

STANDARDS-BASED GRADING 2005-06 and 2006-07

Colleen Paeplov

ABSTRACT

Wake County Public School System (WCPSS) uses a standards-based grading policy at the elementary school level that aligns with North Carolina’s Student Accountability Standards and the WCPSS Promotion/Intervention policy. This evaluation examined the distribution of grades given in 2005-06 and 2006-07 and the correlation of students’ grades and End-of-Grade (EOG) scores in 2006-07. The study found that over three fourths of students were graded as proficient in reading and mathematics. In reading and mathematics grade K-2 students earned higher grades than their 3-5 counterparts; overall, mathematics grades were lower than reading across the K-5 grade levels. Fourth-quarter reading and mathematics grades were correlated with reading and mathematics’ EOG scores. Mathematics grades had a stronger correlation to the newly revised mathematics EOG scores than did fourth-quarter reading grades to the current reading EOG scores. Furthermore, a student scoring proficient within the classroom has a greater likelihood of being proficient on their EOG than those with below grade-level scores.

BACKGROUND

For many years, researchers have been concerned with the imprecision of the U.S. grading system. As evidence of this imprecision, researchers point to the standardized test results for students earning A’s and B’s in high-poverty areas being comparable to students in more affluent schools earning C’s and D’s (U.S. Department of Education, 1994).

Some educators have enhanced report cards by providing information on a student’s proficiency on various standards as an initial step toward standardizing grades (Perlstein, 2003). Districts across the United States have been expanding and changing the traditional A-F letter grading system in favor of detailed standards-based reporting.

Through either a narrative evaluation format, or with numbers or symbols, these standards-based report cards provide detailed information on skills a student is expected to master (Manzo, 2001). Traditionally, teachers are given leeway in determining whether work merits an A or F. Teachers form tests that

<u>Key Topics</u>	
Background -----	p. 1-3
Research Questions-----	p. 4
Methodology -----	p. 4-6
Results -----	p. 6-17
Discussion/Recommendations----	p. 17-18

reflect the required objectives as they see them and may “tweak grades at the margins” by including homework or class participation (Perlstein, 2003). This grade adjustment may increase a student’s grade from a C to a B to reward the student’s efforts. The proportion of the grade based on quizzes, tests, homework, class participation, and other extra credit criteria can vary from teacher to teacher and across schools. “Grades, in other words, have signified a bit more than a sum of a child’s knowledge and academic abilities. Altering that will not be easy” (Perlstein, 2003, p. 1). According to measurement specialists, student achievement would be more accurately reflected and teacher subjectivity decreased by separating work habits from student performance (Cross & Frary, 1999).

Due to concerns that students’ grades were not consistently assigned across teachers, that student mastery of the N.C. Standard Course of Study was not reflected on report cards, and that parents were not properly informed of student progress, the Wake County Public School System (WCPSS) implemented a standards-based grading report card at the elementary school level in 2001-02. During the first year, WCPSS piloted a new performance-based report card utilizing standards-based grading at five elementary schools. Administrators, teachers, and parents at the five schools worked for two years to develop a new elementary school report card designed to promote consistency in the grading process and provide parents better information regarding their child’s “progress toward mastering the state’s Standard Course of Study” (K-5 Standards-Based Grading and Reporting: Fact Sheet, 2003, p. 1). In 2003-04, 22 additional schools implemented the new standards-based grading report card. In 2005-06, the balance of WCPSS’ 88 elementary schools adopted this new grading system. The number of elementary schools was expanded to 93 in 2006-07 due to the addition of five new schools.

The new standards-based report card was implemented to align with North Carolina’s Student Accountability Standards and the WCPSS Promotion/Intervention policy. This new tool was designed to provide parents and students with details about a student’s performance on grade level standards. The new standards-based report card measures a student’s academic progress against a uniform standard, rather than using subjective assessments. The report card includes a section for teacher comments designed to capture student strengths and areas of need (Understanding the Elementary School Report Card, 2005; K-5 Standards-Based Grading and Reporting: Fact Sheet, 2003). “The student performance levels of 1 to 4 indicate whether students have met the expectations set by the state in the Standard Course of Study and indicate whether the student has the necessary skills and concepts to be successful in the next quarter or next grade” (K-5 Standards-Based Grading and Reporting: Fact Sheet, 2003, p.1).

Rather than averaging the student’s cumulative work, standards-based grading assesses a student’s mastery of an objective and assigns a 1-4 rating for that objective. Work habits are not considered a part of the student’s performance on a given objective but are captured under a separate 1-3 rating. Since this new grading method represents a departure from the A-F grading scale traditionally used, the degree to which this policy has been understood and implemented is of interest.

The 1-4 performance levels do not correspond to the A-F scale; rather, they represent a criterion-referenced standard, in which student performance is measured against established criteria with

differentiated levels of quality. Assessment of student performance is measured not against the performance of other students but against the set performance standards (Guskey, 2001).

Within the 1-4 performance levels, level 3 signifies mastery of an objective; however, what constitutes a level 4 (extension of grade level standards) is not always easily understood. Teachers must make sure that they provide level 4 opportunities for each objective to ensure that students are able to receive a level 4. However, the level of experience a teacher has with standards-based grading may impact his or her ability to ensure they are including level 4 opportunities for each objective. Once student performance on each objective is assessed, these objectives are evaluated to produce the final grade by subject, to be recorded on a student's report card.

In 2003, while standards-based grading was being rolled out, schools' Instructional Resource Teachers (IRTs) were required to take a minimum of 12 hours of training on the standards-based report cards and provide training to the teachers at their school using a train-the-trainers model (Webb-Gibbs, 2003). In 2004, a full-day training on standards-based grading was offered at WCPSS's Continuous Improvement Conference. In 2005, the training offered at the Continuous Improvement Conference was 1½ hours of training geared toward new teachers and administrators. In 2005-06, WCPSS provided a training package to all schools with the expectation that IRTs would train new teachers on standards-based grading. This *K-5 Report Card: Facilitators Guide* includes a training script, grading rubric, and sample report cards. In addition to the description of levels of performance found on K-5 Profile cards (used to record a student's performance on literacy and mathematic objectives), in 2006-07, standards-based grading rubrics were added to literacy cards in grades 2-5 to provide guidance to teachers in assessing students and assigning grades. K-2 quarterly mathematic assessments also have rubrics for each objective.

In 2004-05 WCPSS surveyed 400 randomly selected K-5 teachers regarding the implementation and perceived effectiveness of standards-based grading. Overall, the majority of teachers found standards-based grading to be beneficial: 64% reported that they were better informed regarding student performance relative to the standard course of study; 58% felt it was more obvious why a student might be considered for retention; 52% felt it was a better method of communicating work habits, conduct, and homework; and 56% felt it was a better method for communicating with parents. While 94% of staff reported that they understood the level of achievement needed at each level of standards-based grading, 67% of teachers incorrectly reported that students need to master material beyond their grade level to receive a level 4. This was more prevalent among first-year teachers than their more experienced counterparts (75% versus 55% respectively) (Sarazen, 2005). In 2005 and 2006, in-depth interviews with two new teachers, an experienced teacher, and an IRT experienced with standards-based grading confirmed the surveys findings that while in general teachers struggled to articulate what constituted a level 4, new teachers found this particularly challenging. The survey also found that a higher percentage of K-2 teachers consistently rated standards-based grading more beneficial than did 3-5 teachers. However, a higher percentage of 3-5 teachers felt students, parents, and staff understood standards-based grading either fully or fairly well (Sarazen, 2005).

Discovering how grades are distributed, the correlation of grades to standardized test scores, and whether students' grades can predict their EOG scores will provide valuable insights into how standards-based grading has moved from policy to practice and the usefulness of the grading system.

RESEARCH QUESTIONS

Standards-based grades and EOG scores are two different measures of students' mastery of the N.C. Standard Course of Study. This study explores the distribution of grades assigned to students under standards-based grading and the extent to which grades are related to EOG scores.

1. What is standards-based grading?
2. How is the assessment of student performance separated from the assessment of work habits?
3. What is the distribution of 1-4 grades across grade levels and schools?
4. Are students' grades correlated with their EOG level scores?
5. Can a student's reading and mathematics grades be used to predict their performance on the EOG?

METHODOLOGY

Data Analysis

The data used for this study are the 2005-06 and 2006-07 student mathematics and reading grades and 2006-07 mathematics and reading EOG scores. Descriptive and correlational statistics were used to inform the study of standards-based grading and the connection of grades to the N.C. Standard Course of Study. Additionally, regression analyses were conducted to determine if student grades could predict performance on the EOG. Students' grades were correlated to EOG scores to determine the degree to which standard-based grades reflect students' knowledge of the Standard Course of Study. According to North Carolina's Department of Public Instruction (DPI), the EOG tests measure students' performance on the grade-level goals and objectives as stated in the N.C. Standard Course of Study (North Carolina Testing Program, 2003). While each form of the test represents a sample of objectives within the N.C. Standard Course of Study for that grade level, across all forms the curriculum is covered. Thus, EOG scores were used to measure mastery of the Standard Course of Study and were analyzed to determine the correlation of grades to the standard course of study as measured by the EOGs and whether those grades can predict a student's EOG score.

Reliability

Reliability refers to the ability to replicate results if the study conditions are repeated. According to North Carolina's DPI:

In testing, if use is to be made of some piece of information, then the information should be stable, consistent, and dependable ... if decisions about individuals are to be made on the basis of test data, then it is desirable that the test results be

reliable and tests exhibit a reliability coefficient of at least 0.85 (Bazemore, VanDyk, & Kramer, 2006, p. 62).

