



Executive Summary

The intent of the No Child Left Behind (NCLB) Act of 2001 is to hold schools accountable for ensuring that all their students achieve mastery in reading and math, with a particular focus on groups that have traditionally been left behind. Under NCLB, states submit accountability plans to the U.S. Department of Education detailing the rules and policies to be used in tracking the adequate yearly progress (AYP) of schools toward these goals.

This report examines South Carolina's NCLB accountability system—particularly how its various rules, criteria and practices result in schools either making AYP—or not making AYP. It also gauges how tough South Carolina's system is compared with other states. For this study we selected 36 schools from around the nation, schools that vary by size, achievement, and diversity, among other factors, and determined whether each would make AYP under South Carolina's system as well as under the systems of 27 other states. We used school data and proficiency cut score¹ estimates from academic year 2005–2006, but applied them against South Carolina's AYP rules for academic year 2007–2008 (shortened to “2008” in this report).

Here are some key findings:

- We estimate that **15 of 18 elementary schools and all 18 middle schools** in our sample failed to make AYP in 2008 under South Carolina's accountability system.
- This high failure rate is partly explained by our sample, which intentionally includes some schools with a relatively large population of low-performing students. But it's also partly explained by South Carolina's ambitious proficiency standards (or cut scores), which are among the most rigorous in our state sample.

¹ A cut score is the minimum score a student must receive on NWEA's Measures of Academic Progress (MAP) that is equivalent to performing proficient on the South Carolina Palmetto Achievement Challenge Tests (PACT).

- Looking across the 28 state accountability systems examined in the study, we find that the number of elementary schools making AYP in South Carolina is exceeded in 15 other sample states (South Carolina ties four other states with only three elementary schools making AYP). In addition, South Carolina is one of five states with no passing middle schools in the sample (see Figure 1).
- Unlike many other states in the study, most schools in South Carolina fail to meet math and reading targets for their overall populations. Again, this is likely due to the state's difficult proficiency standards.
- In South Carolina, as in most states, schools with fewer subgroups attain AYP more easily than schools

South Carolina's accountability system has several unique characteristics that land it in the middle of the state distribution in terms of the number of schools making AYP. First, South Carolina's cut scores are relatively difficult to achieve compared to other states in the study. Most all of them are above the 50th percentile and some are even around the 70th percentile. However, South Carolina adds the equivalent of one standard error to individual test scores, which essentially lowers the difficulty of the proficiency cut score (other states apply confidence intervals which have a similar effect). South Carolina also utilizes different subgroup sizes for different groups. The minimum subgroup size is 40 for racial, ethnic, and low-income groups (which is fairly standard), and 50 for students with disabilities (SWDs) and for students with limited English proficiency (LEP) (which is higher than in most other states). The latter means that fewer students are held separately accountable for performance in South Carolina that would similar schools in other states.

