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 Rhode Island'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 Rhode Island'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 Rhode Island's system as well as under the systems of twenty-seven other states. We used school data and proficiency cut score[1] estimates from academic year 2005–2006, but applied them against Rhode Island's AYP rules for academic year 2007–2008 (shortened to "2008" in this report).

Here are some key findings:

- We estimate that **11 of 18 elementary schools** and **16 of 18 middle schools** in our sample **failed to make AYP** in 2008 under Rhode Island's accountability system. (This failure rate is partly explained by our sample, which intentionally includes some schools with relatively large populations of low-performing students.)

- Looking across the 28 state **accountability systems examined in the study, we find that the number of elementary schools making AYP in Rhode Island was exceeded in just 8 other sample states** (Rhode Island ties Georgia with 7 elementary schools making AYP). (See Figure 1.)

- Many of the schools in our sample that failed to make AYP in Rhode Island are meeting expected targets for their overall populations but didn't make AYP because of the performance of individual subgroups, particularly students with disabilities (SWDs) and English language learners.[2]

- Two sample schools that failed to make AYP in most other states made AYP in Rhode Island. **This is probably because Rhode Island's minimum subgroup**

Unlike most states, **Rhode Island** measures its student performance with a proficiency index, which gives partial credit for students achieving "partial proficiency." In the short term, the index makes it easier for Rhode Island schools to meet their targets. However, the effect of the index diminishes as the annual targets gradually approach the 100 percent proficiency requirement dictated under NCLB for 2014. Two sample schools make AYP in Rhode Island that fail to make AYP in most other states. This is likely because Rhode Island's minimum subgroup size (45) is larger than in most other states, meaning that schools have fewer accountable subgroups under Rhode Island rules than they would in another state. In addition, Rhode Island's proficiency standards are average when compared to other states, but its annual targets are fairly rigorous. In grades 3-5 reading, for example, Rhode Island requires roughly 84 percent of all subgroups to reach proficiency in order for a school to make AYP in 2008.

---

[1] A cut score is the minimum score a student must receive on the New England Common Assessment Program in order to be considered proficient under Rhode Island's accountability system.

[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.
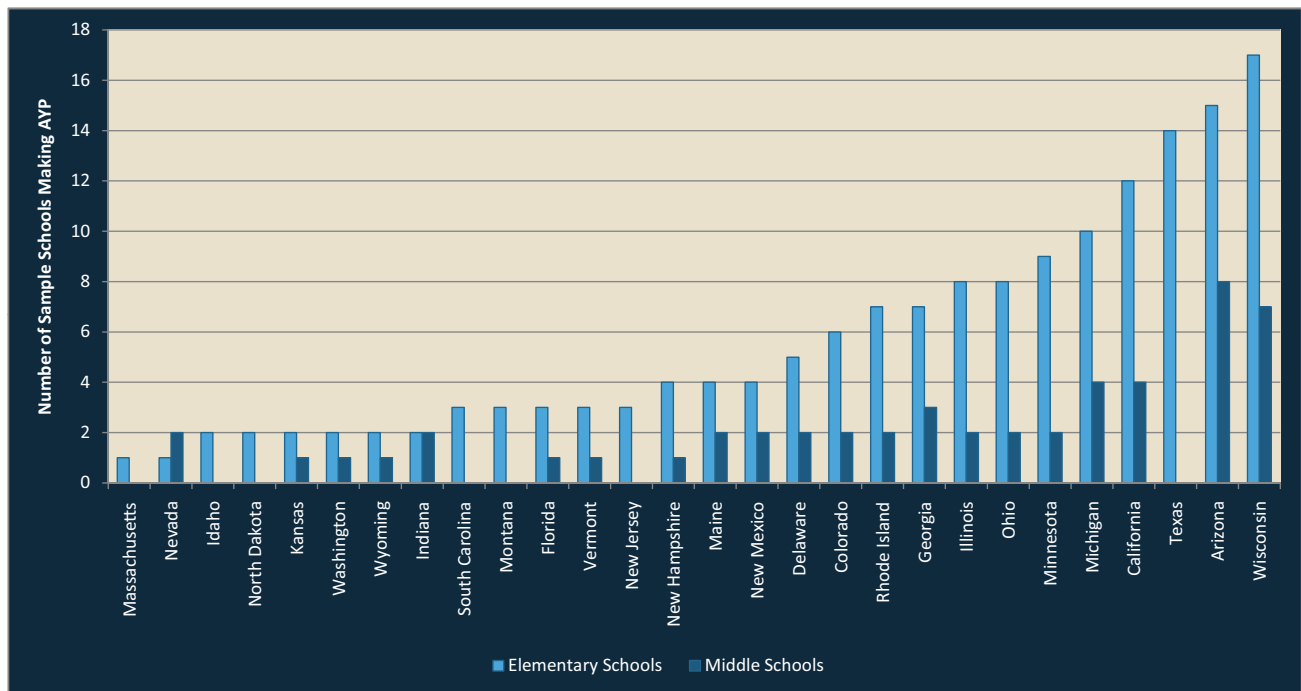
Figure 1. Number of sample schools making AYP by state

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.

sizes are a bit larger than in most other states, meaning that these schools had fewer accountable subgroups under Rhode Island rules.