The metric used to establish the reliability of North Carolina's EOG and EOC tests of reading and mathematics is an internal consistency coefficient, coefficient alpha. Internal reliability "indicates how close the examinee's obtained score would come to the true score if the test were a perfect measuring instrument" (Charter, 2003, p. 1). Both reading and mathematics EOGs have a high degree of reliability. Analysis of the 1993 administration of the EOG test forms A, B, & C produced coefficient alpha 0.90 or greater for reading across grades 3-8 (Sanford, 1996). For mathematics the coefficient alpha indices averaged across forms were also high (> 0.94 for grades 3-8). While the 3rd grade pre-test coefficient alpha was 0.82, this test has half of the items of the grade 3 EOG (Bazemore, et al., 2006). Additional information on the reliability of North Carolina's EOG can be obtained at <http://www.ncpublicschools.org/accountability/testing/technicalreports>. It is important to consider the standard error of measurement when examining individual student level scores. Standard error of measurement is an estimate of the accuracy of a given score on a test using the standard deviation and the reliability of the test. Additional information regarding the standard error of measurement can be found at http://www.wcpss.net/evaluation-research/reports/2000/mment_error.pdf.

Validity

The validity of a test is the degree to which evidence and theory support the interpretation of test scores. Validity provides a check on how well a test fulfills its function. For all forms of test development, the validity of the test is an issue to be addressed from the first stage of development through analysis and reporting of scores. The process of validation involves accumulating evidence to provide a sound scientific basis for the proposed test score interpretations. Those interpretations of test scores are evaluated rather than the test itself (Bazemore et al., 2006, p. 87).

North Carolina's EOG tests are curriculum-based tests designed to measure the state's Standard-Course of Study. North Carolina's curriculum is updated approximately every five years, but the process of creating, field-testing, and administering the exams is continuous (North Carolina Testing Program, 2003). Validity of EOG is measured both in terms of content validity and criterion-related validity. North Carolina teachers write almost all test items and all items are reviewed by at least two teachers. Instructional validity is evaluated by DPI through questionnaires to teachers eliciting feedback on the appropriateness of test content. Criterion-related validity is assessed using Pearson correlation coefficient to measure the association between EOGs and items related to teacher judgment. "The correlation coefficients for the North Carolina EOG and EOC Tests of Mathematics range from 0.49 to 0.89 indicating a moderate to strong correlation between EOG scale scores and its associated variables" (Bazemore, et al., 2006, p. 89). Additional information on the validity of North Carolina's EOGs can be obtained at <http://www.ncpublicschools.org/accountability/testing/technicalreports>.

Population Demographics

The population considered in this study includes elementary students who were enrolled within WCPSS elementary schools with grades recorded in the electronic grade book data collection system (89 of the 93 elementary schools). Currently, grades are collected centrally within the electronic grading system. The population of this evaluation was limited to schools utilizing the electronic system. The population of the 89 schools used in this study is representative of the overall K-5 population in WCPSS. Table 1 displays the 2006-07 demographics for grades K-5 students attending WCPSS elementary schools that utilize the electronic grading data collection system.

Table 1
WCPSS Schools K-5 Demographics by Grade Level 2006-07

	Kindergarten	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Total
Asian	614	624	529	516	538	536	3,357
Black/African Am.	2,437	2,504	2,498	2,421	2,521	2,430	14,811
American Indian	33	24	30	30	25	27	169
Hispanic/Latino	1,559	1,485	1,266	1,157	1,068	1,007	7,542
White	5,403	5,453	5,183	5,337	5,126	4,730	31,232
Multiracial	594	538	506	465	422	417	2,942
Total	10,640	10,628	10,012	9,926	9,700	9,147	60,053
*FRL	3247	3,284	2,948	2,877	2,834	2,541	17,731
*SWD	885	1,093	1,207	1,349	1,426	1,349	7,309
*LEP	1,769	1,557	861	699	820	686	6,392

Data Source: WCPSS Student Locator November 2007 data

Note: Students may appear in more than one category: race and FRL, SWD, and/or LEP.

Note 2: Differences in totals reflect students with missing data for one or more variables.

Note 3: LEP = Limited English Proficient; SWD = Student with Disabilities; and FRL = Free or Reduced-Price Lunch

RESULTS

Question 1: What is standards-based grading?

The standards-based report card measures a student's academic progress against a uniform standard, rather than using subjective assessments. The report card includes a section for teacher comments designed to capture student strengths and areas of need (Understanding the Elementary School Report Card, 2005; K-5 Standards-Based Grading and Reporting, 2003). "The student performance levels of 1 to 4 indicate whether students have met the expectations set by the state in the Standard Course of Study and indicate whether the student has the necessary skills and concepts to be successful in the next quarter or next grade" (K-5 Standards-

Based Grading and Reporting: Fact Sheet, 2003, p.1). Table 2 displays the definitions for each of the 1-4 student performance levels.

Table 2
WCPSS Student Performance Levels

Level	Meaning
1	Insufficient performance of targeted grade level standards with support.
2	Inconsistent and needs support to meet targeted grade level standards.
3	Demonstrates proficiency of targeted grade level standard.
3*	Demonstrates proficiency of targeted grade level standards with evidence of application.
4	Extends targeted grade level standards.

Data Source: K-5 Standards-Based Grading and Reporting: Fact Sheet, 2003, p.1

Standards-based grading assesses a student's mastery of an objective (a minimum of three observations are required for each objective); and each observation is assigned a 1-4 rating. Student scores are not averaged; rather, when a student has three recorded observations with a performance level 3 or better for an objective, then the student has mastered that objective. Work habits are captured under a separate 1-3 rating.

Although standards-based grading has reduced the level of teacher subjectivity by measuring students against state-defined standard objectives, the process of combining objectives into subject grades for report cards still requires teacher judgment in reconciling students' final grades. WCPSS' *K-5 Standards-Based Grading and Reporting Guide to Reporting Progress* addresses the composition of student grades and the teacher's role in producing them. Teachers assess and evaluate student performance in terms of practice (formative) and evaluation (summative) evidence to determine a student's grade. Within WCPSS's elementary schools, students' grades are not based on percentages; rather, after an objective has been observed three times, students receive a 1-4 performance level reflecting the student's mastery of the N.C. Standard Course of Study (K-5 Standards-Based Grading and Reporting Guide to Reporting Progress, 2003). Teachers must consider all reading objectives to determine the student's final reading grade, which is then reflected on the student's report card.

Question 2: How is student performance separated from work habits?

As a result of the move to standards-based grading, WCPSS' K-5 elementary report cards now separate student performance from conduct and work habits. As required by WCPSS' Homework Regulations and Procedures, "Homework is considered practice in grades K-5; therefore, it is reflected in the Work Habits grade" (WCPSS 5510 R&P).

In reporting on conduct, the teacher can indicate whether the student meets expectations in cooperating with others, respecting others, and observing

rules and procedures. In reporting on work habits, the teacher can indicate whether the student uses time wisely, listens carefully, completes assignments, writes legibly, works independently or seeks help when needed, and completes work (Understanding the Elementary School Report Card, 2005, p. 1).

Students are assigned a score of 1-3 for their conduct and work habits, where a 3 indicates the student met expectations, a 2 indicates the student met expectations inconsistently, and a 1 indicates that the student failed to meet expectations (Understanding the Elementary School Report Card, 2005).

Question 3: What is the distribution of 1-4 grades across grade levels and schools?

Grade Level

The electronic grade book requires teachers to record 1, 2, 3, 3*, or 4 by subject quarterly. The distribution of grades by grade level and school reflect the recorded fourth-quarter mathematics and reading scores. Tables 3 and 4 show fourth-quarter grades—1-4 student performance levels—by grade level in reading and mathematics in 2005-06 and 2006-07. The grades of students attending schools not utilizing the electronic grade book are not available at the district level and are thus not represented in these tables. In 2006-07, nearly all elementary schools (95%) reported grades within the electronic grade book; however, in 2005-06 considerably fewer schools had used this tool. Although the number of student grades recorded in 2005-06 was lower than the number of grades reported in 2006-07, the distribution of 1-4 grades by grade level remained relatively constant.

Table 3 depicts the number of students receiving a 1-4 performance level by grade level. The modal (most often occurring) reading score of the 20,656 students with a valid score in 2005-06 and for the 58,137 students with a valid score in 2006-07 was 3. The second most common performance level was a level 4 among K-2 students and a 3* among 3-5 students. Thus, students in grades 3-5 were less likely to receive a level 4 than were K-2 students. Whether this is due to a more difficult curriculum at grades 3-5 or tougher grading standards among teachers is not discernable from these data.

Table 3
Fourth-Quarter Reading Grades
by Grade Level 2005-06 and 2006-07

Grade	Year	Fourth-Quarter Reading Grade					Total
		1	2	3	3*	4	
Kindergarten	2005-06	141	463	1,488	836	850	3,778
	2006-07	483	1,263	4,294	2,015	2,262	10,317
Grade 1	2005-06	268	353	1,171	709	1,000	3,501
	2006-07	852	1,166	3,556	1,896	2,883	10,353
Grade 2	2005-06	148	362	1,435	709	795	3,449
	2006-07	444	1,153	4,078	1,839	2,085	9,599
Grade 3	2005-06	68	528	1,563	679	560	3,398
	2006-07	223	1,549	4,411	1,786	1,622	9,591
Grade 4	2005-06	61	553	1,473	630	514	3,231
	2006-07	188	1,438	4,560	1,848	1,359	9,393
Grade 5	2005-06	42	440	1,632	686	499	3,299
	2006-07	147	1,112	4,372	1,740	1,513	8,884
K-5 Total	2005-06	728	2,699	8,762	4,249	4,218	20,656
	2006-07	2,337	7,681	25,271	11,124	11,724	58,137

Note: **Blue** font indicates the most common performance level within a grade level.