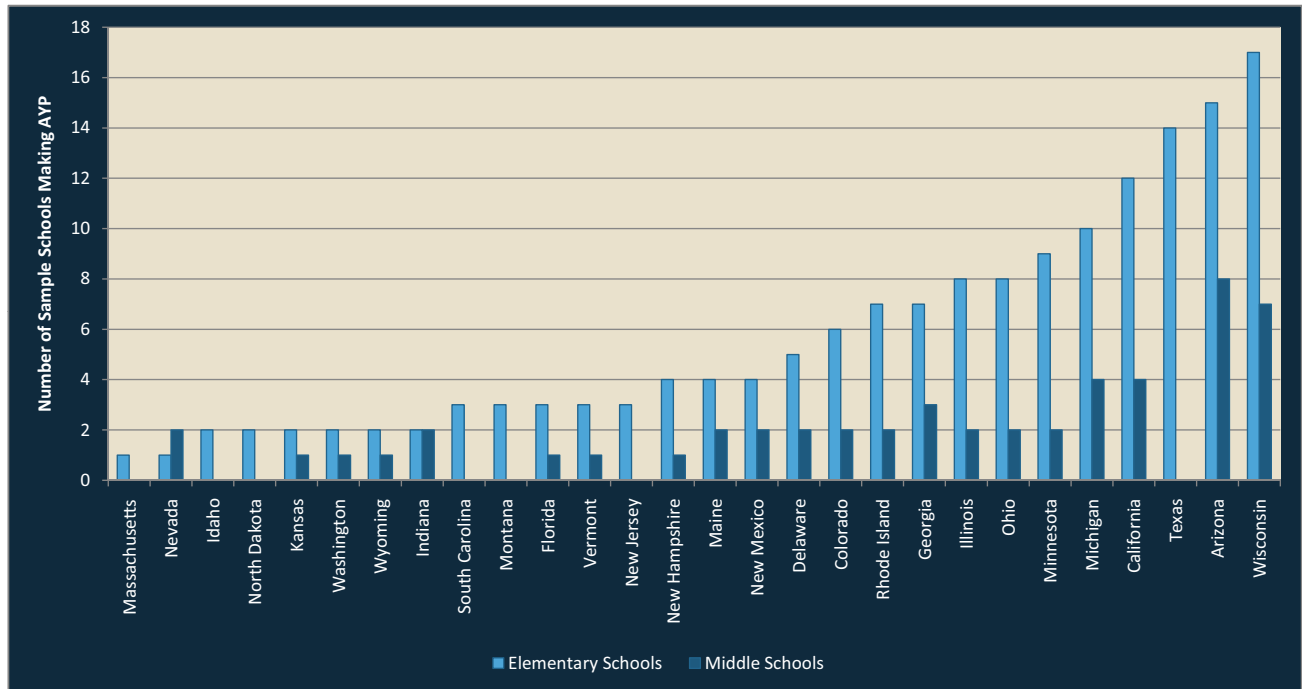


Figure 1. Number of sample schools making AYP by state

Note: Middle schools were not included for Texas and New Jersey; absence of a middle school bar in those states means “not applicable” as opposed to zero. States like Idaho and North Dakota, however, have zero passing middle schools.

with more subgroups, even when their average student performance is much lower. In other words, schools with greater diversity and size face greater challenges in making AYP.

- As is the case in other states, middle schools in South Carolina have greater difficulty reaching AYP than do elementary schools, primarily because their student populations are larger and therefore, have more qualifying subgroups—not because their student achievement is lower than in the elementary schools.
- A strong predictor of a school making AYP under South Carolina’s system is whether it has enough qualifying students with disabilities (SWDs) and students with limited English proficiency (LEP).

South Carolina has a large “minimum *n* size” (50) for these two groups, meaning schools in the state may have fewer accountable subgroups than would similar schools in other states.² Still, when enough students exist to comprise these subgroups, every single school with a SWD or LEP subgroup failed to make AYP. Likewise, all but one school with enough qualifying low-income students failed to meet its AYP targets.³

Introduction

The Proficiency Illusion (Cronin et al.2007a) linked student performance on South Carolina’s tests and those of 25 other state tests to the Northwest Evaluation Association’s (NWEA’s) Measures of Academic Progress (MAP),

² It’s important to note that students in subgroups not meeting the minimum *n* sizes are still included for accountability purposes in the overall student calculations; they simply are not treated as their own subgroup.

³ SWDs are defined as those students following individualized education plans. We should also note that our subgroup findings for LEP students and SWDs may be more negative than actual findings, mostly because of the likely differences between how LEP students and SWDs are treated in the Measures of Academic Progress (MAP), the assessment we used in this study, and in the South Carolina Palmetto Achievement Challenge Tests (PACT), the standardized state test. Specifically, the U.S. Department of Education has issued new NCLB guidelines in recent years that exclude small percentages of LEP students and SWDs from taking the state test or that allow them to take alternative assessments. In this study, however, no valid MAP scores were omitted from consideration.

a computerized adaptive test used in schools nationwide. This single common scale permitted cross-state comparisons of each state's reading and math proficiency standards to measure school performance under the No Child Left Behind (NCLB) Act of 2001. That study revealed profound differences in states' proficiency standards (i.e., how difficult it is to achieve proficiency on the state test), and even across grades within a single state.

Our study expands on *The Proficiency Illusion* by examining other key factors of state NCLB accountability plans and how they interact with state proficiency standards to determine whether schools make AYP. Specifically, we estimate how a single set of schools, drawn from around the country, would fare under the differing rules for determining Adequate Yearly Progress (AYP) in 28 states (the original 25 in *The Proficiency Illusion* plus 3 others for which we now have cut score estimates). In other words, if we could somehow move these entire schools—with their same mix of characteristics—from state to state, how would they fare in terms of making AYP? Will schools with high-performing students consistently make AYP? Will schools with low-performing students consistently fail to make AYP? If AYP determinations for schools are not consistent across states, what leads to the inconsistencies?

NCLB requires every state, as a condition of receiving Title I funding, to implement an accountability system that aims to get 100% of its students to the proficient level on the state test by academic year 2013–14. In the intervening years, states set annual measurable objectives (AMOs). This is the percentage of students in each school, and in each subgroup within the school (such as low-income,⁴ African-American, among others), that must reach the proficient level in order for the school to make AYP in a given year. These AMOs vary by state (as do, of course, the difficulty of the proficiency standards).

States also determine the minimum number of students that must constitute a subgroup in order for its scores to be

analyzed separately (also called the minimum n [number of students in sample] size). The rationale is that reporting the results of very small subgroups—fewer than ten pupils, for example—could jeopardize students' confidentiality and risk presenting inaccurate results. (With such small groups, random events, like one student being out sick on test day, could skew the outcome.) Because of this flexibility, states have set widely varying n sizes for their subgroups, from as few as 10 youngsters to as many as 100.

Many states have also adopted confidence intervals—basically margins of statistical error—to try to account for potential measurement error within the state test. In some states, these margins are quite wide, which has the effect of making it easier to achieve an annual target.

All of these AYP rules vary by state, which means that a school making AYP in Wisconsin or Ohio, for example, might not make it under Nevada's or Idaho's rules (U.S. Department of Education, 2008).

What We Studied

We collected students' MAP test scores from the 2005–2006 academic year from 18 elementary and 18 middle schools around the country. We also collected the NCLB subgroup designations for all students in those schools—in other words, whether they had been classified as members of a minority group, such as English language learners,⁵ among other subgroups.

