



## 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 Texas’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 Texas’s system is compared with those of 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 Texas’s system as well as under the systems of 27 other states. We used school data and proficiency cut score<sup>1</sup> estimates from academic year 2005–2006, but applied them against Texas’s AYP rules for academic year 2007–2008 (shortened to “2008” in this report).

Here are some key findings:

- We estimate that **4 of 18 elementary schools** in our sample **failed to make AYP** in 2008 under Texas’s accountability system.
- Looking across the **28 state accountability systems** examined in the study, we find that the number of

elementary schools making AYP in Texas was exceeded in just 2 other sample states (Arizona and Wisconsin). (Note that middle schools were not examined in Texas, unlike other states, since eighth grade cut scores were not available.)

- Part of the reason that so many schools make AYP in Texas is that its **proficiency standards are relatively easy, compared to other states. Schools also have fewer accountable subgroups in Texas, likely because the state has a relatively large minimum “n size” for holding subgroups accountable.**
- Nearly all the schools in our sample that failed to make AYP in Texas are meeting expected targets for their overall populations<sup>2</sup> but failing because of the performance of individual subgroups, particularly students with disabilities (SWDs) and students with limited English proficiency (LEP).<sup>3</sup>

Just four of 18 elementary schools in our sample fail to make AYP in 2008 under **Texas’s** accountability system. Looking across the 28 state accountability systems examined in the study, we find Texas to be among the least restrictive in terms of how many sample schools make AYP. This is likely due to a number of factors. First, Texas’s proficiency standards (or cut scores) are relatively easy. Almost all of Texas’s cut scores are below the 35th percentile. Second, Texas has a relatively large minimum *n* size for subgroup reporting, meaning that schools in Texas will have fewer accountable subgroups than would similar schools in other states. Unlike most other states, though, Texas does not report a confidence interval around its proficiency rates, but we generally found that they had limited impact on schools’ AYP status in the study anyway.

<sup>1</sup> A cut score is the minimum score a student must receive on the Texas Assessment of Knowledge and Skills in order to be considered proficient under Texas’s accountability system.

<sup>2</sup> 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.

<sup>3</sup> SWDs are defined as those students following individualized education plans. Also, note that we use “LEP students” and “English language learners” interchangeably to refer to students in the same subgroup.