Purple font indicates the second most common performance level within a grade level.

Data Source: 2005-06 and 2006-07 grade files

The modal mathematics scores of the 20,672 students with a valid score in 2005-06 was 3, which was also true for the 58,134 students with a valid score in 2006-07. As shown in Table 4, the pattern of fourth-quarter mathematics grades was similar to reading, in that performance level 3 was the most common grade received by K-5 students and that K-2 students were more likely to receive a higher performance level than 3-5 students. However, unlike reading, the second most common mathematics grade received among K-2 students was 3*, while among 3-5 students the second most common grade received was a 2. Thus, it appears grading standards were more difficult for mathematics than reading across the K-5 grade levels and K-2 students were more likely to receive at-grade-level grades than their 3-5 counterparts.

Table 4
Fourth-Quarter Mathematics Grades
by Grade Level 2005-06 and 2006-07

Grade	Year	Fourth-Quarter Mathematics Grade					Total
		1	2	3	3*	4	
Kindergarten	2005-06	71	440	2,028	841	410	3,790
	2006-07	248	1,195	5,902	1,953	1,020	10,318
Grade 1	2005-06	93	524	1,792	687	403	3,499
	2006-07	356	1,573	5,325	1,991	1,102	10,347
Grade 2	2005-06	111	605	1,612	720	403	3,451
	2006-07	348	1,771	4,417	1,864	1,197	9,597
Grade 3	2005-06	62	505	1,635	640	553	3,395
	2006-07	231	1,839	4,415	1,659	1,451	9,595
Grade 4	2005-06	62	599	1,545	597	435	3,238
	2006-07	215	1,833	4,565	1,562	1,220	9,395
Grade 5	2005-06	44	535	1,599	656	465	3,299
	2006-07	184	1,637	4,231	1,573	1,257	8,882
K-5 Total	2005-06	443	3,208	10,211	4,141	2,669	20,672
	2006-07	1,582	9,848	28,855	10,602	7,247	58,134

Note: **Blue** font indicates the most common performance level within a grade level.

Purple font indicates the second most common performance level within a grade level.

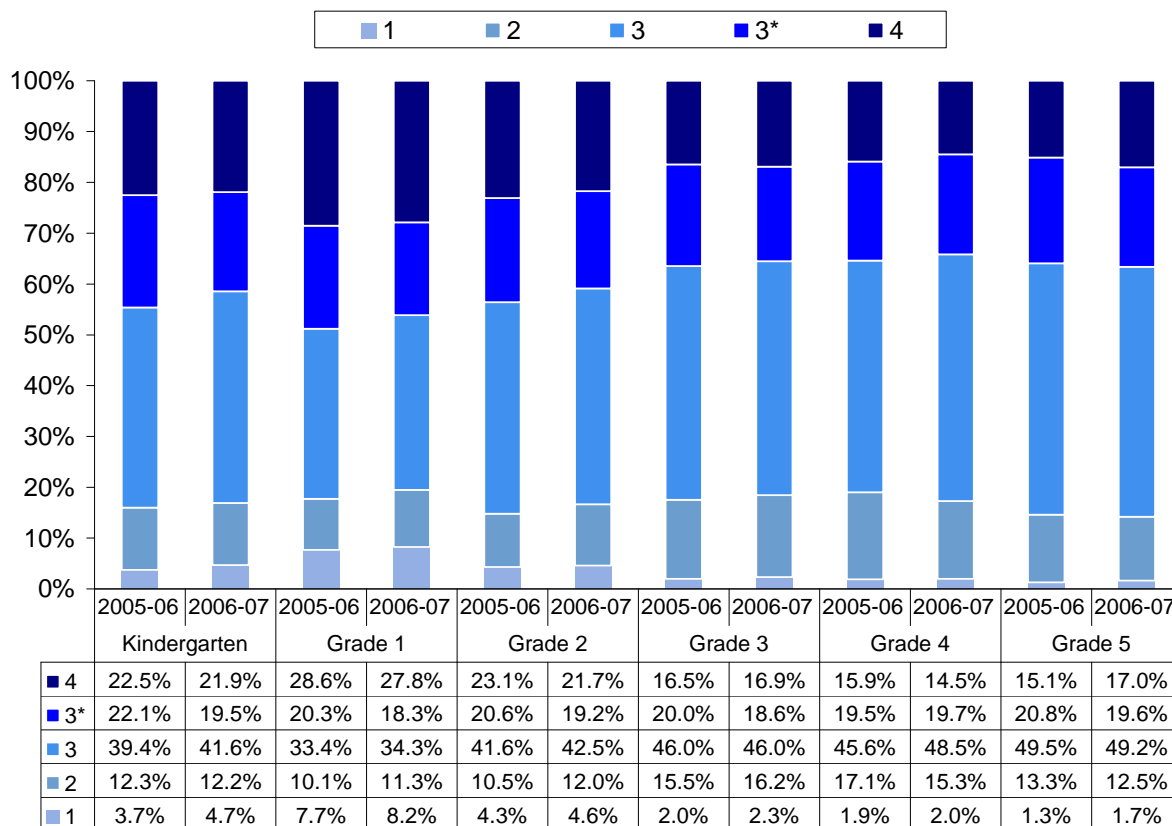
Data Source: 2005-06 and 2006-07 grade files

Figures 1 and 2 display the distribution of grades students received by performance level and grade level for reading and mathematics in terms of percentages in 2005-06 and 2006-07. As depicted in Figure 1, across grade levels the pattern of student grades was relatively consistent. The percentage of students receiving reading grades reflecting mastery of grade level material (3, 3*, or 4) was relatively consistent across the elementary grades (ranging from 81%-86%). However, there were some interesting differences between grade spans K-2 and 3-5. While students in the lower elementary grades, K-2, were more likely to have a fourth-quarter reading grade at level 4 than were students in grades 3-5, students in K-2 were also more likely to receive a grade at level 1 than those in grades 3-5. Thus, scores were more widely distributed at K-2.

- The percentage of students with a level 1 (insufficient performance on standards) varied by grade level from less than 2% of students in grade 5 to more than 8% of students in grade 1.
- Among students with a level 2 (inconsistent performance and required support to meet standard), the range was from 11% of grade 1 students to more than 16% of grade 3 students.
- Level 3 (demonstrates proficiency) was the most common grade across all grade levels, varying from one third of grade 1 students to nearly half of students in grade 5.

- Across grade levels, approximately 20% of students received a 3* (proficiency with evidence of application), varying from 18% at grade 1 to 20% at grade 5.
- The percentage of students with a level 4 (extends targeted standards) ranged from 15% of students in grade 4 to 28% of students in grade 1.

Figure 1
Percentage of Students by Grade Level and
Grade Performance Level in Reading
2005-06 and 2006-07



Data Source: 2005-06 and 2006-07 grade files

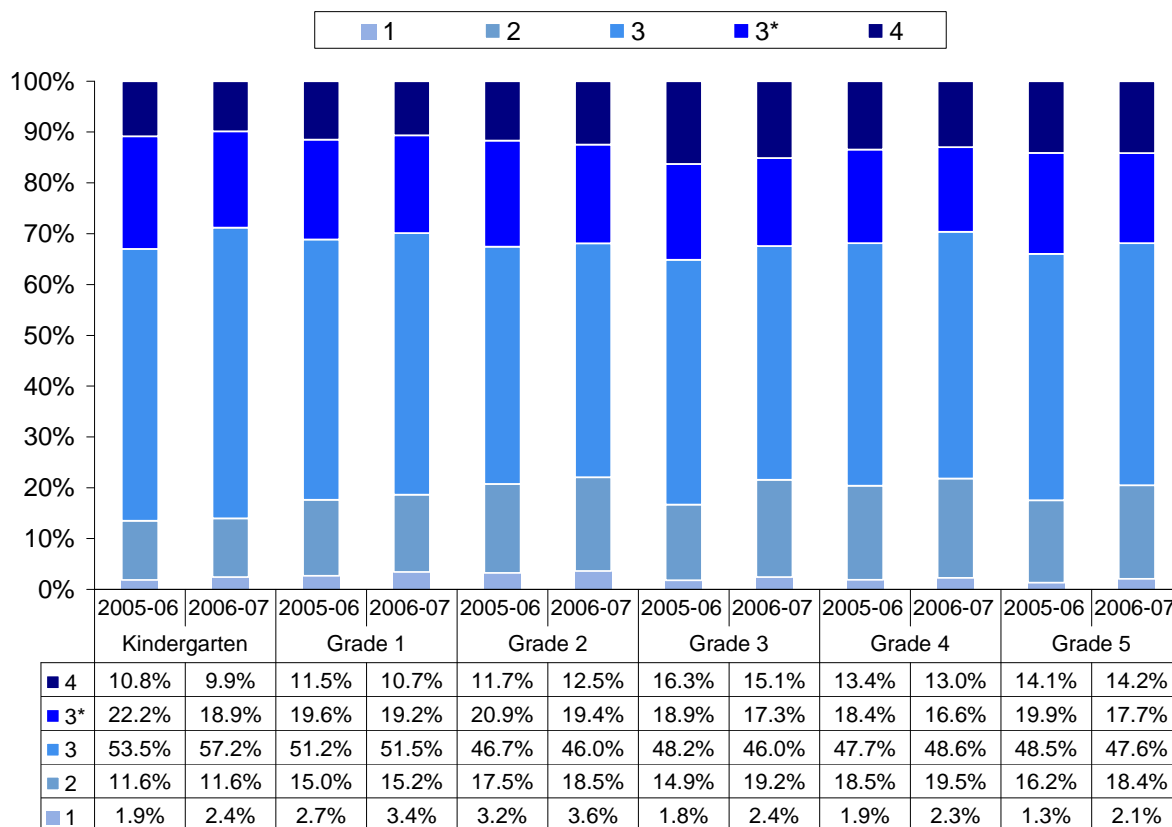
Interpretation Example: Among kindergarten students in 2006-07, 21.9% received a level 4, 19.5% received a 3*, and 41.6% received a level 3, thus, only 16.9% of kindergarten students were not proficient.