The schools were not selected as a representative sample of the nation's population. Instead, we selected the schools because they exhibited a range of characteristics on measures such as academic performance, academic growth, and socioeconomic status (the latter calculated by the percentage of students receiving free or reduced-price lunches). Appendix 1 contains a complete discussion of the methodology for this project along with the characteristics of the school sample.⁶

⁴ Low-income students are those who receive a free or reduced-price lunch.

⁵ Note that we use “students with limited English proficiency (LEP)” or “LEP students” and “English language learners” interchangeably to refer to students in the same subgroup.

⁶ We gave all schools in our sample pseudonyms in this report.

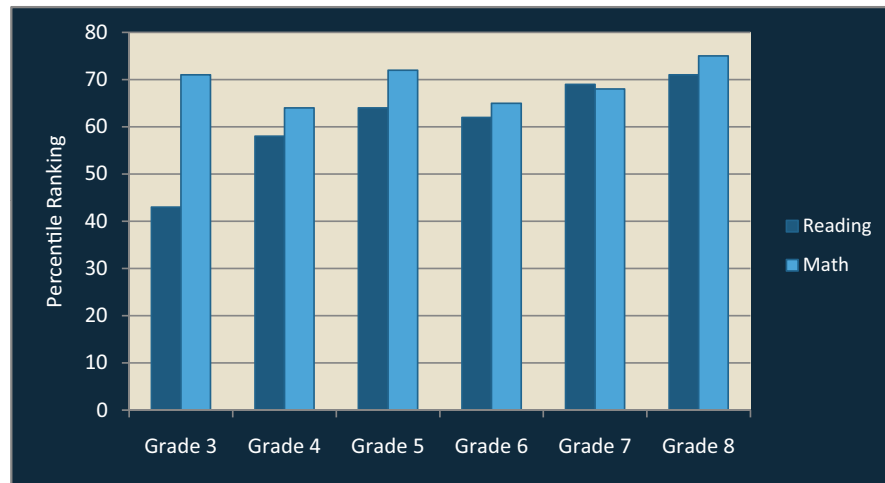


Figure 2. South Carolina reading and math cut score estimates, expressed as percentile ranks (2006)

Note: This figure illustrates the difficulty of South Carolina's cut scores (or proficiency passing scores) for its reading and mathematics tests, as percentiles of the NWEA norm, in grades three through eight. Higher percentile ranks are more difficult to achieve. All of South Carolina's cut scores are at or below the 75th percentile.

Proficiency cut score estimates for the South Carolina Palmetto Achievement Challenge Tests (PACT) are taken from *The Proficiency Illusion* (as shown in Figure 2), which found that South Carolina's definitions of proficiency generally ranked well above the average compared with the standards set by the other 25 states in that study. These cut scores were used to estimate whether students would have scored as proficient or better on the South Carolina test, given their performance on MAP. Student test data and subgroup designations were then used to determine how these 18 elementary and 18 middle schools would have fared under South Carolina AYP rules for 2008. (In other words, the school data and our proficiency cut score estimates are from academic year 2005–2006, but we are applying them against South Carolina's 2008 AYP rules.)

Table 1 shows the pertinent South Carolina AYP rules that were applied to elementary and middle schools in this study. South Carolina's minimum subgroup size is 40 for race/ethnicity and low-income groups, and 50 for students with disabilities and for students with limited English proficiency. While 40 is roughly comparable to

the number used by most other states in the current study, 50 is a bit larger.⁷ This means that schools in South Carolina may have fewer accountable subgroups than would similar schools in other states. Furthermore, while the majority of states examined in the study apply confidence intervals to their measurements of student proficiency rates, South Carolina adds the equivalent of one standard error to individual test scores, essentially lowering the difficulty of the proficiency cut score.⁸

Note that we were unable to examine the impact of NCLB's "safe harbor" provision. This provision permits a school to make AYP even if some of its subgroups fail, as long as it reduces the number of nonproficient students within any failing subgroup by at least 10% relative to the previous year's performance. Because we had access to only a single academic year's data (2005–2006), we were not able to include this in our analysis. As a result, it is possible that some of the schools in our sample that failed to make AYP according to our estimates would have made AYP under real conditions.

Furthermore, attendance and test participation rates are

⁷ Keep in mind, however, that school size and n size are related (e.g., a small n size make sense for small schools).

⁸ By adding a standard error to individual student scores, South Carolina essentially adds a few points to the score and effectively lowers the proficiency standard (or cut score). This is done to correct for potential measurement error of the state testing instrument, which is the same argument used by other states that use confidence intervals when reporting school or group proficiency rates. For the stated purpose (i.e., correcting for measurement error), South Carolina's approach is probably more appropriate. However, the current study did not systematically examine whether standard errors or confidence intervals provided greater assistance to schools.