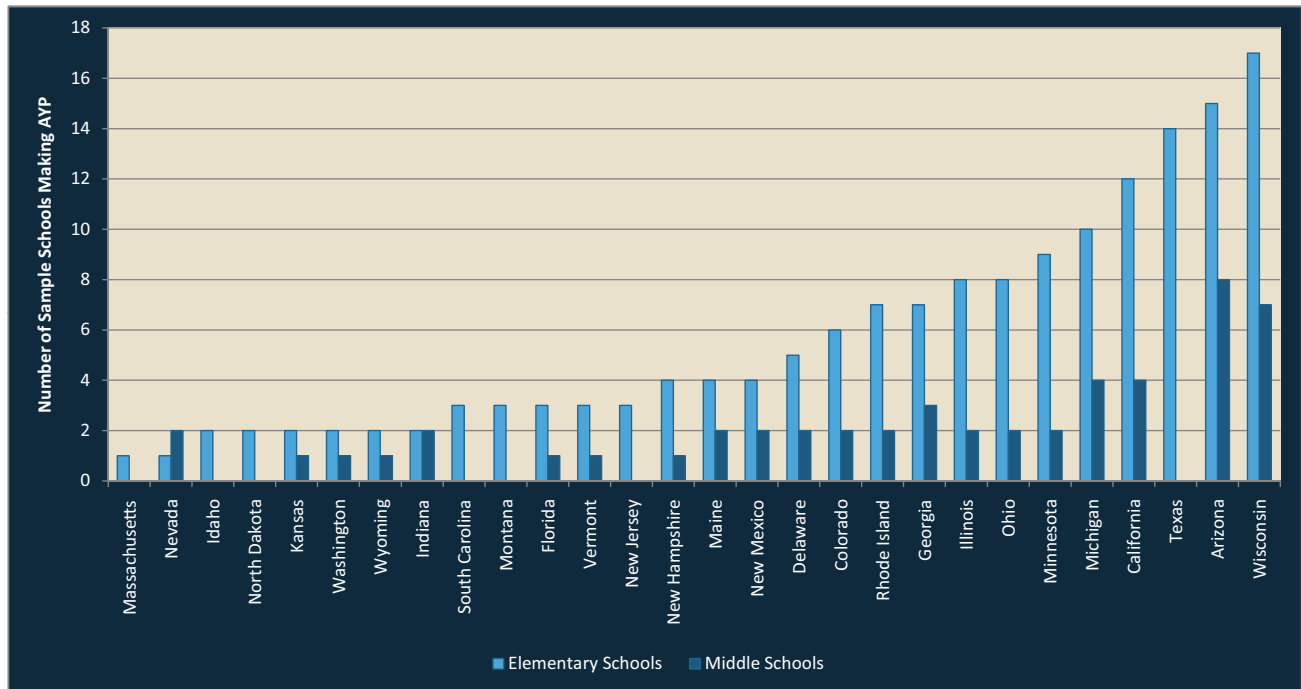


Figure 1. Number of sample schools making AYP by state

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.

- Ten sample elementary schools that failed to make AYP in most other states made AYP in Texas. Again, this is likely due to the state’s easy proficiency standards and large minimum subgroup size.
- In Texas, as is the case in most states, schools with fewer subgroups attain AYP more easily than 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.
- A strong predictor of a school making AYP under Texas’s system is whether it has enough SWDs or LEP students to qualify as a separate subgroup. Every single school with these subgroups failed to make AYP.<sup>4</sup>

## Introduction

*The Proficiency Illusion* (Cronin et al. 2007a) linked student performance on Texas’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 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

<sup>4</sup> It should be noted that our subgroup findings for Limited English Proficient (LEP) and students with disabilities may be slightly more negative than would be seen under real world conditions. This is mostly due to the differences in testing practices between how LEP students and students with disabilities are treated in the Texas Assessment of Knowledge and Skills (TAKS) state assessment and in the NWEA’s Measures of Academic Progress (MAP), the assessment used in this study. Specifically, the U.S. Department of Education has issued NCLB guidelines permitting schools to exclude small percentages of LEP or disabled students from taking state tests, or providing them alternate assessments. In the current study, however, no valid MAP scores were omitted from consideration.

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 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 income<sup>5</sup> 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 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 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**

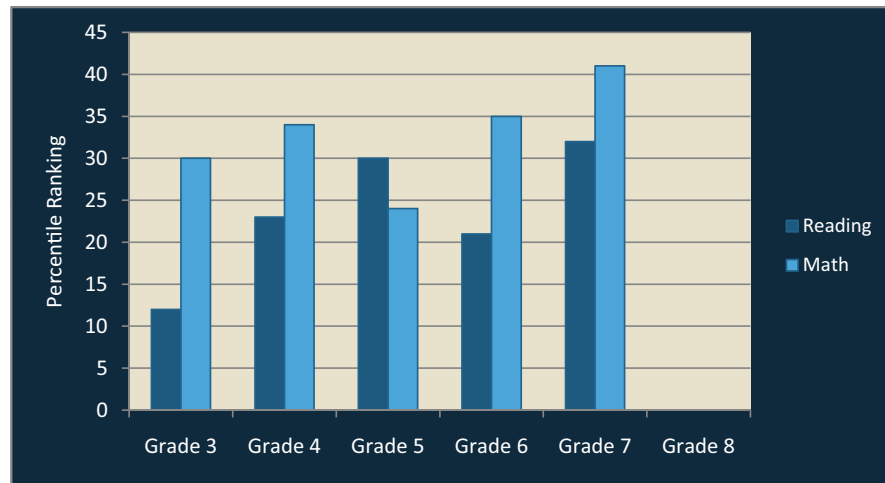
We collected students' MAP test scores from the 2005–2006 academic year from 18 elementary and 18 middle schools around the country. We also collected the NCLB subgroup designations for all students in those schools—in other words, whether they had been classified as members of a minority group, such 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.<sup>6</sup>

Proficiency cut score estimates for the Texas Assessment of Knowledge and Skills (TAKS) are taken from *The Proficiency Illusion* (as shown in Figure 2), which found that Texas's definitions of proficiency were below the average, or less difficult, 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 Texas test, given their performance on MAP. Student test data and subgroup designations are then used to determine how these 18 elementary schools would have fared under Texas AYP rules for 2008. In

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

<sup>6</sup> We gave all schools in our sample pseudonyms in this report.