Figure 2 presents the distribution of fourth-quarter mathematics grades among WCPSS elementary students. The distribution of 1-4 mathematics grades by grade level remained relatively constant between 2005-06 and 2006-07.

Across grade levels, the percentage of students receiving a grade reflecting mastery (3, 3*, or 4) ranged from 78% to 86%. In mathematics, similar to reading, the pattern of student grades was relatively consistent across grade levels. However, unlike reading, students in the lower elementary grades, K-2, were slightly less likely to have a fourth-quarter mathematics grade at level 4 than were students in grades 3-5. The percentage of students at level 4 was lower in mathematics than reading at all grade levels and more evenly distributed by grade.

- The percentage of students with a level 1 (insufficient performance on standards) varied by grade level from less than 2% of students in grade 5 to 4% of students in grade 2.
- Among students with a level 2 (inconsistent and required support to meet standard) the range was slightly larger, ranging from 12% of kindergarten students to 20% of grade 4 students.
- Level 3 was the most common grade across all grade levels, varying from 46% of grade 2 and 3 students to 57% of kindergarten students.
- Across grade levels, the percentage of students receiving a 3* was fairly consistent, varying from 17% at grade 4 to 19% at grade 2.
- The percentage of students with a level 4 ranged from 10% of kindergarten students to 15% of students in grade 3.

Figure 2
Percentage of Students by Grade Level and
Grade Performance Level in Mathematics
2005-06 and 2006-07



Data Source: 2005-06 and 2006-07 grade files

Interpretation Example: The percentage of students in grade 2 with a level 4 grade increased slightly, from 11.7% in 2005-06 to 12.5% in 2006-07.

School

Among 93 WCPSS elementary schools in 2006-07, the percentage of students with a 1, 2, 3, 3*, or 4 for their fourth-quarter mathematics grade varied by school. Appendix A and B displays the grade performance level by elementary school for reading and mathematics. Due to the large number of elementary schools the mathematics and reading results are separated into three figures for each subject. The percentage of students at each reading performance level varied by school ranging from:

- Less than 1% to 13% of students at level 1,
- 4% to 25% for level 2,
- 29% to 61% of students at level 3,
- Less than 1% to 31% for level 3*, and
- 3% to 40% of students at level 4.

The percentage of students at each mathematics performance level also varied by school ranging from:

- Less than 1% to 9% of students at level 1,
- 5% to 33% for level 2,
- 32% to 63% of students at level 3,
- 0% to 29% for level 3*, and
- 9% to 23% of students at level 4.

Variations by school may be due to real differences in performance or differences in applying standards. One notable difference was in the range of percentages of students receiving a level 4 was considerably greater for reading (37 points) than mathematics (14 points). By school differences of 30 percentage points or more seem larger than expected thus may deserve further study.

Question 4: Are student grades correlated with their End-of-Grade (EOG) scores?

Students in kindergarten through grade 2 are not assessed by an EOG. Thus, the remainder of this study will focus on students in grades 3, 4, and 5 in 2006-07. The 2006-07 school year was selected since the data for that year are more complete, with a greater number of schools entering data into the electronic grading system.

A Spearman's rho was computed to evaluate the relationship between EOG level and fourth-quarter grade. A moderately positive relationship exists between EOG level and fourth-quarter grade (see Table 5 for results of the correlation analysis). The correlation coefficient demonstrates how well the two variables are correlated: the closer to 1 denotes a more linear relationship while the closer to 0 denotes little or no relationship.

In mathematics, the correlations between a student's grade and EOG score were similar across grade levels, students' fourth-quarter mathematics grades had a higher correlation to their mathematics EOG ($>.72$) than their reading grade did to the reading EOG (see Table 5).

Table 5
Correlation of Fourth-Quarter Grade and EOG
by Grade Level 2006-07

	Grade Level	Number of students	EOG
Reading Grade (1-4)	3	n=9,542	.67
	4	n=9,335	.63
	5	n=8,831	.61
Math Grade (1-4)	3	n=9,543	.74
	4	n=9,337	.72
	5	n=8,827	.72

Data Source: 2006-07 grade file and 2006-07 End-of-Year Student Roster.

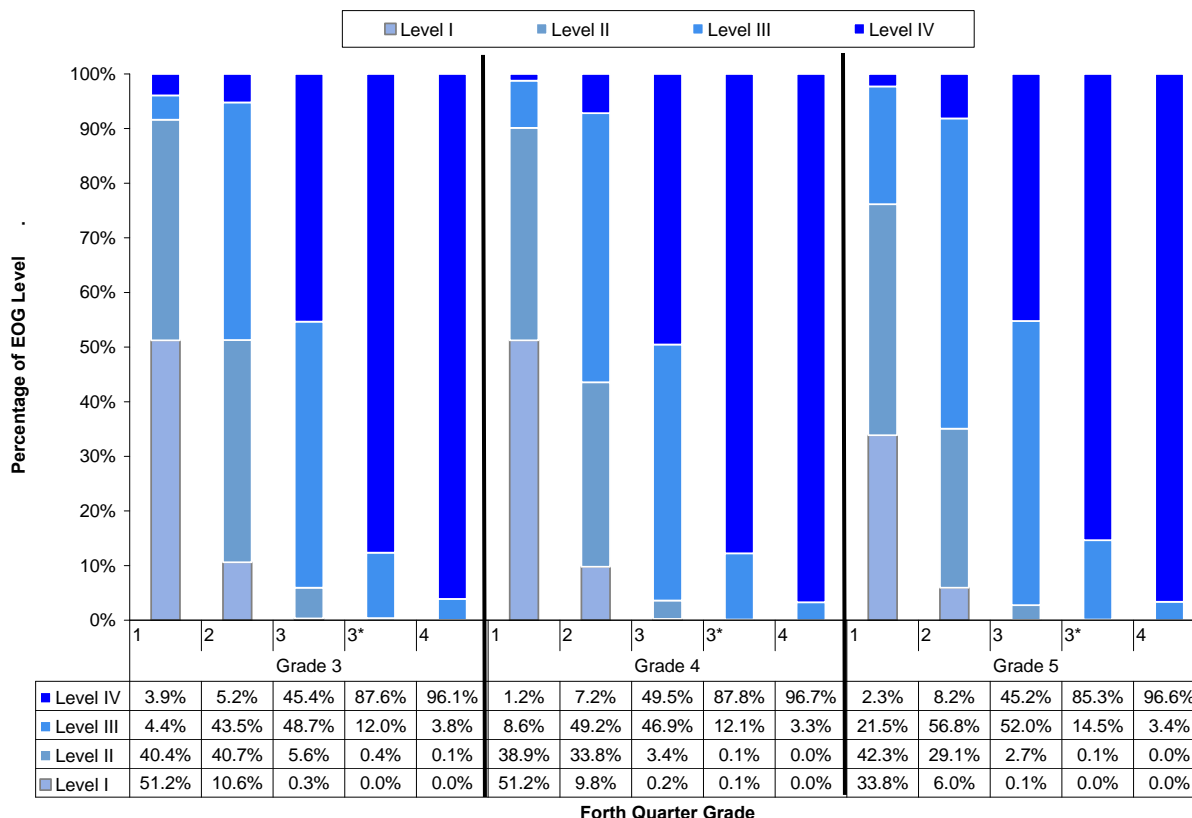
Interpretation Example: Among fifth grade students' fourth-quarter mathematics grade had a higher correlation to their mathematics EOG (0.72) than their reading grade did to the reading EOG (0.61).

To further examine the relationship between student reading grade and reading EOG score Figures 3 and 4 display the percentage of students at each performance grade level by EOG level. The percentage of students by performance level and EOG level were similar across grade levels. For reading, the vast majority (>96%) of students who were on grade level based on their performance level within the classroom were also at or above grade level on their EOG. However, for students who were below grade level based on classroom grades, the percentage that scored proficient on the EOG ranged by grade level from 44% to 61%. Thus, while a student performing on grade level within the classroom was highly likely to be successful on the EOG, for students performing below grade level within the classroom approximately half were successful on the EOG.

While the performance level within the classroom and EOG levels were not intended to have a one-to-one correspondence, the fact that the more than three fourths (>85%) of students scoring a level 3* scored a level IV on their EOG, may indicate that classroom teachers have grading practices which are more stringent than the EOG standards. This may also call in to question the necessity of having a 3* as part of the standards-based performance levels.

- Of students with a reading grade of level 1, approximately half of students in grades 3 and 4 and one-third of grade 5 students scored a level I on their reading EOG.
- Among students with a level 2 on their fourth-quarter reading grade, approximately 50% or more scored on grade level on their reading EOG.
- Of students with a level 3 reading grade nearly all (>94%) were on grade level scoring a level III or IV on their reading EOG.
- The vast majority (>85%) of students with a 3* scored a level IV on their EOG.
- Almost all students (>96%) with a 4 on their fourth-quarter grade scored a level IV on their EOG in grades 3-5.

Figure 3
Percentage of Students by
Grade Performance Level in Reading and Reading EOG Level
2006-07



Data Source: 2006-07 grade file and 2006-07End-of-Year Student Roster.