Table 1. South Carolina AYP rules for 2008

Subgroup minimum <i>n</i>	Race/ethnicity: 40	
	SWDs: 50	
	Low-income students: 40	
	LEP students: 50	
CI	Applied to proficiency rate calculations?	
	Not used, but one standard error added to individual test scores	
AMOs	Baseline proficiency levels as of 2002 (%)	2008 targets (%)
READING/LANGUAGE ARTS		
Grade 3	17.6	58.8
Grade 4	17.6	58.8
Grade 5	17.6	58.8
Grade 6	17.6	58.8
Grade 7	17.6	58.8
Grade 8	17.6	58.8
MATH		
Grade 3	15.5	57.8
Grade 4	15.5	57.8
Grade 5	15.5	57.8
Grade 6	15.5	57.8
Grade 7	15.5	57.8
Grade 8	15.5	57.8

Sources: U.S. Department of Education (2008); Council of Chief State School Officers (2008).

Abbreviations: SWDs = students with disabilities; LEP = limited English proficiency; CI = confidence interval; AMOs = annual measurable objectives

beyond the scope of the study. Note that most states include attendance rates as an additional indicator in their NCLB accountability system for elementary and middle schools. In addition, federal law requires 95% of each school’s students—and 95% of the students in each subgroup—to participate in testing.

To reiterate, then, AYP decisions in the current study are modeled solely on test performance data for a single academic year. For each school, we calculated reading and math proficiency rates (along with any confidence intervals) to determine whether the overall school population and any qualifying subgroups achieved the AMOs. We deemed that a school made AYP if its overall student body and all its qualifying subgroups met or exceeded

its AMOs. Again, Appendix 1 supplies further methodological detail.

How Did the Sample Schools Fare Under South Carolina’s AYP Rules?

Figure 3 illustrates the AYP performance of the sample elementary schools under South Carolina’s 2008 AYP rules. **Only 3 elementary schools made AYP while 15 failed to make it.** The triangles in Figure 3 show the average academic performance of students within the school, with negative values indicating below-grade-level performance for the average student, and positive values indicating above-grade-level performance. All schools making AYP are in the right half of the figure, meaning

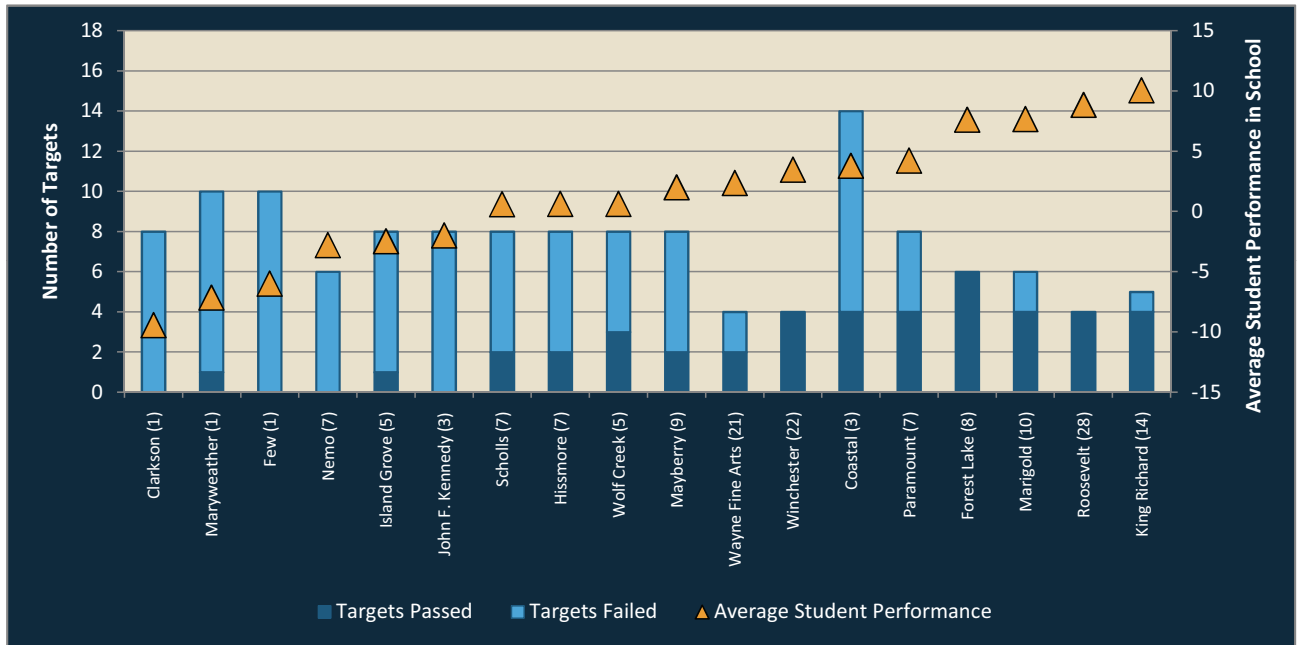


Figure 3. AYP performance of the elementary school sample under the South Carolina 2008 AYP rules

Note: This figure indicates how each elementary school within the sample fared under South Carolina’s AYP rules (as described in Table 1). The bars show the number of targets that each school has to meet to make AYP under the state’s NCLB rules, and whether they met them (dark blue) or did not (light blue). The more subgroups in a school, the more targets it must meet. Under the study conditions, a school that failed to meet the AMO for even a single subgroup did not make AYP, so any light blue means the school failed. Marigold Elementary, for example, met 4 of its 6 targets, but because it did not meet them all, it did not make AYP. Schools are ordered from lowest to highest average student performance (shown by the orange triangles) which is measured by the average MAP performance of students within the school; its scale is shown on the right side of the figure. Scores below zero (which is the grade level median) denote below-grade-level performance and scores above zero denote above-grade-level performance. One unit does not equal a grade level; however, the higher the number, the better the average performance and the lower the number, the worse the average performance. The number in parentheses after each school name indicates the number of states, out of 28, in which that school would have made AYP in the study.

that the highest performing students were found at these schools.

