EOG proficiency rates for WCPSS students based on risk factor subgroups as defined by NCLB. These risk factor subgroups are: students with disabilities (SWD), economically disadvantaged students (ED⁴), and students with limited English proficiency (LEP). Students who are classified in any of these subgroups have shown an increased difficulty in reaching proficiency on EOG tests (Baenen, Paepflow, Ives, & Reichstetter, 2007).

Less than 40% of the students in these various subgroups in elementary and middle school were at or above grade level on the 2007-08 Reading EOG. In 2008-09, however, all groups made progress, with LEP students showing the largest gain based on initial testing (4.9 percentage points). Once retests were factored in, scores for these three groups in reading increased by an additional 7.4 – 8.5 percentage points (Figure 20). The increase in scores due to retesting was slightly larger than the 6.5 percentage point boost that retesting had for students overall.

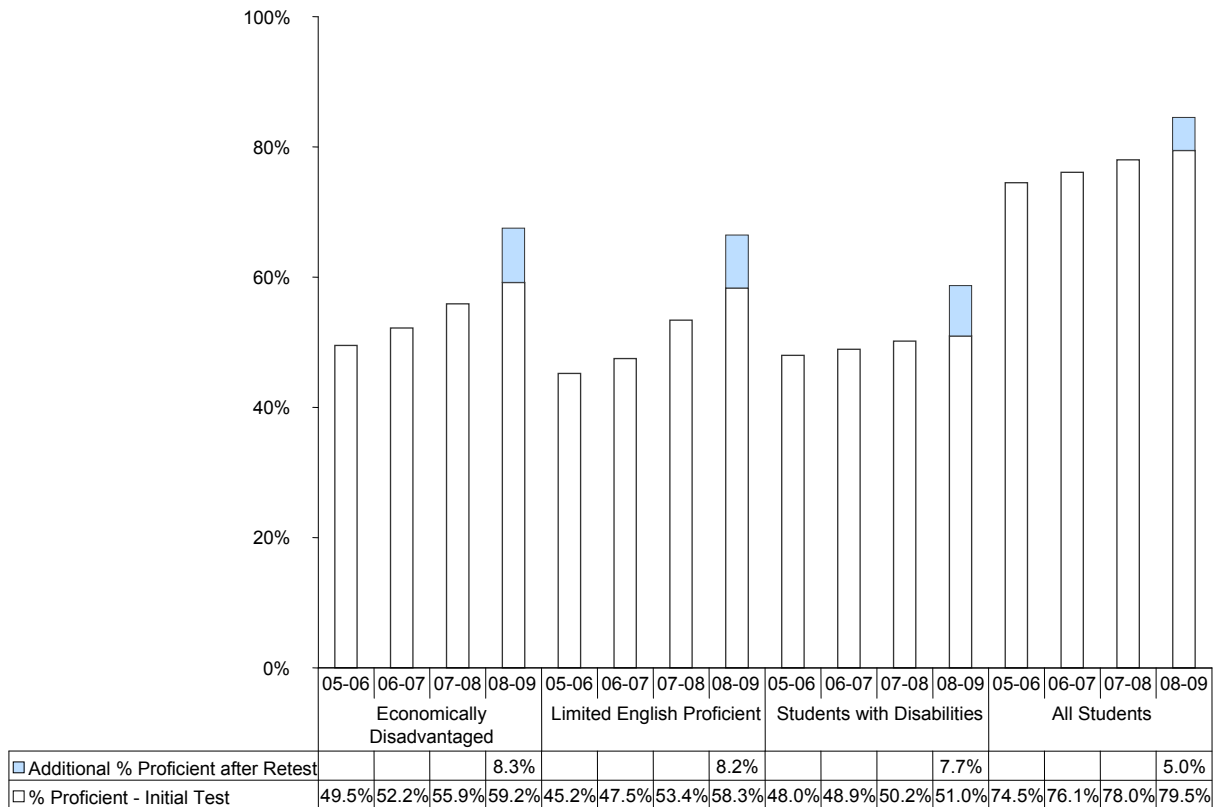
Figure 20
Reading EOG Proficiency Rates by NCLB Subgroup



⁴ “Economically Disadvantaged” is the state’s term for students who are eligible for free or reduced-price lunches.