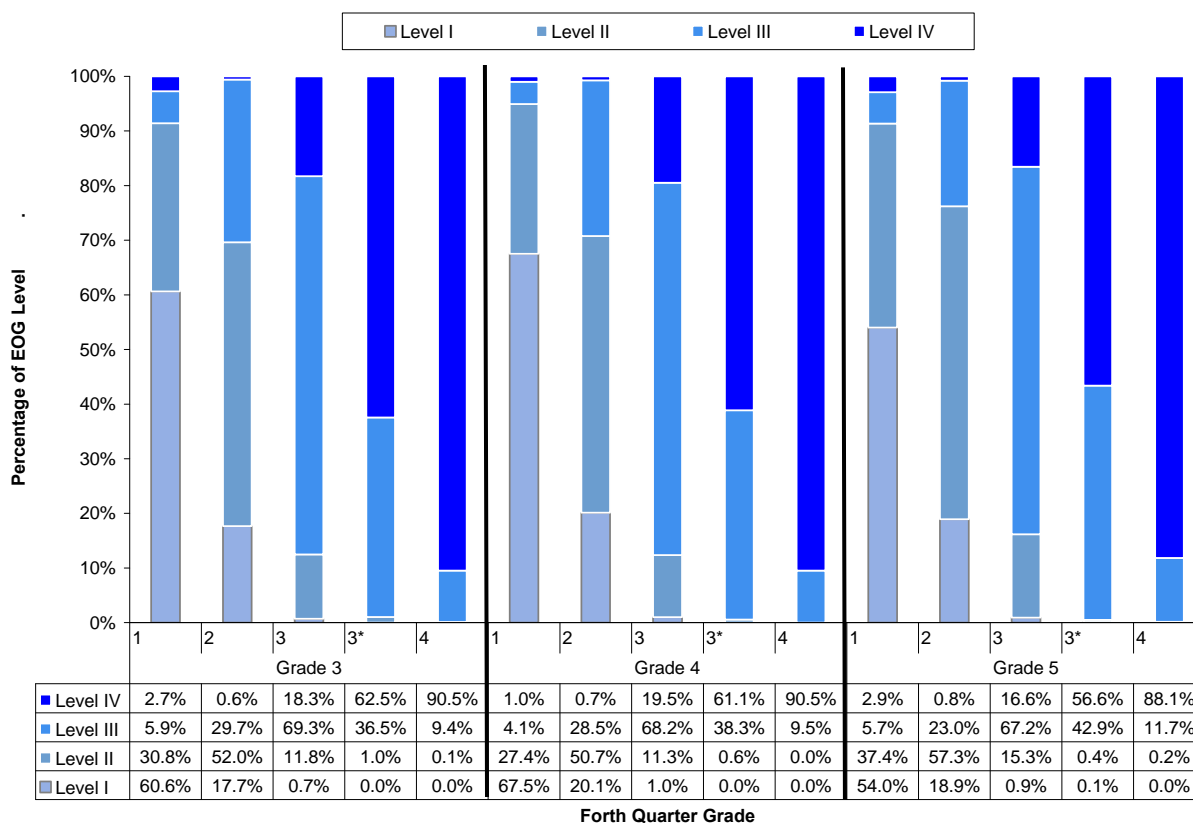
Interpretation Example: Among grade 3 students who had a fourth-quarter reading grade of 3*, 87% scored a level IV and 12% scored a level III on the reading EOG.

The relationship between student mathematics grade and mathematics EOG score was comparable across grade levels. While similar to reading, the vast majority (>90%) of students receiving a grade reflecting mastery (3, 3*, or 4) within the classroom were at or above grade level based on their EOG score, a greater percentage of students performing below grade level within the classroom were also below grade level on their EOG (72% to 78%). However, as with reading, students scoring below grade level on their fourth-quarter grade were more likely to be on grade level based on the EOG (22% to 28%) as compared to the smaller percentage (<10%) of students on grade level within the classroom who are unsuccessful on the EOG. Again, this indicates that classroom teachers may be tougher graders or that standards-based grading are more difficult standards than on the EOG exam.

- The vast majority (>91%) of students with a level 1 mathematics grade scored below grade level on their mathematics EOG, with the majority scoring a level I.
- Among students with a level 2 on their fourth-quarter mathematics grade, approximately 50% scored a level II on their mathematics EOG.

- Of students with a level 3 mathematics grade, the vast majority (>83%) scored at or above grade level on their mathematics EOG, with the majority scoring a Level III.
- Nearly all (approximately 99%) of students with a 3* mathematics grade scored at or above grade level with the majority scoring level IV on their mathematics EOG.
- The vast majority (>88%) of students with a level 4 mathematics grade scored level IV on their mathematics EOG.

Figure 4
Percentage of Students by
Grade Performance Level in Mathematics and Mathematics EOG Level
2006-07



Data Source: 2006-07 grade file and 2006-07 End-of-Year Student Roster.

Interpretation Example: Among grade 3 students with a fourth-quarter mathematics grade of 3*, 62.5% scored a level IV and 36.5% scored a level III on the mathematics EOG.

Question 5: Can a student’s reading and mathematics grades be used to predict their performance on the EOG?

In order to assess the predictive ability of a student’s mid-year classroom performance on their EOG score at the end of the year, a logistic regression analysis was conducted to estimate whether students’ second-quarter grades could predict students’ EOG scores. Student performance on the reading and mathematics EOGs were dichotomized into on-grade level and below grade level. The logistic regression model was used to estimate the predictor variables

(classroom grade, prior EOG score, free or reduced price lunch status, limited English proficiency, and disability status) effect on the likelihood of scoring proficient on the EOG. Students meeting the grade level standard (receiving a 3, 3*, or 4), have significantly greater odds of being on-grade level on the reading and mathematics EOG (see Appendix C). For students who are on grade level based on their reading second-quarter grade the odds of them being on grade level on the reading EOG ranged from 16% greater odds among grade 5 students to 46% greater odds among grade 3 students. While for students on grade level based on their mathematics second-quarter grade, the odds of being on grade level on the mathematics EOG ranged from 12% of grade 4 students to 35% of grade 3 students. Additional regression analysis using the same variables revealed student grades were a much stronger predictor than the demographic variables and worked as well as prior EOG scores.

DISCUSSION AND RECOMMENDATIONS

These results suggest that WCPSS teachers grade students realistically or a little low relative to the standard of grade-level performance as measured on the EOG assessments. Standards-based grading was not formally designed to match the EOG, but it is reasonable to expect a fairly high correlation between grades and EOG scores. This is reassuring that standards-based grading is an appropriate grading system. The fact that students tend to do as well or better on the EOG than in their class grades is a good trend in that students are more likely to be prepared for the EOG tests than if the opposite trend was true. In addition, the fact that second-quarter grades are strong predictors of EOG performance means that teachers are in a good position to use these data mid-year to recommend students for additional help if they are receiving grades below grade level standard (level 1 or 2) in reading or mathematics. It is also reassuring that grades based on classroom performance are better predictors of achievement test scores than demographic variables.

For reading and mathematics the modal grade received was a level 3 across all grade levels. However, one notable difference was that K-2 students were more likely to score higher in both reading and mathematics than their 3-5 counterparts. Whether the curriculum at grades 3-5 is more difficult or there are tougher grading standards among teachers at the upper elementary grades is beyond the scope of this study. In addition, mathematics reflected a more difficult standard than reading across the K-5 grade levels.

The wide range in the percentage of students at each performance level across schools deserves further discussion and study. If schools with similar EOG score patterns have wide variation in grading practices, this suggests the standards are being applied differently across schools. This could mean additional training is needed, or perhaps rubrics and examples that are more extensive be provided to schools. One way to gain more insight on differences and similarities across teachers would be to pair grade-level teams with students with similar EOG scores but dissimilar grade distributions to discuss standards used based on concrete examples of students' work. Further qualitative analysis could be considered to determine why grades are not uniformly distributed across the system or to investigate the differences between schools with the strongest correlation between student grades and EOG scores as compared to schools with the weakest correlation.

Mathematics and reading fourth-quarter grades were highly correlated to EOG scores. Not only were students' classroom mathematics grades lower than reading at all grade levels, but mathematics grades had a stronger correlation to the newly revised mathematics EOG than did reading grades to the current reading EOG. The strong correlation between mathematics grades and EOG scores indicates that the newly revised EOG is a better match to what our students are expected to know within the classroom. In 2007-08, the reading EOG was revised. Checking the correlation of reading grades with the revised reading EOG could be helpful in determining whether the correlation between the two measures of students' knowledge of the North Carolina's Standard Course of Study improves.

The Evaluation and Research Department plans to study the correlation of middle school grades, which are the traditional A-F model, to students' EOGs. Examining the A-F grading system's correlation to EOG scores will provide insight into whether the traditional grading system correlates with, and adequately reflects students' knowledge of the N.C. Standard Course of Study. Comparing elementary and middle school correlations can spur interesting discussions of appropriate grading practices.

