

Executive Summary

The intent of the No Child Left Behind (NCLB) Act of 2001 is to hold schools accountable for ensuring that all of 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 Nevada'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 Nevada's system is compared with other states. For this study, we selected 36 schools from various states around the nation, schools that vary by size, achievement, and diversity, among other factors, and determined whether each would make AYP under Nevada'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 Nevada's AYP rules for academic year 2007–2008 (shortened to “2008” in this report).

Here are some key findings:

- We estimate that **17 of 18 elementary schools** and **16 of 18 middle schools** in our sample **failed to make adequate yearly progress** in 2008 under Nevada's accountability system. This high failure rate is partly explained by our sample, which intentionally includes some schools with relatively large populations of low-performing students. It's also partly because Nevada's minimum subgroup size is relatively small (25) compared to other states; this means more subgroups are held accountable for per-

formance. In fact, a few sample schools that made AYP in most other states did not make it in Nevada, largely owing to the state's *n* size. (This occurred despite Nevada's fairly low annual performance targets, which require barely half of students to be proficient in math and reading in 2008).

- **Looking across the 28 state accountability systems examined in the study, we find that the number of elementary schools that made AYP in Nevada was exceeded by virtually all of the other sample states (Massachusetts and Nevada tie with a single elementary school making AYP). Nevada is one of 10 states with 2 middle schools that made AYP in the sample** (see Figure 1).
- Many schools in our sample that failed to make AYP in Nevada met expected targets for their overall populations² but failed because of the performance of individual subgroups, particularly students with disabilities (SWDs) and English language learners.
- In Nevada, schools with fewer subgroups attained AYP more easily than schools with more subgroups, even when their average student performance is

A couple of key factors combine to place **Nevada** at the low end of the state distribution in terms of the number of schools making AYP. First, Nevada's definitions of proficiency generally ranked at or above average compared to the standards set by the other 27 states in the study. This means that students had to perform at a higher level in order to be deemed proficient in Nevada. Second, Nevada's minimum subgroup size is relatively small (25), meaning that more subgroups are held separately accountable in Nevada than would be in other states. In fact, every single school with a limited English proficient (LEP) or students-with-disabilities (SWD) subgroup failed to make AYP in Nevada.

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 Nevada Criterion Referenced Test.

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 are simply not treated as their own subgroup.