Yet almost without regard to average student performance, the only schools to make AYP are those with relatively few qualifying subgroups—and thus the fewest targets to meet (because each subgroup has separate targets). For example, Winchester made it, but it had only four targets—two in reading and math for its overall population, and two more for the only subgroup to exceed the minimum size (the white subgroup).

Figure 4 illustrates the AYP performance of the sample middle schools under the 2008 South Carolina AYP rules. **None of the sample middle schools made AYP.**

Where do schools fail?

Figures 3 and 4 indicate the number of subgroup targets evaluated in each sample school, and each school’s final

AYP outcome. However, these figures do not indicate which subgroups failed or passed in which school. Information on individual subgroup performance appears in Tables 2 and 3 for elementary and middle schools, respectively.

Tables 2 and 3 show which subgroups qualified for evaluation at each school (i.e., whether the number of students within that subgroup exceeded the state’s minimum *n*), and whether that subgroup passed or failed. Although all schools are evaluated on the proficiency rate of their overall population, potential subgroups that are separately evaluated for AYP include SWDs, students with LEP, low-income students, and the following race/ethnic categories: African American, Asian/Pacific Islander, Hispanic/Latino, American Indian/Alaska Native, and White. Tables 2 and 3 also show whether a school met AYP under the 2008 South Carolina rules, and the total number of states within the study in which that school met AYP.

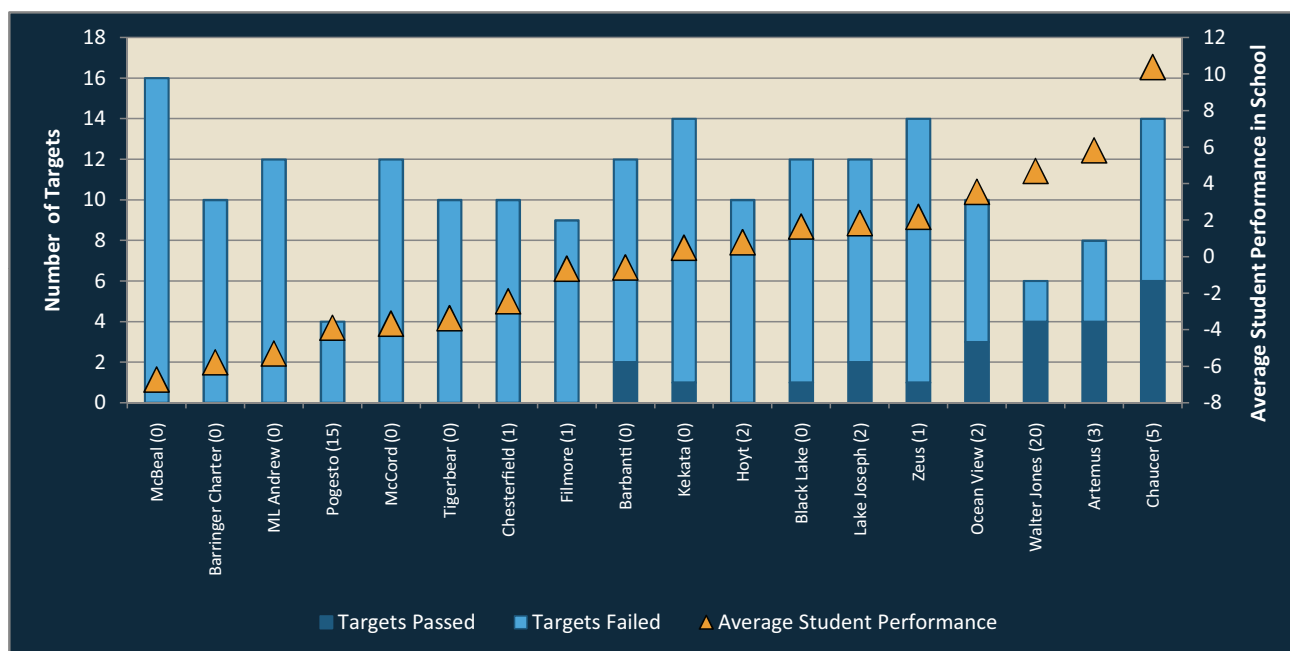


Figure 4. AYP performance of the middle school sample under the South Carolina 2008 AYP rules

Note: This figure indicates how each of the middle schools within the sample fared under South Carolina’s AYP rules (as described in Table 1). The bars show the number of targets that each school had to meet in order to make AYP under the state’s NCLB rules, and whether they met them (dark blue) or did not (light blue). The more subgroups in a school, the more targets it must meet. Under the study conditions, a school that failed to meet the AMO for even a single subgroup does not make AYP, so any light blue means the school failed. Walter Jones, for example, met 4 of its 6 targets, but because it did not meet them all, it did not make AYP. Schools are ordered from lowest to highest average student performance (shown by the orange triangles). This is measured by the average MAP performance of students within the school; its scale is shown on the right side of the figure. Scores below zero (which is the grade level median) denote below-grade-level performance and scores above zero denote above-grade-level performance. One unit does not equal a grade level; however, the higher the number, the better the average performance and the lower the number, the worse the average performance. The number in parentheses after each school name indicates the number of states, out of 28, in which that school would make AYP in the study.