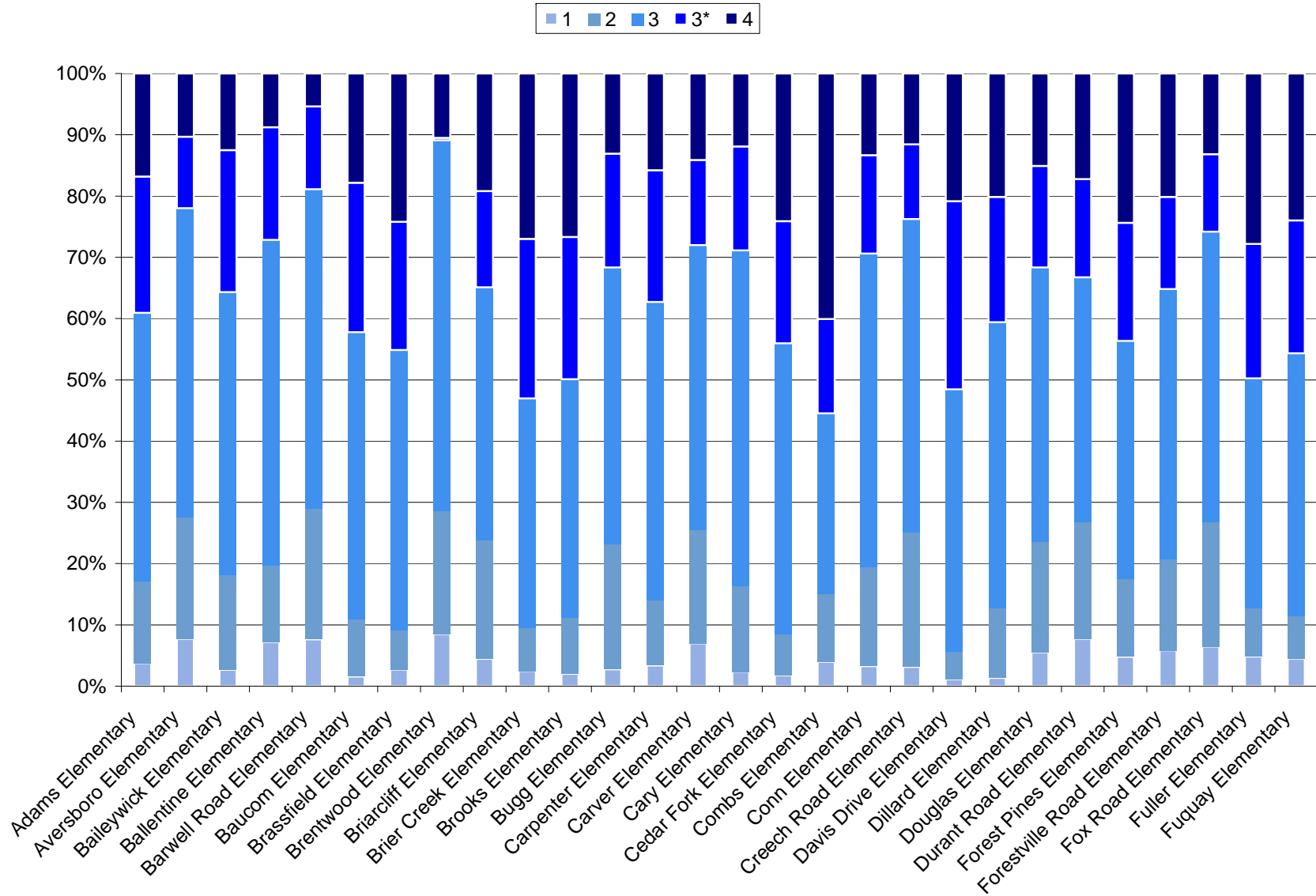
Finally, this evaluation did not address the second stated goal of the standards-based grading policy—to prepare students for success in the next grade or quarter. While the percentage of students who were below grade level within the classroom ranged from 14% to 22%, the range of students retained across grades K-5 ranged from 0.5% at grade 5 to 4.7% of grade 1 students. Thus, students who are not proficient within the classroom are being promoted to the next grade level. While students performing below grade level within the classroom may be promoted to the next grade level with instructional assistance, since a primary goal of standards-based grading is to prepare students to be successful in the next quarter or grade level, the disparity between the percentage of students below grade level and the percentage retained could be examined. One way to evaluate the success of this policy for this second goal would be to examine student retention rates and student success in the next grade.

REFERENCES

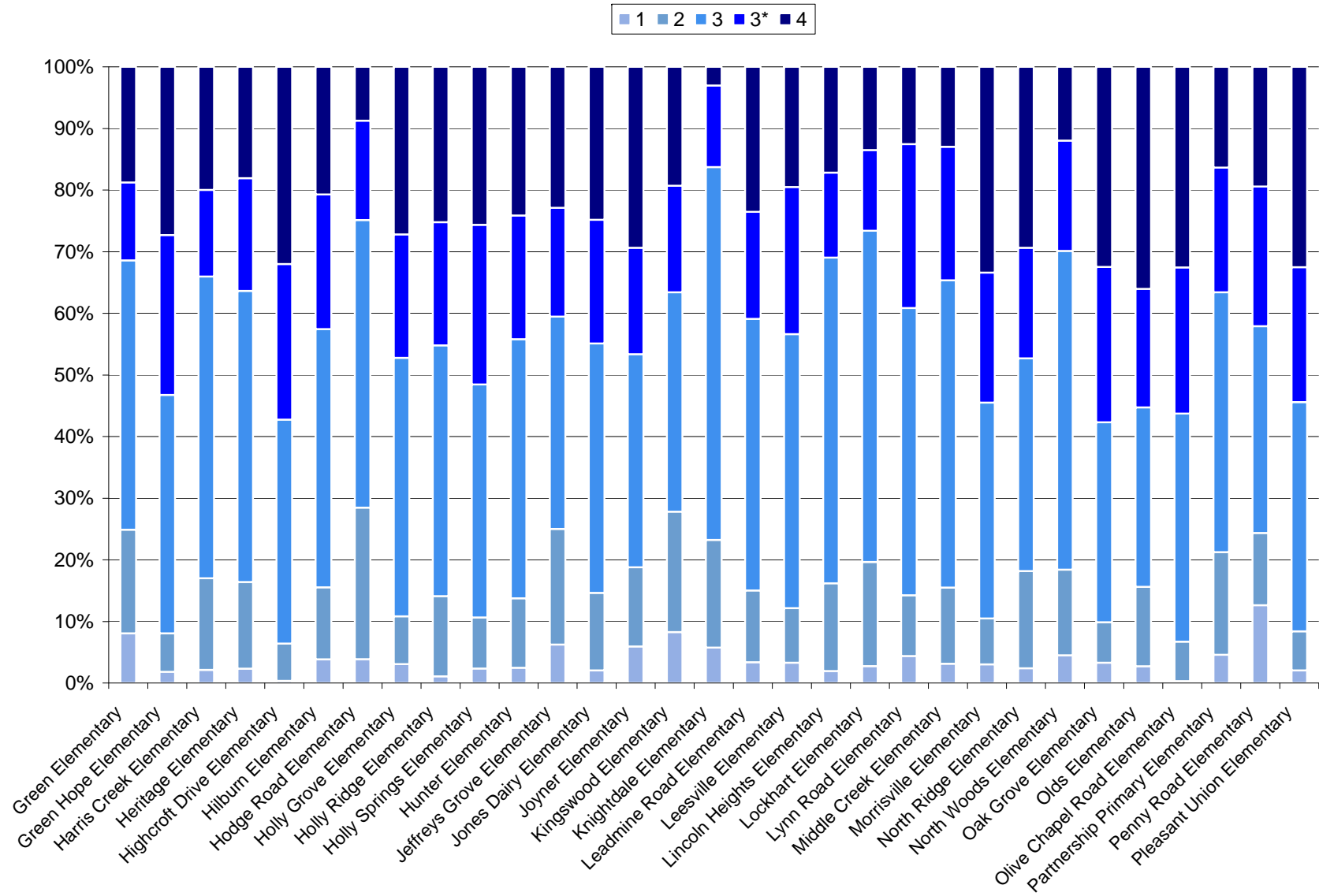
- Bazemore, M., VanDyk P., and Kramer, L. (2006). *The North Carolina mathematics technical report*. North Carolina Department of Public Instruction. Retrieved April 14, 2006, from <http://www.ncpublicschools.org/docs/accountability/testing/mathtechmanual.pdf>
- Charter, R.A. (2003). A breakdown of reliability coefficients by test type and reliability method, and the clinical implications of low reliability. *Journal of General Psychology*. Retrieved April 28, 2006, from http://www.findarticles.com/p/articles/mi_m2405/is_3_130/ai_107124690
- Cross, L. H. & Frary, R. B. (1999). *Hodgepodge grading: Endorsed by students and teachers alike*. *Applied Measurement in Education*, 12(1), 53-71.
- Guskey, T. R. (2001). Helping standards make the grade. *Educational Leadership*, 59(1), 20-27.
- K-5 standards-based grading and reporting: Involves the implementation of a new report card based on the student's performance of the Standard Course of Study*. Principals Edition. (2003). Fact Sheet on new WCPSS Elementary Report Card (NC curriculum). Wake County Public School System.
- Manzo, K. M. (2001). Districts tinker with report cards to make better sense of standards. *Education Week*, 21(4), 1-3.
- North Carolina Department of Public Instruction. (2003). *North Carolina testing program: multiple-choice test development process*. Retrieved April 14, 2006, from <http://www.ncpublicschools.org/docs/accountability/testing/policies/mctestdevelopment/RevisedTestDevelopmentProcessFinalFinal.pdf>
- Perlstein, L. (2003, April 14). Schools seek a higher standard for grades [Electronic version]. *The Washington Post*, 1-3.
- Sanford, E. E. (1996). *Technical report #1: North Carolina end-of-grade tests reading comprehension, mathematics*. North Carolina Department of Public Instruction.
- Sarazen, E. (2005). *Standards-based grading: teacher perspectives*. Unpublished manuscript.
- U.S. Department of Education (1994). *Research report: What do student grades mean? Differences across schools*. Office of Research.
- Wake County Public School System. *Homework policy 5510 R&P*. Retrieved April 24, 2008, from http://www2.wcpss.net/departments/c-and-i/homework_policy.htm
- Wake County Public School System. *Understanding the elementary school report cards*. Retrieved November 14, 2005, from <http://www.wcpss.net/curriculum-instruction/report-cards/elementary.html>
- Webb-Gibbs, K. (2003, May 14). Different grading for area schools: Ballentine, Lincoln Heights schools to participate. Fuquay-Varina Independent, pp.1, 14.