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

Note: This figure illustrates the difficulty of Texas'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. All of Texas's cut scores are below the 45th percentile. Cut scores for eighth grade were not available.

other words, the school data and our proficiency cut score estimates are from academic year 2005–2006, but we are applying them against Texas's 2008 AYP rules. Note that in Texas, unlike most of the other state reports, the 18 sample middle schools were not examined, since the eighth grade cut scores were not available for Texas. Consequently, for Texas, only the performance of the sample elementary schools was examined.

Table 1 shows the pertinent Texas AYP rules that were applied to elementary schools in this study. Texas's minimum subgroup size is 10% of the population, if that is at least 50 but not more than 200.<sup>7</sup> This is a larger subgroup size than in many of the other states examined, meaning that schools in Texas will have fewer accountable subgroups than would similar schools in other states. Unlike most of the states in the study, Texas does not report a confidence interval around its proficiency rates. This means that schools in Texas will have greater difficulty achieving their targets than would schools that do use confidence intervals.

**Note that we were unable to examine the effect 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 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 school's 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

<sup>7</sup> In Texas, the minimum subgroup size is 10% of the total school population. Generally, this means that the subgroup size grows with the school size. However, there's also a clause that specifies the minimum subgroup size can't be less than 50 or more than 200. For example, a school with a total population of 1000 would have a minimum subgroup size of 100 (i.e., 10%), but a school with only 400 students would have a minimum subgroup size of 50, since 10% of 400 (i.e., 40) is below the minimum. Similarly, a school with 3,000 students would have a minimum subgroup size of 200, since 10% of 3,000 (i.e., 300) is greater than the maximum value.

Table 1. Texas AYP rules for 2008

Subgroup minimum <i>n</i>	Race/ethnicity: 10% of school population if at least 50 but not more than 200	
	SWDs: 10% of school population if at least 50 but not more than 200	
	Low-income students: 10% of school population if at least 50 but not more than 200	
	LEP students: 10% of school population if at least 50 but not more than 200	
CI	Applied to proficiency rate calculations?	
	Not used	
AMOs	Baseline proficiency levels as of 2002 (%)	2008 targets (%)
READING/LANGUAGE ARTS		
Grade 3	46.8	60.0
Grade 4	46.8	60.0
Grade 5	46.8	60.0
Grade 6	46.8	60.0
Grade 7	46.8	60.0
Grade 8	46.8	60.0
MATH		
Grade 3	33.4	50.0
Grade 4	33.4	50.0
Grade 5	33.4	50.0
Grade 6	33.4	50.0
Grade 7	33.4	50.0
Grade 8	33.4	50.0

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

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 Texas's AYP Rules?

Figure 3 illustrates the AYP performance of the sample elementary schools under Texas's 2008 AYP rules. **Fourteen elementary schools made AYP while only 4 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. Three of the schools not making AYP (Clarkson, Maryweather and Few) are in the left half of the figure, meaning that the lowest performing students were found at these schools.

Yet almost without regard to average student performance, the schools that failed to make AYP were those with relatively more qualifying subgroups—and thus the most targets to meet (because each subgroup has separate targets). For example, Coastal has relatively high performing students when compared to the other schools in the sample. However, it has the highest number of targets (12) and did not make AYP; whereas, Nemo is a school with lower performing students and made AYP, likely due to the low number of targets (6).