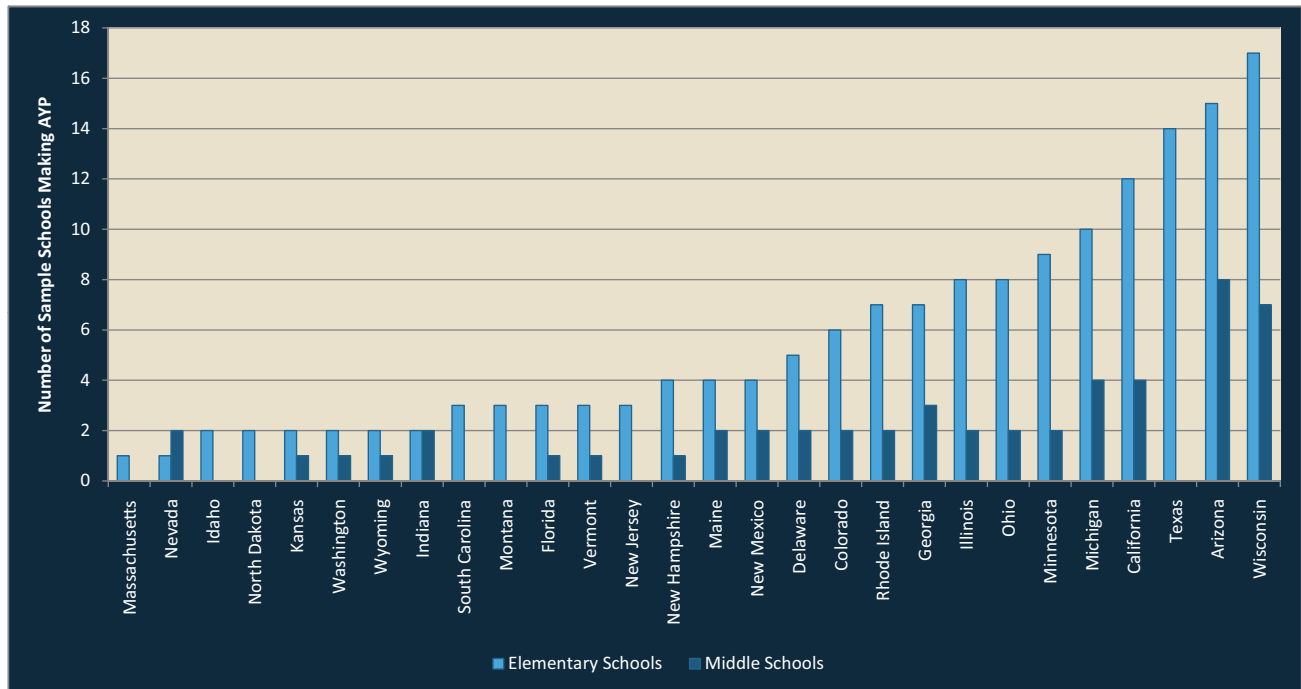


Figure 1. Number of sample schools making AYP by state

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.

much lower. In other words, schools with greater diversity and size face greater challenges in making AYP.

- Every school with a limited English proficient (LEP)³ subgroup failed to make AYP. Likewise, almost all schools with enough qualifying SWDs failed to meet their AYP targets.⁴

Introduction

The Proficiency Illusion (Cronin et al. 2007a) linked student performance on Nevada’s tests and those of 25 other states to the Northwest Evaluation Association’s (NWEA’s) Measures of Academic Progress (MAP), a computerized adaptive test used in schools nationwide. This single common scale permitted cross-state comparisons of each state’s reading and math proficiency stan-

dards 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 the schools in our sample made adequate yearly progress (AYP) in 2008. Specifically, we estimated how a single set of schools, drawn from around the country, would fare under the differing rules for determining 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

³ Note that we use “LEP students” and “English language learners” interchangeably to refer to students in the same subgroup.

⁴ 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 MAP, the assessment we used in this study, and in the Nevada Criterion Referenced Test, 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.

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 2013–2014. 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 income⁵ or African American, among others), that must reach the proficient level in order for the school to make AYP in a given year. The 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 10 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 that makes AYP in Wisconsin or Ohio, for example, might not make it under South Carolina’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 or as English language learners, 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

Proficiency cut score estimates for the Nevada Criterion Referenced Test (Nevada CRT) are taken from *The Proficiency Illusion* (as shown in Figure 2), which found that Nevada’s definitions of proficiency generally ranked about 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 Nevada 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 Nevada 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 Nevada’s 2008 AYP rules.

Table 1 shows the pertinent Nevada AYP rules that we applied to elementary and middle schools in this study. **Nevada’s minimum subgroup size is 25, which is small compared to most other states examined in the study, meaning that Nevada schools will have more accountable subgroups than would similar schools in other states.**⁷

⁵ Low-income students are those who receive a free or reduced-price lunch.

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

⁷ It’s also possible that Nevada’s schools are small and that an *n* size of 25 makes sense for that state.

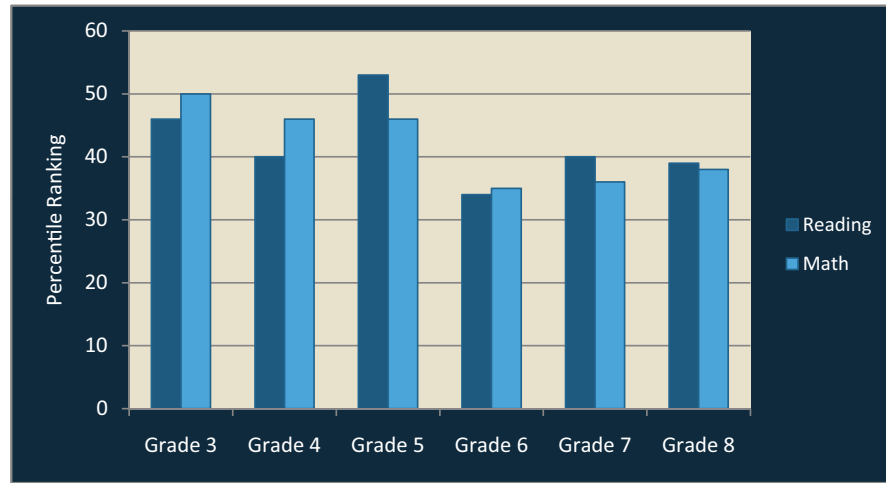


Figure 2. Nevada reading and math cut score estimates, expressed as percentile ranks (2006)

Note: This figure illustrates the difficulty of Nevada's cut scores (or proficiency passing scores) for its reading and math tests, as percentiles of the NWEA norm, in grades three through eight. Higher percentile ranks are more difficult to achieve. Almost all of Nevada's cut scores are at or below the 50th percentile.