APPENDIX A

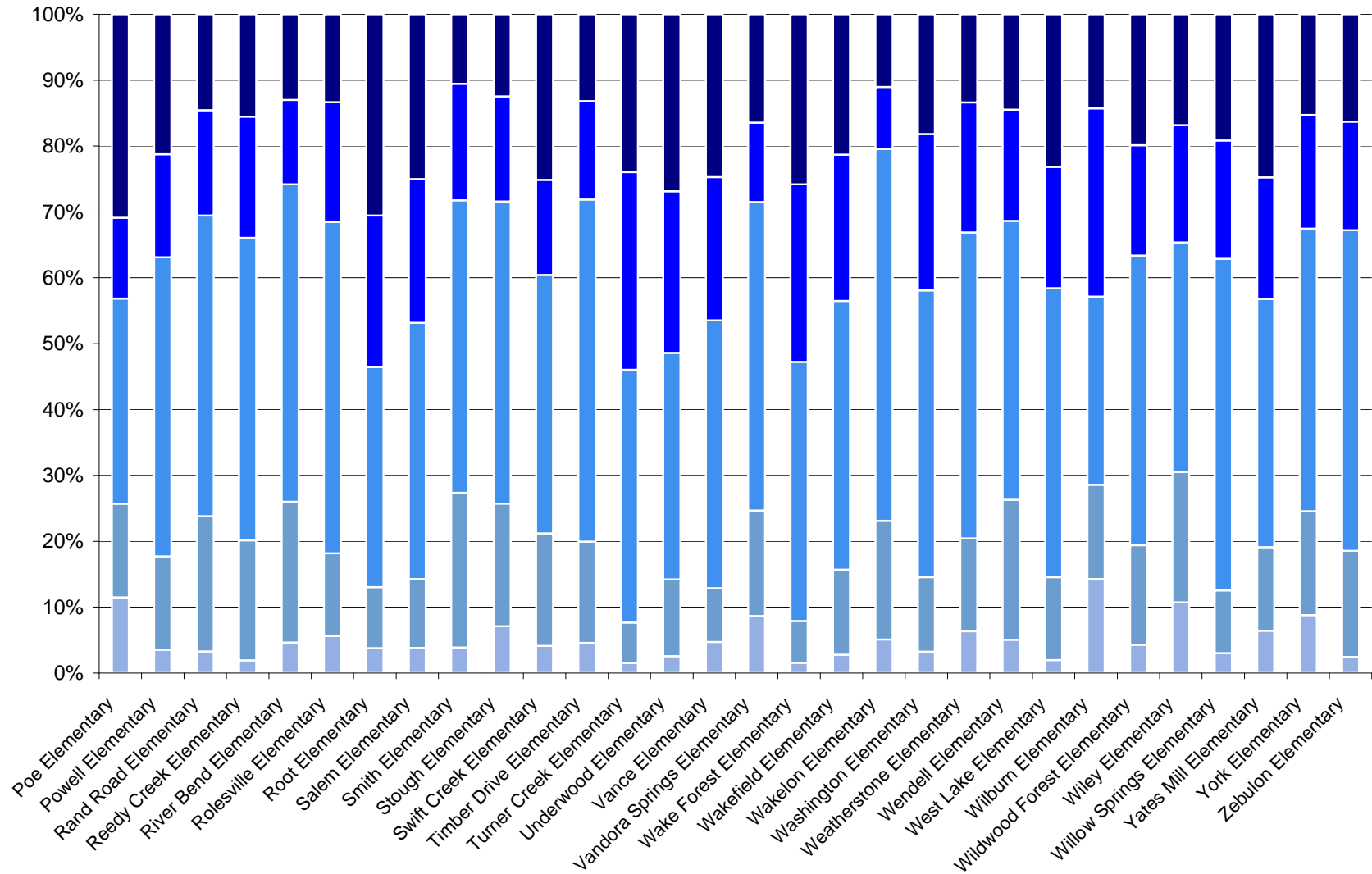
Percentage of Students by School and Grade Performance Level in Reading 2006-07