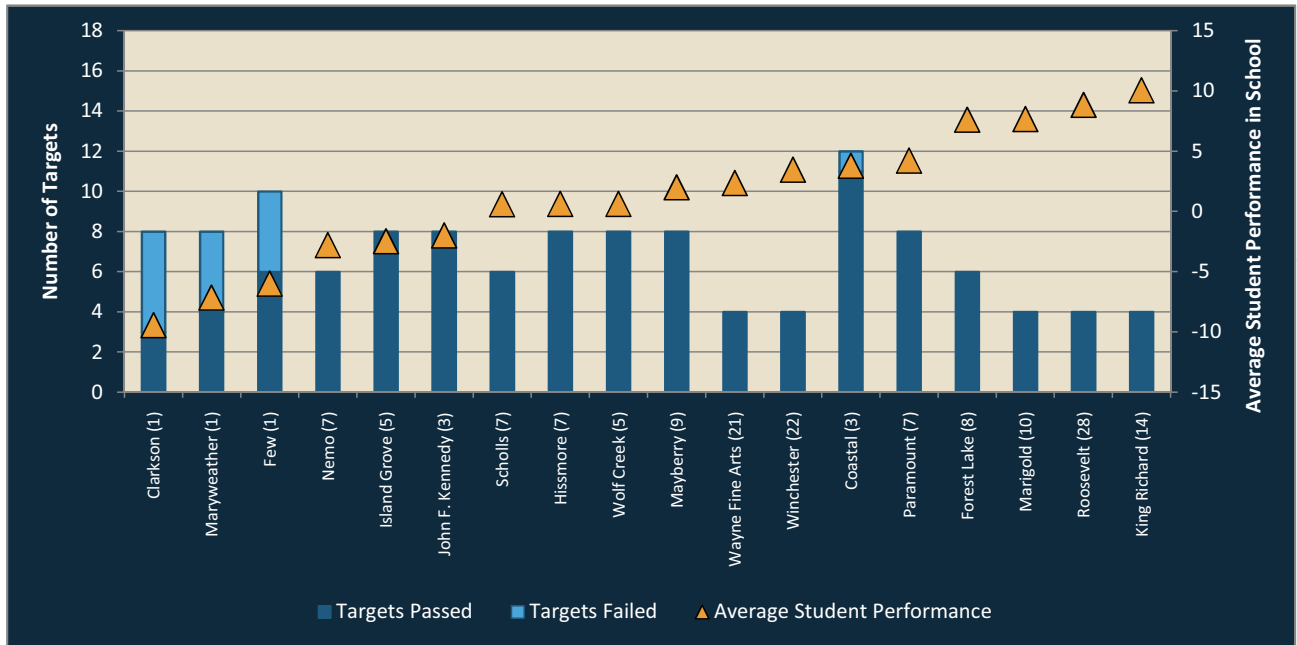


Figure 3. AYP performance of the elementary school sample under Texas's 2008 AYP rules

Note: This figure indicates how each elementary school within the sample fared under Texas'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 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. Coastal Elementary, for example, meets 11 of its 12 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), 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.

### Where do schools fail?

Figure 3 illustrates that schools with low or middling performance can still make AYP when the school has fewer targets to meet because it has fewer subgroups. This figure does not indicate which subgroups failed or passed in which school. Table 2 lists information on individual subgroup performance.

Table 2 shows 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. Table 2 also shows whether a school met AYP under the Texas rules, and the total number of states within the study in which that school met AYP.

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

- Only 2 schools have enough SWDs to comprise a separate subgroup. Only three schools have enough LEP students to comprise a separate subgroup. None of these schools made AYP.
- One elementary school (Clarkson) failed to meet the reading targets for its overall school population. No elementary schools failed to meet their overall math targets.
- One failing elementary school (Coastal) met its targets for every subgroup except for SWDs.
- All low income subgroups met their math targets.

Table 3 summarizes the performance of the various subgroups. First, the performance of LEP students is proving challenging for schools under Texas's system; all three schools with large enough LEP populations to qualify as

Table 2. Elementary school subgroup performance of sample schools under the 2008 Texas 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	AYP	Targets				
Clarkson	56.2%	56.1%	Y	N			N	N	Y	N					Y	N					8	3	38%	N	1	
Maryweather	59.8%	62.1%	Y	Y			N	N	Y	N					Y	N					8	4	50%	N	1	
Few	69.1%	66.3%	Y	Y	N	N	N	N	Y	Y					Y	Y					10	6	60%	N	1	
Nemo	68.8%	80.5%	Y	Y					Y	Y									Y	Y	6	6	100%	Y	7	
Island Grove	72.7%	77.0%	Y	Y					Y	Y					Y	Y				Y	Y	8	8	100%	Y	4
JFK	75.5%	73.9%	Y	Y					Y	Y	Y	Y								Y	Y	8	8	100%	Y	3
Scholls	82.8%	78.8%	Y	Y					Y	Y										Y	Y	6	6	100%	Y	7
Hissmore	82.5%	82.8%	Y	Y					Y	Y	Y	Y								Y	Y	8	8	100%	Y	7
Wolf Creek	72.9%	79.0%	Y	Y					Y	Y					Y	Y				Y	Y	8	8	100%	Y	5
Alice Mayberry	82.4%	83.7%	Y	Y					Y	Y	Y	Y								Y	Y	8	8	100%	Y	9
Wayne Fine Arts	83.3%	91.4%	Y	Y																Y	Y	4	4	100%	Y	21
Winchester	82.1%	86.3%	Y	Y																Y	Y	4	4	100%	Y	22
Coastal	84.9%	79.4%	Y	Y	Y	N			Y	Y	Y	Y			Y	Y				Y	Y	12	11	92%	N	3
Paramount	82.9%	82.5%	Y	Y					Y	Y					Y	Y				Y	Y	8	8	100%	Y	7
Forest Lake	90.9%	90.3%	Y	Y					Y	Y										Y	Y	6	6	100%	Y	8
Marigold	92.8%	89.2%	Y	Y																Y	Y	4	4	100%	Y	10
Roosevelt	95.6%	96.3%	Y	Y																Y	Y	4	4	100%	Y	28
King Richard	90.5%	91.5%	Y	Y																Y	Y	4	4	100%	Y	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.