Table 1. Nevada AYP rules for 2008

Subgroup minimum <i>n</i>	Race/ethnicity: 25	
	SWDs: 25	
	Low-income students: 25	
	LEP students: 25	
CI	Applied to proficiency rate calculations?	
	Yes; 95% CI used	
AMOs	Baseline proficiency levels as of 2002 (%)	2008 targets (%)
READING/LANGUAGE ARTS		
Grade 3	32.4	51.7
Grade 4	32.4	51.7
Grade 5	32.4	51.7
Grade 6	n/a	58.0
Grade 7	n/a	58.0
Grade 8	n/a	58.0
MATH		
Grade 3	37.3	56.3
Grade 4	37.3	56.3
Grade 5	37.3	56.3
Grade 6	n/a	54.6
Grade 7	n/a	54.6
Grade 8	n/a	54.6

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; n/a = not available

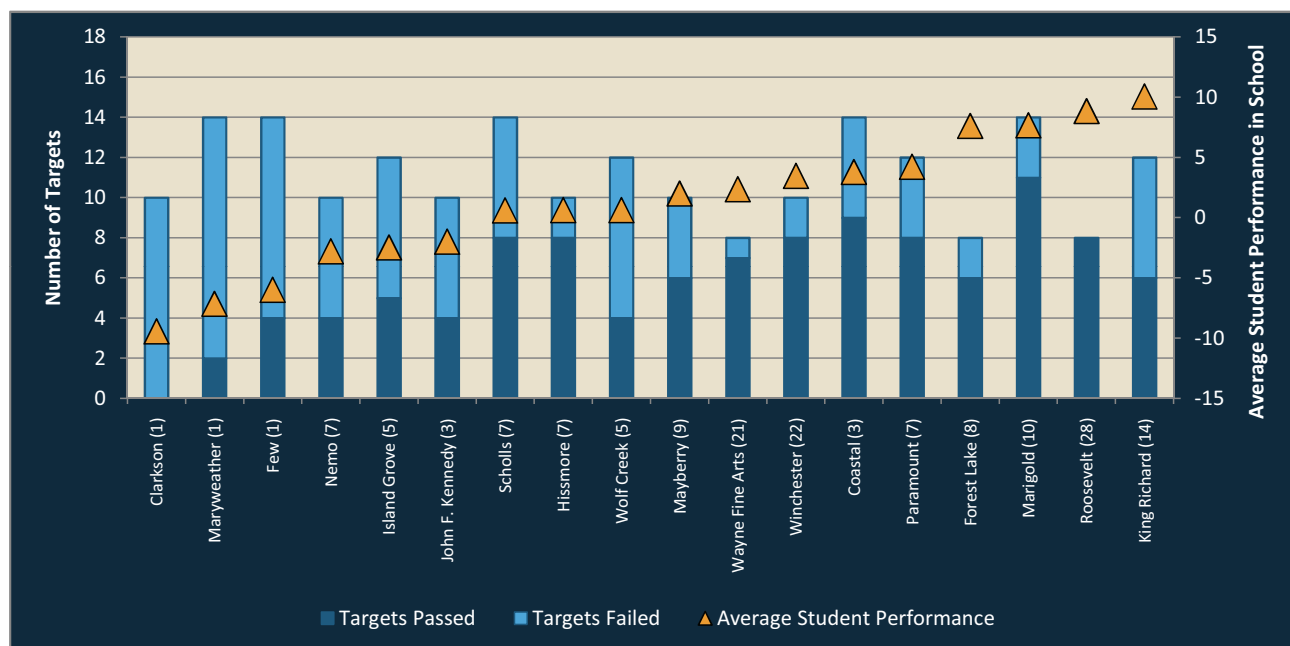


Figure 3. AYP performance of the elementary school sample under Nevada's 2008 AYP rules

Note: This figure indicates how each of the elementary schools within the sample fared under Nevada's AYP rules (as described in Table 1). The bars show the number of targets that each school has to meet in order to make AYP under the state's NCLB rules, and whether they met them (dark blue) or did not meet them (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 AMOs for even a single subgroup didn't make AYP, so any light blue means that the school failed. Wayne Fine Arts Elementary, for example, met seven of its eight targets, but because it didn't meet them all, it didn't 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, and 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.

Nevada, like the majority of states in the study, applies 95% confidence intervals to its measurements of student proficiency rates.⁸ So, for instance, even though schools are supposed to get 51.7% of their grade 3 students to the proficient level on the state reading test (and 51.7% of the grade 3 students in each subgroup), applying the confidence interval means that the real target can be lower, particularly with smaller groups.