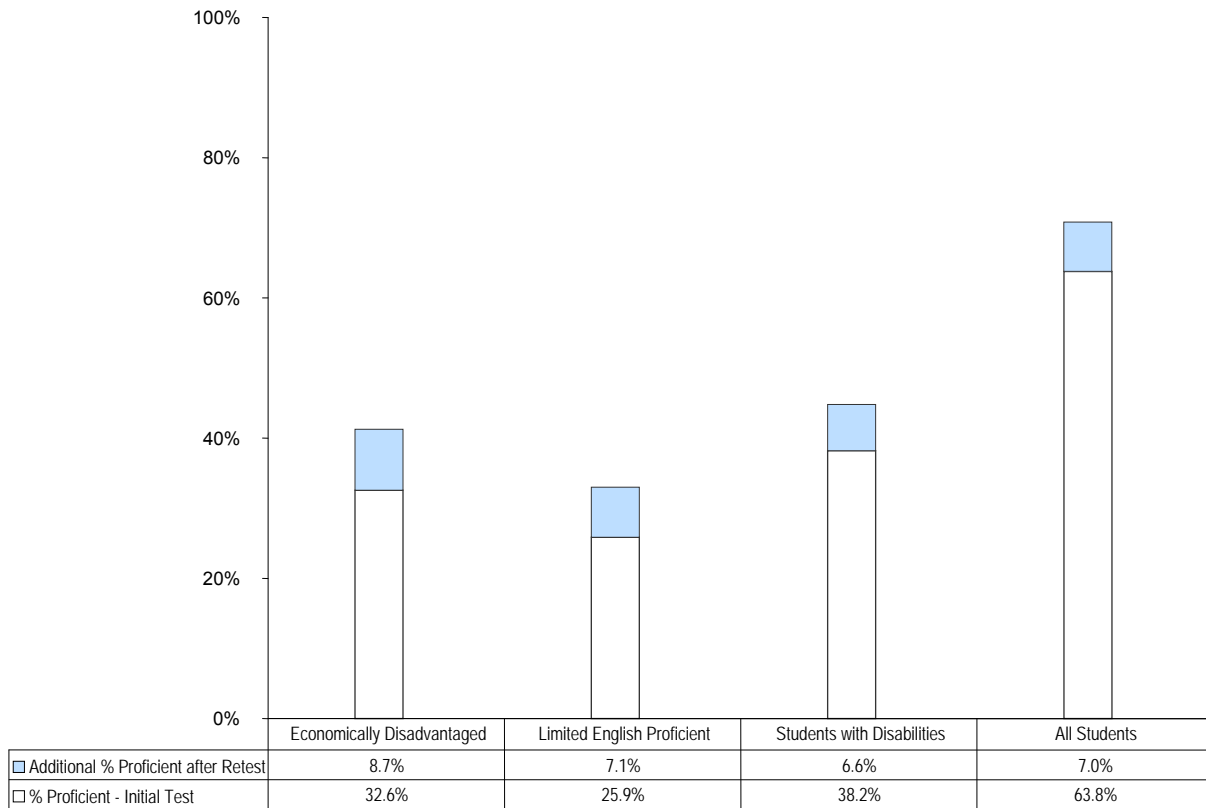
Proficiency rates based on initial testing also increased in mathematics for each of these three student groups in 2008-09, marking the fourth consecutive year of improvement for each. As in reading, LEP students posted the largest gain in 2008-09 (4.9 percentage points), followed by ED students (3.3 points; Figure 21). Retesting results added around 8 additional percentage points to each group’s proficiency rate in 2008-09, larger than the 5 percentage point gain seen among students overall.

Figure 21
Mathematics EOG Proficiency Rates by NCLB Subgroup



The first year of official reporting of EOG science results in 2008-09 showed that less than 40% of students in the three NCLB subgroups in grades 5 and 8 scored proficient based on initial testing, as compared to 63.8% systemwide (Figure 22). The addition of retest results added between 6.6 and 8.7 percentage points to those results for those three groups, with LEP students scoring at only 33% proficient even after including the results of retesting. The 37.5 percentage point gap in science between LEP students and the district as a whole is larger than the gap for any other NCLB subgroup in any subject.