separate subgroups fail to meet their reading and math targets for these students. SWDs are also struggling to meet the state's targets. Neither of the two schools with qualifying SWD subgroups made AYP.

### Characteristics of Schools that Did and Didn't Make AYP

A close look at Figure 3 indicates that Texas's NCLB accountability system is, in some respects, behaving like those in other states. For example, among the elementary

schools in our sample, Roosevelt, Winchester, and Wayne Fine Arts all made AYP in the greatest number of states—28, 22, and 21, respectively. And these schools all made AYP in Texas, too. But Texas is also home to quite a few anomalies. First, consider JFK Elementary School (Figure 3). Even with its relatively low average performance it made AYP in Texas, but failed to do so in 25 of 28 states. Its AYP success in Texas is most likely attributable to its relatively small number of targets under Texas's minimum subgroup size rule (see Table 2), along with Texas's relatively easy proficiency cut scores, compared to other states.

**Table 3.** Summary of subgroup performance of sample elementary schools under the 2008 Texas 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	1	2
Students with limited English proficiency	3	3	3
Low-income students	13	0	2
African-American students	4	0	0
Asian/Pacific Islander students	0	0	0
Hispanic students	7	0	2
American Indian/Alaska Native students	0	0	0
White students	15	0	0

**Table 4.** Comparisons between schools that did and didn't make AYP in Texas, 2008

	Elementary Schools	
	Made AYP	Failed to make AYP
<b>Number of schools in sample</b>	14	4
<b>Average student body size</b>	281	387
<b>Average % low income</b>	37	79
<b>Average % nonwhite</b>	31	76
<b>Average performance<sup>†</sup></b>	2.92	-4.69
<b>Average % growth<sup>‡</sup></b>	117	109
<b>Average number of targets to meet</b>	6	10

<sup>†</sup> 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.

<sup>‡</sup> 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.

This is consistent with the patterns shown in Table 4, which compares schools that made 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 pop-

ulations, fewer subgroups (and thus fewer targets to meet), and much 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 country to see how these schools would fare under Texas’s AYP rules and annual measurable objectives for 2008. Among this sample, 14 elementary schools in Texas—14 from an elementary school sample of 18—would have made AYP in Texas (this study did not include examination of Texas middle schools). Looking across the 28 state accountability systems examined in the study, this puts Texas at the high end of the distribution in terms of the number of schools making AYP (see Figure 1). **The fairly large number of schools making AYP in Texas may be due to the fact that Texas’s proficiency standards are relatively easy, compared to other states and because the state has a relatively large minimum *n* size for subgroup reporting, meaning fewer groups are held accountable than might be the case in other states.**<sup>8</sup> In fact, only two schools have enough SWDs to comprise a separate subgroup and only three schools have enough students with LEP to comprise a separate subgroup.

Because the overriding goal of the federal 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, Texas’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 students. All but one of the sample schools met the Texas reading and math targets for their student populations as a whole. 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 weren’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 making AYP? Does it make sense that having fewer subgroups enhances the likelihood of making AYP? Is it “fair,” in Texas’s case, that so few SWDs and students with LEP are counted separately, meaning schools have to meet fewer targets? And in the rare cases when they do count separately, that they consistently fail to meet their annual targets? These will be critical considerations for Congress as it takes up NCLB re-authorization 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 AMOs. A few schools would have also passed under the new growth-model pilots currently under way in

<sup>8</sup> Keep in mind, however, that school size and *n* size are related (larger *n* sizes may make sense for larger schools).

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.