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 re-

sult, it's 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 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

⁸ We also conducted an analysis to show the effect of confidence intervals on the reading and math proficiency rates for elementary and middle schools. We describe those results later in the report.

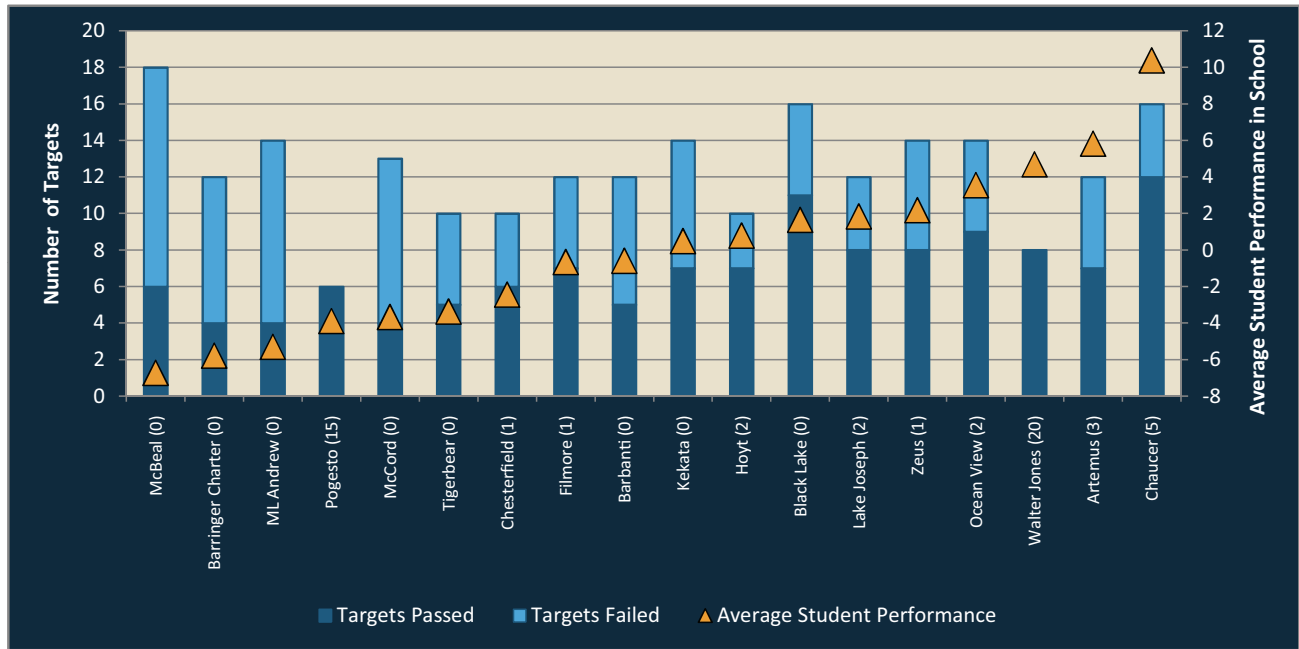


Figure 4. AYP performance of the middle school sample under Nevada's 2008 AYP rules

Note: This figure shows how each of the middle schools within the sample fared under Nevada'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 meet them (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 AMOs for even a single subgroup did not make AYP, so any light blue means that the school failed. Hoyt Middle School, for example, met 7 of its 10 targets, but because it didn't meet them all, it didn't 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, and 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.

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 Nevada's AYP Rules?

Figure 3 illustrates the AYP performance of the sample elementary schools under Nevada's 2008 AYP rules. **Only one elementary school made AYP and 17 failed.** 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. The only school making AYP (Roosevelt) is in the right half of the figure, meaning that relatively high performing students were found at that school. Roosevelt was also one of the only high performing schools with relatively few subgroups (and hence, targets to meet).

Figure 4 illustrates the AYP performance of the sample middle schools under the 2008 Nevada AYP rules. **Of 18 middle schools in our sample, only 2 passed**—one low-performance school (Pogesto) and one high-performance school (Walter Jones), both of which have relatively few qualifying subgroups.

Figures 5 and 6 indicate the degree to which schools' math proficiency rates are aided by Nevada's confidence interval for elementary and middle schools, respectively. On this figure, the darker portions of the bars show the actual proficiency rates at each school and the lighter portions of the bars show the degree to which these proficiency rates were increased by applying the confidence interval. The orange lines show the AMOs needed to meet AYP. These figures show that three sample elementary schools (Few, Nemo, and Island Grove) and two middle schools (ML Andrew and Pogesto) were assisted by the confidence interval. However, all of these schools but Pogesto already failed to make AYP because of low subgroup performance (see Figures 3 and 4), and therefore did not make AYP.