Figure 22
Science EOG Proficiency Rates by NCLB Subgroup



DISCUSSION

New tests in reading for 2007-08 dramatically impacted WCPSS proficiency rates, but those results began to bounce back in 2008-09, as improvement was seen at each grade level and for every major student subgroup included in this report. Results for mathematics also improved across the board in 2008-09, continuing a steady trend evident since the current edition of the mathematics EOG test was introduced in 2005-06. Since 2008-09 was the first official year of EOG science testing for accountability purposes, trends are yet to be enumerated in that subject, although the overall proficiency rates generally trail those in reading and mathematics. As science tests become part of the accountability milieu over the next few years, it will be important to see whether the steady upward progress seen in reading and mathematics is reflected in that subject as well.

Despite these overall positive trends, historic achievement gaps persist between students in some subgroups and their peers across all subjects. Black/African-American students, Hispanic/Latino students, and students in the three subgroups defined by NCLB – students with disabilities, economically disadvantaged students, and students with limited English proficiency – still trail their peers in terms of EOG proficiency by large margins. It is important to note, however, that the general increase in proficiency rates in WCPSS appears to be disproportionately driven by improvements in these very subgroups. Since those historically-underperforming subgroups are gaining in proficiency at a faster rate than other students, then it is clear that those achievement gaps are closing in WCPSS.

What is also clear, however, is that those gaps are not closing rapidly enough. For example, if current trends in elementary math proficiency continue in a linear fashion, the Black/African-American proficiency rate will not reach the *current* proficiency rate for White students for another 14 years. So while virtually all of the trend results in this report are moving in a positive and more equitable direction, the pace clearly needs to accelerate.

The addition of retest results does accelerate the closing of achievement gaps in a sense, as evidenced by the fact that lower-scoring subgroups appear to benefit the most from the opportunity to take the test a second time. However, using the higher of two scores for students who fail to score proficient on their first attempt introduces some interpretation problems into the results. Testing a student twice produces in essence two separate estimates of that student's "true" score, and the state's testing rules dictate that the highest of the two scores stand as the official result for each student. This implies that the higher score is the more accurate one. Since from a statistical perspective neither score can be assumed to be more accurate than the other, choosing the highest score by default will artificially inflate the results for some students who are given that second opportunity. If the goal of retesting is to prevent students who have truly mastered the material from being classified as non-proficient, then this approach is exactly the correct one. However, that best-out-of-two scenario then compromises the ability to use those scores in the aggregate to draw conclusions about the relative performance of subgroups, because some students are allowed to increase their score via retesting while others are not. The benefits of retesting in this situation will accrue disproportionately to those groups that score lower on the initial test, and consequently the gap between those lower-performing groups and

their higher-performing peers will shrink. This phenomenon is precisely the reason why this report clearly separates the examination of initial testing results from retest-adjusted results, thereby allowing for a more consistent and comprehensive interpretation of group-level results.

In summary, both proficiency rates and average scale scores on EOG tests are continuing to rise across the board in WCPSS, and achievement gaps are closing between historically lower and higher-performing subgroups. The positive trends evident in these results are indicative of the hard work put in by students, teachers, and administrators throughout the district who have continued to focus relentlessly on learning and teaching for all students in the face of rising standards and high community expectations. Finding ways to accelerate these trends remains a challenge, however, as the number and percentage of students who score below proficient on EOG tests suggest that there is still significant room for improvement. Continuing to track data on student performance, including the effects of retesting and the introduction of the new science test over time, will allow us to more easily assess the effects of our efforts in raising achievement for all students and preparing them for future success in high school and beyond.

REFERENCES

- Baenen, N., Paeplow, C., Ives, S., Reichstetter, R. (2007). *WCPSS students with multiple academic risks: achievement patterns and school experiences*. Raleigh, NC: Wake County Public School System, Evaluation and Research Department.
- Holdzkom, D. (Ed.). (2008). *Wake County Public Schools (WCPSS) elementary school student outcomes: 2006-07*. Raleigh, NC: Wake County Public School System.
- Stevens, W. (2009). *End-of-Grade multiple choice test results, 2007-08*. Raleigh, NC: Wake County Public School System.