Percentage of Students by School and Grade Performance Level in Reading 2006-07

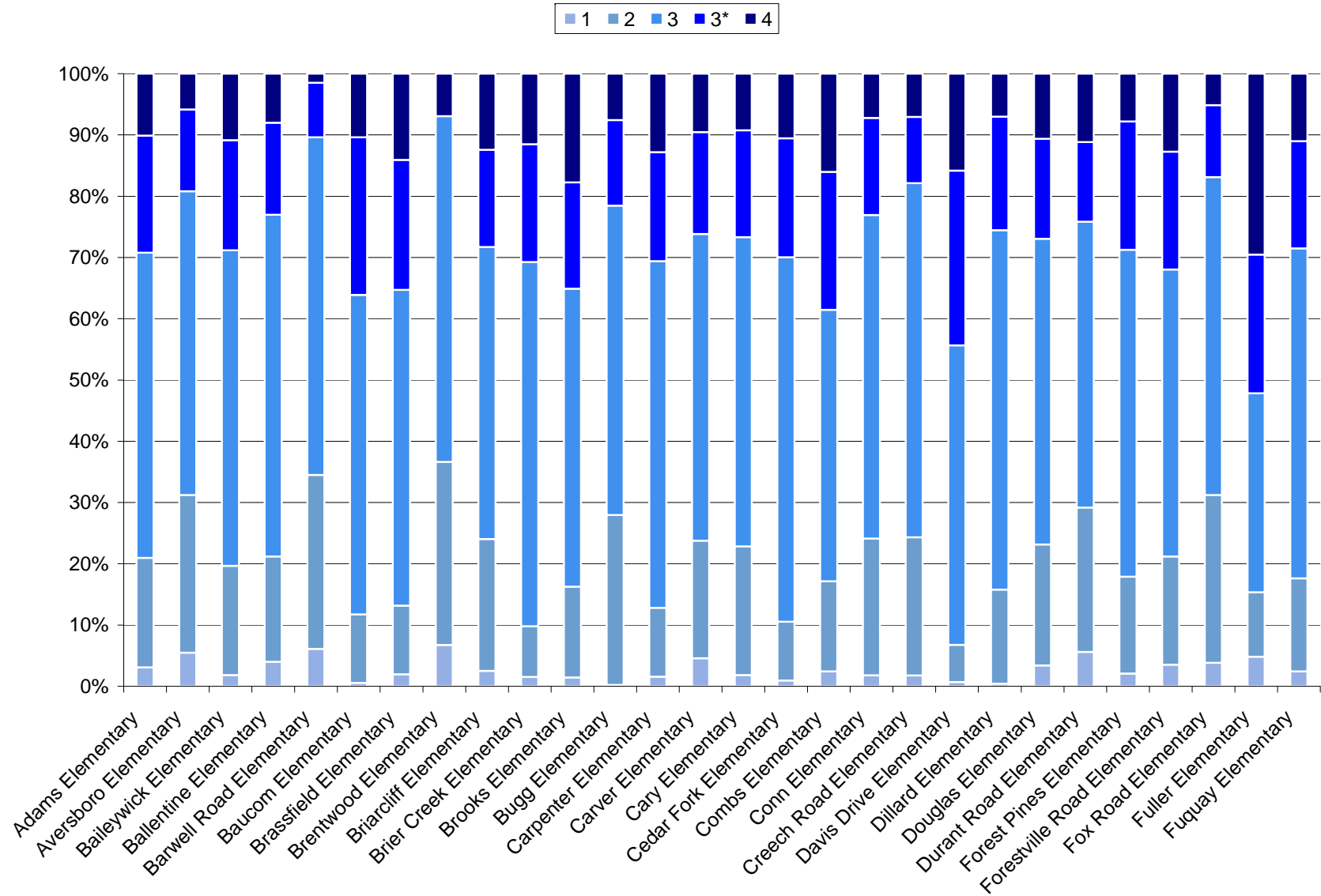


Percentage of Students by School and Grade Performance Level in Reading 2006-07

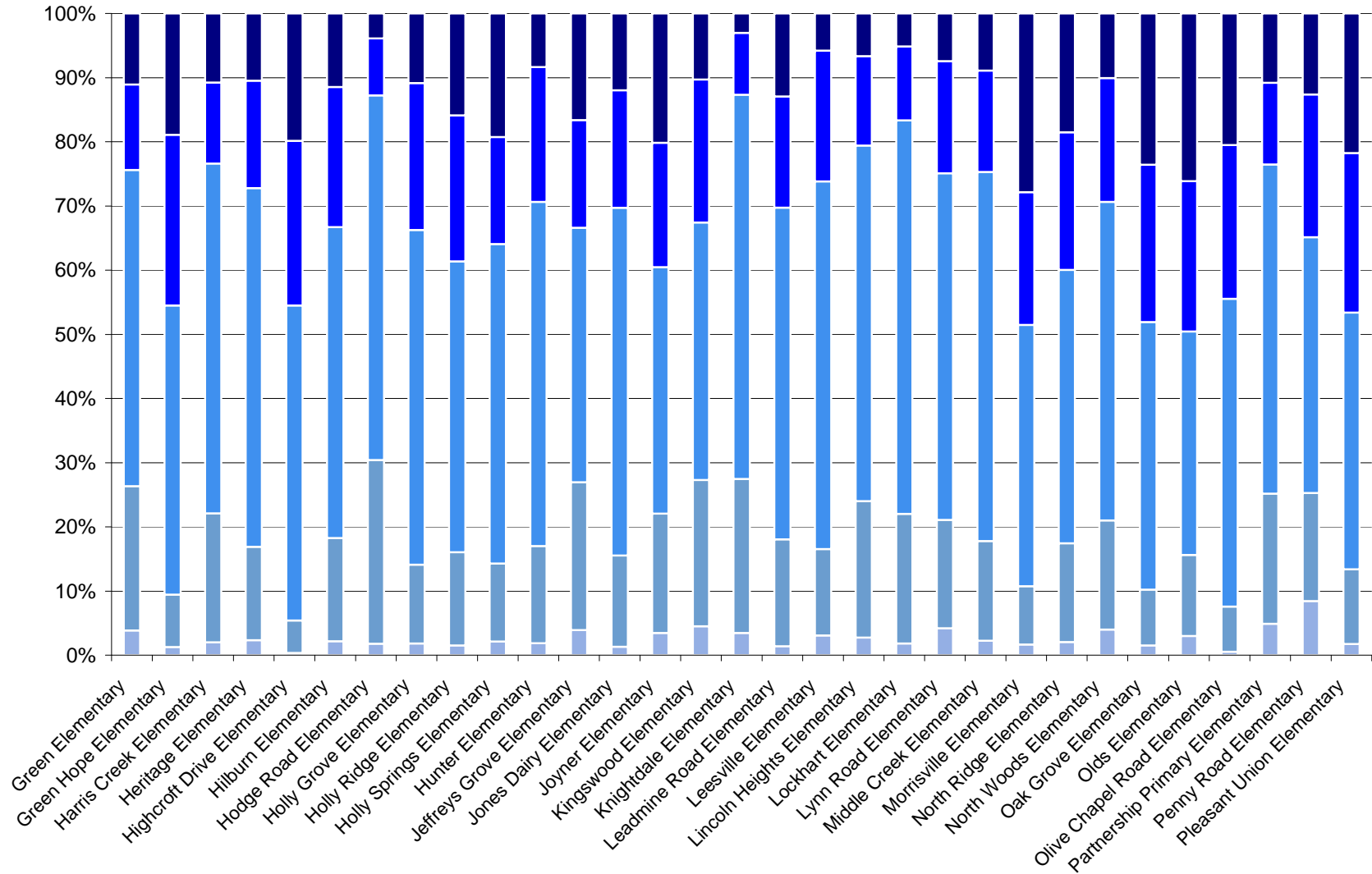


APPENDIX B

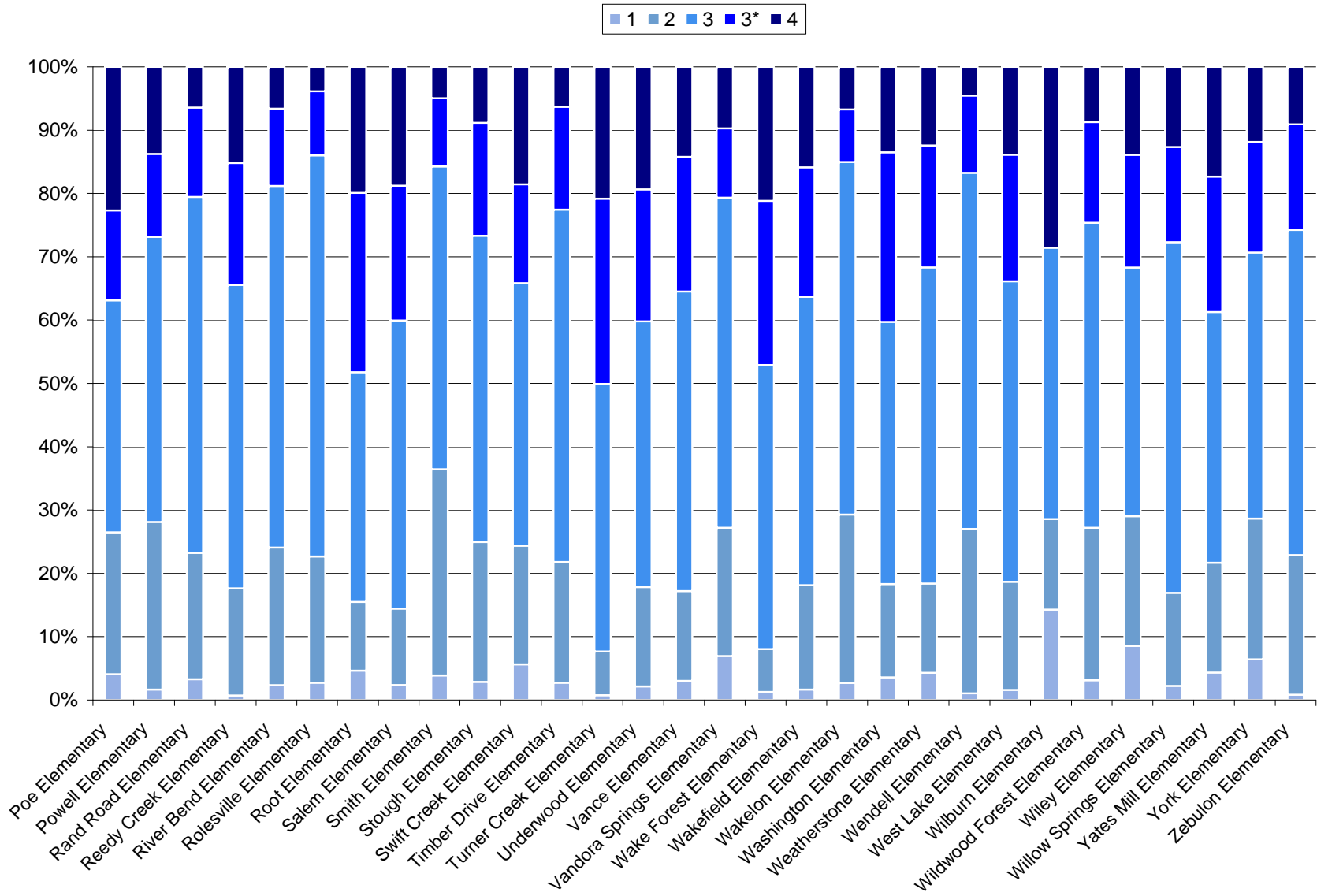
Percentage of Students by School and Grade Performance Level in Mathematics 2006-07



Percentage of Students by School and Grade Performance Level in Mathematics 2006-07



Percentage of Students by School and Grade Performance Level in Mathematics 2006-07



APPENDIX C

Summary of Logistic Regression Analysis for Variables Predicting Reading and Mathematics EOG Performance 2006-07

Dependent Variable = Reading EOG Level

Variable	B	SE B	Wald	Sig.	Odds Ratio	Direction of Odds (increase/decrease)
Grade 3 (N=8938)						
Reading EOG Pretest Level	1.22	0.06	441.98	<.0001	3.37	↑
Reading 2 nd Quarter Grade	1.46	0.10	208.02	<.0001	4.34	↑
Free or Reduced-Price Lunch	-.069	0.09	64.54	<.0001	0.50	↓
Limited English Proficiency	-0.24	0.12	3.99	.0459	0.79	↓
Students with Disabilities	-0.80	0.09	73.09	<.0001	0.45	↓
Grade 4 (N=8049)						
Reading EOG Level	1.60	0.08	393.07	<.0001	4.96	↑
Reading 2 nd Quarter Grade	1.41	0.13	112.17	<.0001	4.09	↑
Free or Reduced-Price Lunch	-0.37	0.11	11.57	.0007	0.69	↓
Students with Disabilities	-0.53	0.11	21.43	<.0001	0.59	↓
Grade 5 (N=7690)						
Reading EOG Level	1.70	0.09	330.17	<.0001	5.49	↑
Reading 2 nd Quarter Grade	1.16	0.16	53.61	<.0001	3.19	↑
Free or Reduced-Price Lunch	-0.47	0.14	11.02	.0009	0.62	↓
Limited English Proficiency	-0.71	0.18	15.03	.0001	0.49	↓
Students with Disabilities	-0.92	0.14	42.32	<.0001	0.40	↓

Dependent Variable = Mathematics EOG Level

Variable	B	SE B	Wald	Sig.	Odds Ratio	Direction of Odds (increase/decrease)
Grade 3 (N=8935)						
Math EOG Pretest Level	1.73	0.06	855.92	<.0001	5.65	↑
Math 2 nd Quarter Grade	1.35	0.08	309.25	<.0001	3.84	↑
Free or Reduced-Price Lunch	-0.67	0.07	85.26	<.0001	0.51	↓
Students with Disabilities	-0.40	0.09	19.34	<.0001	0.67	↓
Grade 4 (N=8056)						
Math EOG Level	2.02	0.07	914.72	<.0001	7.57	↑
Math 2 nd Quarter Grade	1.12	0.08	182.01	<.0001	3.06	↑
Free or Reduced-Price Lunch	-0.60	0.08	56.50	<.0001	0.55	↓
Students with Disabilities	-0.74	0.10	58.19	<.0001	0.48	↓
Grade 5 (N=7729)						
Math EOG Level	1.98	0.07	915.88	<.0001	7.25	↑
Math 2 nd Quarter Grade	1.21	0.08	219.12	<.0001	3.36	↑
Free or Reduced-Price Lunch	-0.67	0.08	70.57	<.0001	0.51	↓
Students with Disabilities	-0.47	0.10	22.76	<.0001	0.63	↓

Data Source: 2006-07 grade file and 2006-07 end-of-year student roster.

Note 1: Among grade 4 students in reading and 3-5 in mathematics, LEP did not significantly contribute to the model and were thus dropped.

Note 2: Odds Ratio > 1 = increased odds of EOG being on grade level; odds ratio = 1 means odds were unchanged; and odds ratio < 1 = decreased odds of EOG being on grade level.

Interpretation Example: Of the 8,938 grade 3 students, those who were on grade level in 2nd quarter reading had increased their odds (odds ratio 4.34 > 1) of being on grade level on the EOG.