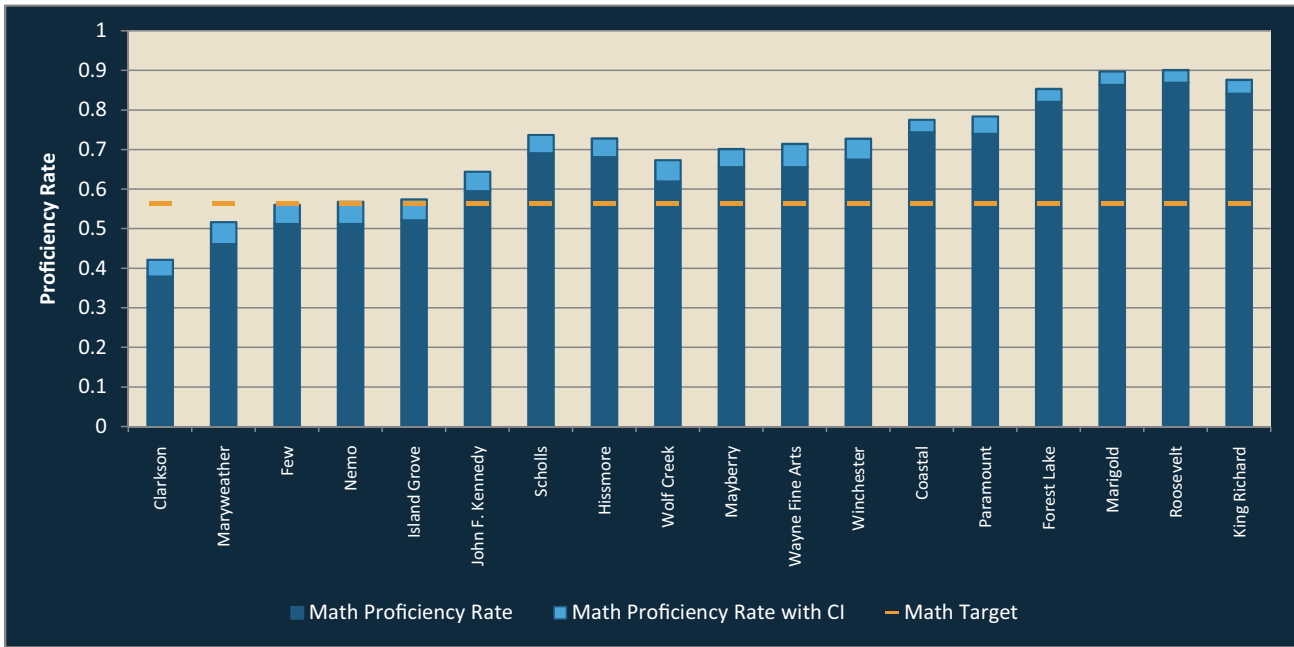


Figure 5. Impact of the confidence interval on elementary school math proficiency rates for 2008

Note: This figure shows the reported proficiency rate for the student population as a whole and the impact of the confidence interval on meeting annual targets. The darker portions of the bars show the actual proficiency rate achieved, while the lighter (upper) portions of the bars show the margin of error as computed by the confidence interval. The figure shows that three of the sample elementary schools (Few, Nemo, and Island Grove) were assisted by the confidence interval. Annual targets (the orange lines) are considered to be met by the confidence interval if they fall within the light blue portion.