The school-by-school findings in Tables 2 and 3 show that:

- In South Carolina, most schools failed to meet math and reading targets for their overall populations, unlike many other states in the study.
- Seven elementary schools meet both the reading and math targets for their overall populations.
- Three middle schools met both the reading and math targets for their overall populations.
- Overall, in most of the cases where there were enough students to comprise a subgroup, these groups didn’t meet their targets.

Tables 4 and 5 summarize subgroup performance for elementary and middle schools, respectively. First, one can see that nearly all subgroups at both the elementary and middle school levels struggle with South Carolina’s reading and math requirements, perhaps because South Carolina’s proficiency cut scores are very difficult compared

to the other states in the sample. Every school with sufficient numbers of SWDs, LEP students, low income students, African American, and Hispanic subgroups failed to make AYP.

Characteristics of Schools that Did and Didn’t Make AYP

A close look at Figures 3 and 4 indicates that South Carolina’s NCLB accountability system is, in some respects, behaving like those in other states. For example, among the elementary schools in our sample, Roosevelt and Winchester made AYP in the greatest number of states—28 and 22 respectively. And these schools made AYP in South Carolina, too. Likewise, the elementary and middle schools that fail to make AYP in the greatest number of states also failed to make AYP in South Carolina.

But South Carolina is home to at least one anomaly. First, consider Forest Lake Elementary (see Figure 3). It failed to make AYP in 20 of the 28 states in our sample,

Table 2. Elementary school subgroup performance of sample schools under the 2008 South Carolina AYP rules

SCHOOL PSEUDONYM	Overall Proficiency Rate		Overall		SWDs		LEP Students		Low-income Students		AA		Asian		Hispanic		AI/AN		White		AYP Targets Required	Targets MET	% of Targets Met	School Met AYP?	Number of states in which school met AYP?	
	Math	Reading	M	R	M	R	M	R	M	R	M	R	M	R	M	R	M	R	M	R						
Clarkson	27.0%	23.1%	N	N			N	N	N	N					N	N					8	0	0%	N	1	
Maryweather	30.6%	38.4%	N	N			N	N	N	N					N	N				N	Y	10	1	10%	N	1
Few	36.1%	37.8%	N	N	N	N	N	N	N	N					N	N					10	0	0%	N	1	
Nemo	35.3%	53.5%	N	N					N	N									N	N	6	0	0%	N	7	
Island Grove	39.9%	56.8%	N	N					N	N					N	N				N	Y	8	1	13%	N	4
JFK	43.1%	48.5%	N	N					N	N	N	N							N	N	8	0	0%	N	3	
Scholls	54.1%	56.1%	N	N					N	N	N	N							Y	Y	8	2	25%	N	7	
Hissmore	53.2%	58.4%	N	N					N	N	N	N							Y	Y	8	2	25%	N	7	
Wolf Creek	55.0%	58.9%	N	Y					N	N					N	N			Y	Y	8	3	38%	N	5	
Alice Mayberry	54.1%	57.4%	N	N					N	N	N	N							Y	Y	8	2	25%	N	9	
Wayne Fine Arts	48.3%	67.2%	N	Y															N	Y	4	2	50%	N	21	
Winchester	58.0%	68.7%	Y	Y															Y	Y	4	4	100%	Y	22	
Coastal	64.8%	61.5%	Y	Y	N	N	N	N	N	N	N	N			N	N			Y	Y	14	4	29%	N	3	
Paramount	68.8%	67.9%	Y	Y					N	N					N	N			Y	Y	8	4	50%	N	7	
Forest Lake	76.7%	76.6%	Y	Y					Y	Y									Y	Y	6	6	100%	Y	8	
Marigold	75.9%	75.1%	Y	Y					N	N									Y	Y	6	4	67%	N	10	
Roosevelt	74.7%	83.4%	Y	Y															Y	Y	4	4	100%	Y	28	
King Richard	76.0%	81.3%	Y	Y					N										Y	Y	5	4	80%	N	14	

Abbreviations: M = math; R = reading; N = no; Y = yes; SWDs = students with disabilities; AA = African American; Asian/Pacific Islander = Asian; Hispanic/Latino = Hispanic; American Indian/Alaska Native = AI/AN.

Note: Schools are ordered from lowest (Clarkson) to highest (King Richard) average student performance as measured by combined and weighted math and reading performance on the MAP assessment (not shown in table). A blank space underneath a subgroup means that subgroup contained fewer than the minimum number of students required for evaluation, so it wasn't counted. A "Y" in blue means that the group met the AMOs and an "N" in peach means that the group did not meet the AMOs. The two rightmost columns show (1) whether that school met AYP (i.e., it met the targets for its overall population and all required subgroups); and (2) the total number of states in the study for which that school met AYP.

yet made AYP in South Carolina. Examining Table 2, one can see that Forest Lake didn't meet the minimum numbers for the LEP or SWD subgroups, perhaps because South Carolina's minimum "n" for these groups is higher than in most of the other states examined. With fewer accountable subgroups, Forest Lake made AYP, even when other schools with higher average performance failed.