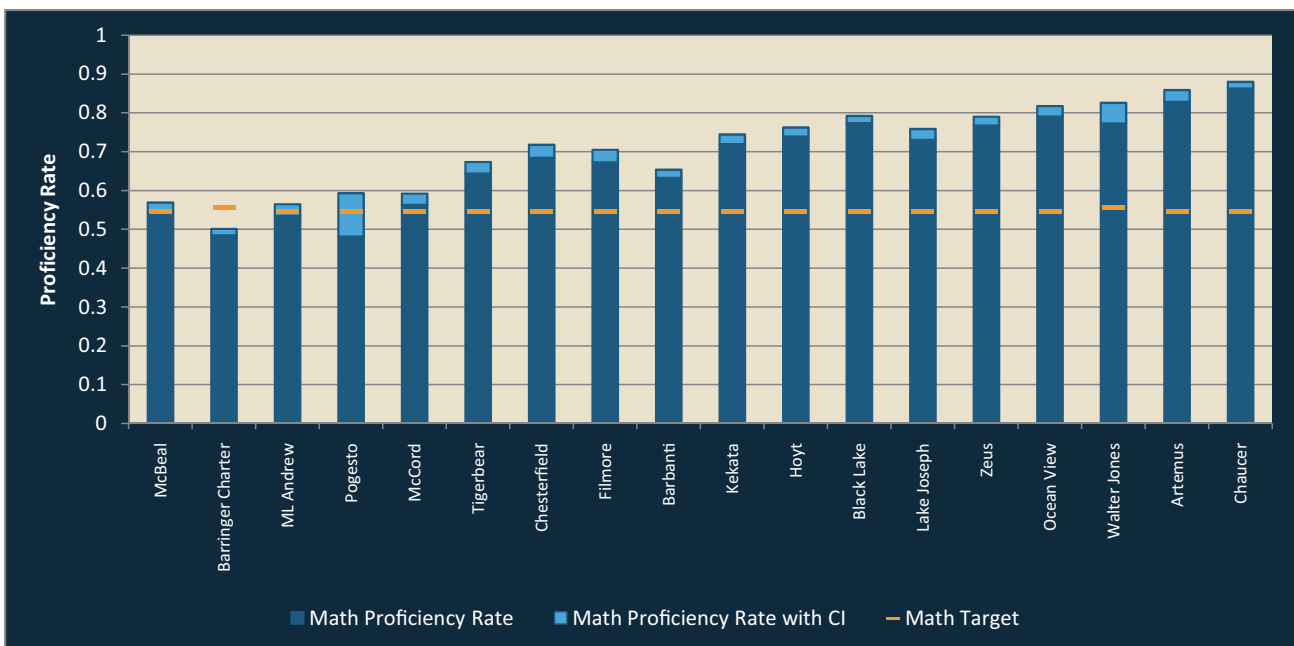


Figure 6. Impact of the confidence on middle school math proficiency rates for 2008s

Note: This figure shows the reported proficiency rate for the student population as a whole and the impact of the confidence interval on meeting annual targets. The darker portions of the bars show the actual proficiency rate achieved, while the lighter (upper) portions of the bars show the margin of error as computed by the confidence interval. The figure shows that two of the sample middle schools (ML Andrew and Pogesto) were assisted by the confidence interval. Annual targets (the orange lines) are considered to be met by the confidence interval if they fall within the light blue portion.

Table 2. Elementary school subgroup performance of sample schools under the 2008 Nevada 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	37.9%	23.7%	N	N	N	N	N	N	N	N					N	N					10	0	0%	N	1	
Maryweather	46.1%	37.9%	N	N	N	N	N	N	N	N	N			N	N					Y	Y	14	2	14%	N	1
Few	51.2%	38.8%	N	N	N	N	N	N	N	N	Y	Y		Y	N					Y	N	14	4	29%	N	1
Nemo	51.2%	52.6%	Y	Y	N	N			N	N	N	N								Y	Y	10	4	40%	N	7
Island Grove	52.1%	56.4%	Y	Y	N	N	N	N	N	Y				N	N					Y	Y	12	5	42%	N	4
JFK	59.5%	46.6%	Y	N	N	N			Y	N	N	N								Y	Y	10	4	40%	N	3
Scholls	69.0%	56.1%	Y	Y	N	N	Y	N	Y	Y	N	N		Y	N					Y	Y	14	8	57%	N	7
Hissmore	68.1%	57.6%	Y	Y	N	N			Y	Y	Y	Y								Y	Y	10	8	80%	N	7
Wolf Creek	61.9%	58.9%	Y	Y	N	N	N	N	N	N				N	N					Y	Y	12	4	33%	N	5
Alice Mayberry	65.5%	57.4%	Y	Y	N	N			Y	N	Y	N								Y	Y	10	6	60%	N	9
Wayne Fine Arts	65.5%	67.8%	Y	Y					Y	Y	N	Y								Y	Y	8	7	88%	N	21
Winchester	67.5%	67.8%	Y	Y	N	N			Y	Y				Y	Y					Y	Y	10	8	80%	N	22
Coastal	74.4%	63.9%	Y	Y	N	N	N	N	Y	Y	Y	N		Y	Y					Y	Y	14	9	64%	N	3
Paramount	74.0%	66.8%	Y	Y	Y	Y	N	N	Y	N				Y	N					Y	Y	12	8	67%	N	7
Forest Lake	82.0%	76.1%	Y	Y	N	N			Y	Y										Y	Y	8	6	75%	N	8
Marigold	86.3%	76.9%	Y	Y	Y	N	Y	N	Y	Y			Y	Y	Y	N				Y	Y	14	11	79%	N	10
Roosevelt	86.9%	83.7%	Y	Y					Y	Y				Y	Y					Y	Y	8	8	100%	Y	28
King Richard	84.1%	82.7%	Y	Y	N	N	N	N	N	N				Y	Y					Y	Y	12	6	50%	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.

The effect of confidence intervals on reading proficiency rates for elementary and middle schools is much the same (not shown). In reading, only one elementary school (John F. Kennedy) and one middle school (Chesterfield) met the overall targets with the confidence interval, although we know from Figures 3 and 4 that these two schools still failed to meet all of their subgroup targets. In short, **the application of the confidence interval had only modest effect on whether the sample ele-**

mentary and middle schools met Nevada's overall reading and math targets.⁹

Where Do Schools Fail?

Figures 3 and 4 illustrate that schools with low or mid-dling performance can still pass AYP when the school has fewer targets to meet because it has fewer subgroups. These figures do not, however, indicate which subgroups

⁹ In the current analyses, confidence intervals were applied to both the overall school population and to all eligible subgroups in our sample schools. Thus, the ultimate impact of the confidence interval is likely larger than the impact depicted in Figures 5 and 6. However, we chose not to *show* how the confidence interval impacted subgroup performance because it would have added greatly to the report's length and complexity.

Table 3. Middle school subgroup performance of sample schools under the 2008 Nevada 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	54.2%	51.1%	Y	N	N	N	N	N	N	N	N	N	Y	Y	N	N	Y	N	Y	Y	18	6	33%	N	0
Barringer Charter	48.5%	46.8%	N	N	N	N			N	N	N	N			Y	Y			Y	Y	12	4	33%	N	0
ML Andrew	53.5%	54.5%	Y	N	N	N	N	N	N	N	N	N			Y	N			Y	Y	14	4	29%	N	0
Pogesto	48.1%	57.4%	Y	Y					Y	Y									Y	Y	6	6	100%	Y	15
McCord Charter	56.3%	59.4%	Y	Y	N	N	N		N	N	N	N			N	N			Y	Y	13	4	31%	N	0
Tigerbear	64.3%	53.7%	Y	N	N	N			Y	N	Y	N							Y	Y	10	5	50%	N	0
Chesterfield	68.5%	56.3%	Y	Y	N	N			Y	N	Y	N							Y	Y	10	6	60%	N	1
Filmore	67.3%	64.0%	Y	Y	N	N	N	N	Y	Y					Y	N			Y	Y	12	7	58%	N	1
Barbanti	63.3%	59.9%	Y	Y	N	N	N	N	N	N					Y	N			Y	Y	12	5	42%	N	0
Kekata	71.9%	64.0%	Y	Y	N	N	N	N	Y	N	Y	N			Y	N			Y	Y	14	7	50%	N	0
Hoyt	73.9%	67.0%	Y	Y	N	N			Y	Y	Y	N							Y	Y	10	7	70%	N	2
Black Lake	77.3%	67.5%	Y	Y	N	N	Y	N	Y	N	Y	N	Y	Y	Y	Y			Y	Y	16	11	69%	N	0
Lake Joseph	73.1%	70.6%	Y	Y	N	N	N	N	Y	Y					Y	Y			Y	Y	12	8	67%	N	2
Zeus	76.8%	70.5%	Y	Y	N	N	N	N	Y	Y	Y	Y			N	N			Y	Y	14	8	57%	N	1
Ocean View	79.1%	79.7%	Y	Y	Y	N	N	N	Y	N			Y	Y	Y	N			Y	Y	14	9	64%	N	2
Walter Jones	77.3%	76.0%	Y	Y					Y	Y					Y	Y			Y	Y	8	8	100%	Y	20
Artemus	82.9%	77.1%	Y	Y	N	N			Y	N			Y	Y	N	N			Y	Y	12	7	58%	N	3
Chaucer	86.3%	85.3%	Y	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y			Y	Y	16	12	75%	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.

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 cat-

egories: 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 Nevada rules, and the total number of states within the study in which that school met AYP.

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

- The majority of schools met their targets for their overall student school populations, but failed to make AYP because of the performance of one or more subgroups.

Table 4. Summary of subgroup performance of sample elementary schools under the 2008 Nevada 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	16	14	15
Students with limited English proficiency	10	8	10
Low-income students	18	7	9
African-American students	9	5	6
Asian/Pacific Islander students	1	0	0
Hispanic students	12	4	8
American Indian/Alaska Native students	0	0	0
White students	17	0	1

Table 5. Summary of subgroup performance of sample middle schools under the 2008 Nevada 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	16	15	16
Students with limited English proficiency	11	10	10
Low-income students	18	5	11
African-American students	11	4	9
Asian/Pacific Islander students	5	0	0
Hispanic students	14	4	9
American Indian/Alaska Native students	1	0	1
White students	18	0	0

- Four elementary schools failed to meet the reading targets and three failed to meet the math targets for their overall school populations.
- Four middle schools failed to meet the reading targets and one (Barringer Charter) failed to meet the math target for its overall school population.
- Low-income students tended to perform better on

their math targets than their reading targets especially at the middle school level .