That fewer accountable subgroups is a good predictor of making AYP is consistent with the patterns shown in Table 6, which compares elementary schools (there were

no passing middle schools) that made and didn't make AYP on a number of academic and demographic dimensions. Within the sample, schools that made AYP do indeed show higher average student performance, but they also differ in the following ways: they have smaller student populations, fewer subgroups (and thus fewer targets to meet), and lower percentages of low income students.

Concluding Observations

This study examined the test performance data of students from 18 elementary and 18 middle schools across the

Table 3. Middle school subgroup performance of sample schools under the 2008 South Carolina AYP rules

SCHOOL PSEUDONYM	Overall Proficiency Rate		Overall		SWDs		LEP Students		Low-income Students		AA		Asian		Hispanic		AI/AN		White		AYP Targets Required	Targets MET	% of Targets Met	School Met AYP?	Number of states in which school met AYP?
	Math	Reading	M	R	M	R	M	R	M	R	M	R	M	R	M	R	M	R	M	R					
McBeal	28.9%	33.7%	N	N	N	N	N	N	N	N	N	N			N	N	N	N	N	N	16	0	0%	N	0
Barringer Charter	28.5%	37.2%	N	N	N	N			N	N	N	N			N	N					10	0	0%	N	0
ML Andrew	26.4%	34.3%	N	N	N	N			N	N	N	N			N	N			N	N	12	0	0%	N	0
Pogesto	29.6%	20.4%	N	N															N	N	4	0	0%	N	15
McCord Charter	32.3%	40.1%	N	N	N	N			N	N	N	N			N	N			N	N	12	0	0%	N	0
Tigerbear	38.9%	34.9%	N	N	N	N			N	N	N	N							N	N	10	0	0%	N	0
Chesterfield	40.0%	33.5%	N	N	N	N			N	N	N	N							N	N	10	0	0%	N	1
Filmore	40.4%	42.9%	N	N		N			N	N					N	N			N	N	9	0	0%	N	1
Barbanti	42.4%	43.3%	N	N	N	N	N	N	N	N					N	N			Y	Y	12	2	17%	N	0
Kekata	49.7%	43.9%	N	N	N	N	N	N	N	N	N	N			N	N			Y	N	14	1	7%	N	0
Hoyt	45.1%	46.6%	N	N	N	N			N	N	N	N							N	N	10	0	0%	N	2
Black Lake	52.5%	43.2%	N	N	N	N			N	N	N	N			N	N			Y	N	12	1	8%	N	0
Lake Joseph	49.1%	49.4%	N	N	N	N	N	N	N	N					N	N			Y	Y	12	2	17%	N	2
Zeus	53.7%	50.2%	N	N	N	N	N	N	N	N	N	N			N	N			Y	N	14	1	7%	N	1
Ocean View	52.8%	58.9%	N	Y			N	N	N	N					N	N			Y	Y	10	3	30%	N	2
Walter Jones	63.4%	70.3%	Y	Y					N	N									Y	Y	6	4	67%	N	20
Artemus	66.2%	63.2%	Y	Y					N	N					N	N			Y	Y	8	4	50%	N	3
Chaucer	68.4%	70.7%	Y	Y	N	N	N	N	N	N			Y	Y	N	N			Y	Y	14	6	43%	N	5

Abbreviations: M = math; R = reading; N = no; Y = yes; SWDs = students with disabilities; AA = African American; Asian/Pacific Islander = Asian; Hispanic/Latino = Hispanic; American Indian/Alaska Native = AI/AN.

Note: Schools are ordered from lowest (McBeal) to highest (Chaucer) average student performance as measured by combined and weighted math and reading performance on the MAP assessment (not shown in table). A blank space underneath a subgroup means that subgroup contained fewer than the minimum number of students required for evaluation, so it wasn't counted. A "Y" in blue means that the group met the AMOs and an "N" in peach means that the group did not meet the AMOs. The two rightmost columns show (1) whether that school met AYP (i.e., it met the targets for its overall population and all required subgroups); and (2) the total number of states in the study for which that school met AYP.

country to see how these schools would fare under South Carolina's AYP rules and annual measurable objectives for 2008. We found that, only three elementary schools and no middle schools—three in all from a total of 36—would have made AYP in South Carolina. Looking across the 28 state accountability systems examined in the study, this puts South Carolina at the low end of the distribution in terms of the number of elementary schools making AYP (see Figure 1). Part of the reason that so few schools make

AYP in South Carolina may be due to its ambitious proficiency standards, which are among the most rigorous in our state sample. Indeed, unlike many other states in the study, most schools in South Carolina fail to meet math and reading targets for their overall populations.⁹

The overriding goal of the No Child Left Behind act (NCLB) is to eliminate educational disparities within and across states, it's important to consider whether

⁹ It does not appear that South Carolina's high proficiency standards have had much impact on the state's performance on the latest (2007) National Assessment of Educational Progress (NAEP). South Carolinian children performed lower than the national average in grade 4 and 8 reading, as well as in grade 4 math.