Tables 4 and 5 summarize the performance of the various subgroups for elementary and middle schools, respectively. The performance of SWDs is proving challenging for schools under Nevada's system, particularly in middle schools, where this subgroup tends to have enough students to meet the state's minimum n of

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

	Elementary Schools		Middle Schools	
	Made AYP	Failed to make AYP	Made AYP	Failed to make AYP
Number of schools in sample	1	17	2	16
Average student body size	262	307	124	951
Average % low income	13	48	42	45
Average % nonwhite	19	42	27	46
Average performance [†]	8.85	0.78	0.40	-0.11
Average % growth [‡]	103	116	109	97
Average number of targets to meet	8	11	7	13

[†] 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.

25. In fact, for SWDs, only one elementary school (Paramount) (and no middle schools) met its targets in reading, and only two elementary schools (Paramount and Marigold) and one middle school (Ocean View) met their targets in math. Students with limited English proficiency are also struggling to meet the state's targets; every school with a large enough LEP population to qualify as a separate subgroup failed to meet its reading targets for these students.

Characteristics of Schools that Did and Didn't Make AYP

A close look at Figures 3 and 4 indicates that Nevada's NCLB accountability system is, in some respects, behaving like those in other states. For example, Roosevelt and Walter Jones were among those schools that made AYP in the greatest number of states—28 and 20. And these schools made AYP in Nevada, too.

But Nevada is also home to a few anomalies. First, consider Winchester Elementary (see Figure 3). It made AYP in 22 of the 28 states in our sample, but not in Nevada. In examining Table 2, we can see that Winchester met

the minimum numbers for the SWD subgroup. Many other states within the sample had no SWD subgroup because of their larger minimum subgroup size. With more accountable subgroups, Winchester didn't meet all its targets; hence it failed to make AYP. This is likely to be also true for Wayne Fine Arts Elementary, which failed for a single subgroup.

That fewer subgroups make it more likely to make AYP is consistent with the patterns shown in Table 6, which compares the schools that did and didn't make AYP on a number of academic and demographic dimensions. Within the sample, schools that make AYP do indeed show higher average student performance, but they also differ in the following ways: they have much smaller student populations (especially at the middle school level), fewer subgroups (and thus fewer targets to meet), and much lower percentages of nonwhite students.

Concluding Observations

This study examined the test performance data of students from 18 elementary and 18 middle schools across the country to see how these schools would fare under

Nevada's AYP rules (and AMOs) for 2008. We found that only 1 elementary school and 2 middle schools—3 in all, from a sample of 36—would have made AYP in Nevada.

Looking across the 28 state accountability systems examined in the study, this puts Nevada near the lower end of the sample distribution in terms of the number of schools making AYP (see Figure 1). We find that the number of elementary schools that made AYP in Nevada is exceeded in 26 other sample states. The high number of schools that didn't make AYP in Nevada is partly because Nevada's minimum subgroup size is relatively small compared to other states; this means more subgroups are held separately accountable for performance. In fact, a few sample schools that made AYP in most other states did not make it in Nevada, largely because of the state's *n* size. (This occurred despite Nevada's fairly low annual targets, which require barely half of students to be proficient in 2008).

Because the overriding goal of NCLB is to eliminate educational disparities within and across states, it's important to consider whether states' annual decisions about the progress of individual schools are consistent with this aim. In some respects, Nevada's NCLB accountability system is working exactly as Congress intended: identifying as "needing attention" schools with relatively high test score averages that mask low performance for partic-

ular groups of students, such as low-income or Hispanic students. Most sample schools made AYP in Nevada for their student populations as a whole, without considering subgroups. In the pre-NCLB era, such schools might have been considered effective or at least not in need of improvement, even though sizable numbers of their students aren't meeting state standards. Disaggregating data by race, income, and so on has made those students visible. That is surely a positive step.

Yet NCLB's design flaws are also readily apparent. Does it make sense that the size of a school's enrollment has so much influence over whether or not a school makes adequate yearly progress? Does it make sense that having 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 massive failure of these students to meet Nevada'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 students and SWDs, as for other students, but when almost no school is able to meet the goal, perhaps that indicates that 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.

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

¹⁰ See footnote 4.

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.