Table 4. Summary of subgroup performance of sample elementary schools under the 2008 South Carolina AYP rules

SUBGROUP	Number of schools with qualifying subgroups	Number of schools where subgroup failed to meet math target	Number of schools where subgroup failed to meet reading target
Students with disabilities	2	2	2
Students with limited English proficiency	4	4	4
Low-income students	15	14	13
African-American students	5	5	5
Asian/Pacific Islander students	0	0	0
Hispanic students	7	7	7
American Indian/Alaska Native students	0	0	0
White students	16	5	2

Table 5. Summary of subgroup performance of sample middle schools under the 2008 South Carolina AYP rules

SUBGROUP	Number of schools with qualifying subgroups	Number of schools where subgroup failed to meet math target	Number of schools where subgroup failed to meet reading target
Students with disabilities	14	13	14
Students with limited English proficiency	7	7	7
Low-income students	17	17	17
African-American students	10	10	10
Asian/Pacific Islander students	1	0	0
Hispanic students	13	13	13
American Indian/Alaska Native students	1	1	1
White students	17	8	11

states' annual decisions about the progress of individual schools are consistent with this aim. In some respects, South Carolina's No Child Left Behind accountability system is working exactly as Congress intended: identifying as "needing attention" schools with relatively high test score averages that mask low performance for particular groups of students, such as low-income or Hispanic students. In the pre-NCLB era, such schools might have

been considered to be effective or at least not in need of improvement, even though sizable numbers of their pupils aren't meeting state standards. Disaggregating data by race, income, and so on has made those students visible. That is surely a good thing.

Yet NCLB's design flaws are also readily apparent. Does it make sense, in the case of South Carolina, that schools

Table 6. Comparisons between schools that did and didn't make AYP in South Carolina, 2008

	Elementary Schools		Middle Schools	
	Made AYP	Failed to make AYP	Made AYP	Failed to make AYP
Number of schools in sample	3	15	0	18
Average student body size	243	317	n/a	859
Average % low income	20	52	n/a	45
Average % nonwhite	21	45	n/a	44
Average performance†	6.65	0.14	n/a	-0.05
Average % growth‡	131	112	n/a	98
Average number of targets to meet	5	8	n/a	11

† Student performance is measured by NWEA's MAP assessment and is expressed as an index of grade level normative performance. Scores below zero (which is the grade level median) denote below-grade-level performance and scores above zero denote above-grade-level performance. One unit does not equal a grade level; however, the higher the number, the better the average performance and the lower the number, the worse the average performance.

‡ Average growth refers to improvement from fall to spring on the NWEA MAP assessments, averaged across all students within the school. Growth is expressed as an index value relative to NWEA norms and is scaled as a percentage. Thus, 100% means that students at the school are achieving normative levels of growth for their age and grade. Less than 100% growth means that the average student is increasing *by less* than normative amounts, while percentages over 100 mean that the average student is *exceeding* normative growth expectations.

n/a = not applicable

are penalized for the state's high proficiency standards? Does it make sense that fewer subgroups enhances the likelihood of making AYP? Even if actual participation guidelines for English language learners and SWDs are more generous under the current state assessment system,¹⁰ doesn't the failure of these students (especially at the middle school level where more satisfy eligibility requirements) to meet South Carolina's targets indicate

that a new approach is needed for holding schools accountable for the performance of these students? Yes, schools should redouble their efforts to boost achievement for LEP and SWD students, as for other students, but when so few schools are able to meet the goal perhaps that indicates the goal is unrealistic. These will be critical considerations for Congress as it takes up NCLB reauthorization in the future.

Limitations

Although the purpose of our study was to explore how various elements of accountability systems in different states jointly affect a school's AYP status, the study will not precisely replicate the AYP outcome for every single school for several reasons. Because we projected students' state test performance from their MAP scores, and because MAP assessments—unlike state tests—are not required of all students within a school, it's possible that sampling or measurement error (or both) affected school AYP outcomes within our model. Nevertheless, for all but two of the sampled schools, our projections matched NCLB-reported proficiency ratings (in each respective state) to within 5 percentage points.

¹⁰ See footnote 3.

An additional limitation of the study was that it was not possible to consider NCLB's safe harbor provisions, which might have allowed some schools to make AYP even though they failed to meet their state's required AMOs. A few schools would have also passed under the new growth-model pilots currently under way in a handful of states, such as Ohio and Arizona. Others identified as making AYP in our study might actually have failed to make it because they did not meet their state's average daily attendance requirement or because they did not test 95% of some subgroup within their overall student population. At the end of the day, then, it's important to keep in mind that the number of schools that did or did not make AYP in our study do not by themselves measure the effectiveness of the entire state accountability system, of which there are many parts.

Despite these limitations, we believe that the study illuminates the inconsistency of proficiency standards and some of the rules across states. It's also useful for illustrating the challenges that states face as the requirements for AYP continue to ratchet up. The national report contains additional discussion of the study methodology and its limitations.