- Rhode Island's proficiency standards (or cut scores) are above average compared to other states; similarly, the state's annual targets for proficiency are fairly ambitious, particularly at the elementary school level. However, Rhode Island uses a proficiency index, which gives partial credit for students achieving "partial proficiency." **In the short term, the index makes it easier for schools in Rhode Island to meet their targets, although the effect of the index diminishes as the targets approach the 100% proficiency requirement dictated under NCLB for 2014.**[3]

  - In Rhode Island, as in most states, schools with fewer subgroups attain AYP more easily in Rhode Is-

land 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.

- As in other states, middle schools in Rhode Island 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 Rhode Island's system is whether it has enough English language learners to qualify as a separate subgroup. Every school with a subgroup of students with limited English proficiency (LEP)[4] failed to make AYP. Likewise, all but one school

---

[3] Rhode Island is one of six states studied (Massachusetts, Minnesota, Vermont, Wisconsin, and New Hampshire are the others) that uses an index that gives full credit to students who achieve proficient (or better) and partial credit to students performing at lower levels. Consequently, the resultant score in states using this "hybrid" model is always higher than the actual proficiency percentage (giving students partial credit for achieving lower proficiency levels is obviously better than no credit, at least for the schools' ratings). The index provides a fair amount of help when annual targets are below 50%; however, once targets rise above 75%, the index has far less impact.

[4] Note that we use "LEP students" and "English language learners" interchangeably to refer to students in the same subgroup.

(King Richard) with enough qualifying SWDs failed to meet their AYP targets.[5]

Introduction

The Proficiency Illusion (Cronin et al. 2007a) linked student performance on Rhode Island'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 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[6] 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

---

[5] 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 New England Common Assessment Program (NECAP), 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.

[6] Low-income students are those who receive a free or reduced-price lunch.
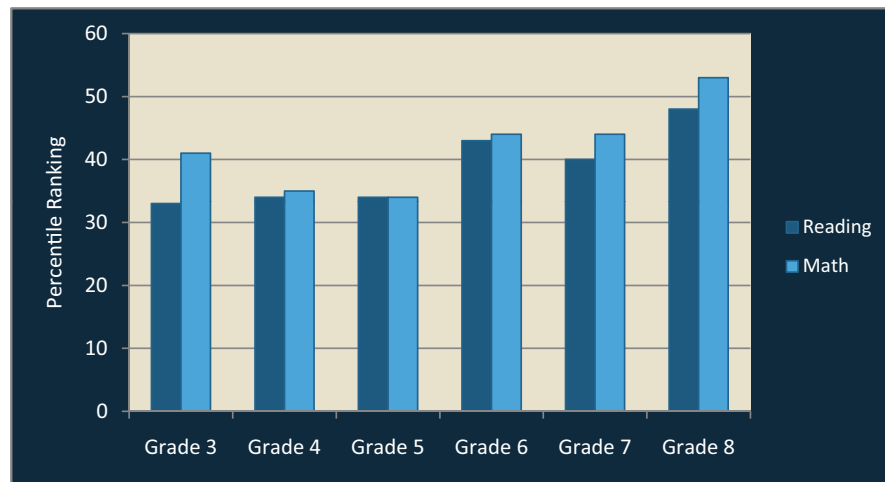
THOMAS B. FORDHAM INSTITUTE

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

Note: This figure illustrates the difficulty of Rhode Island'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 Rhode Island's cut scores are below the 55th percentile.

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.[7]

Proficiency cut score estimates for the New England Common Assessment Program (NECAP) are taken from *The Proficiency Illusion* (as shown in Figure 2), which found that Rhode Island'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 Rhode Island test (NECAP), given their performance on MAP. Student test data and subgroup designations were then used to determine how these 18 elementary and 18 mid-

dle schools would have fared under Rhode Island 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 Rhode Island's 2008 AYP rules.

Table 1 shows the pertinent Rhode Island AYP rules that were applied to elementary and middle schools in this study. Rhode Island's minimum subgroup size is 45, which is slightly larger than the subgroup size for other states we examined.[8] This means that schools in Rhode Island may have fewer subgroups than similar schools in other states.

Rhode Island, like the majority of states examined, applies a 95% confidence interval to its measurements of student proficiency rates. This means even though the AMO might require a school to attain, for instance, 84.1% reading proficiency among its grade 3 students, and 84.1% reading proficiency among its grade 3 students in each subgroup, the real target can be lower, particularly with smaller groups.

Unlike most states, Rhode Island measures its student performance with a proficiency index, which gives partial credit for students achieving "partial proficiency." In the

---

[7] We gave all schools in our sample pseudonyms in this report.

[8] Keep in mind, however, that school size and *n* size are related (e.g., large *n* sizes make sense for larger schools).

Table 1. Rhode Island AYP rules for 2008

Subgroup minimum *n*	Race/ethnicity: 45	
	SWDs: 45	
	Low-income students: 45	
	LEP students: 45	
CI	**Applied to proficiency rate calculations?**	
	Yes; 95% CI used	
AMOs	**Baseline proficiency levels as of 2002 (index)**	**2008 targets (index)**
READING/LANGUAGE ARTS		
Grade 3	76.1	84.1
Grade 4	76.1	84.1
Grade 5	76.1	84.1
Grade 6	68.0	78.6
Grade 7	68.0	78.6
Grade 8	68.0	78.6
MATH		
Grade 3	61.7	74.5
Grade 4	61.7	74.5
Grade 5	61.7	74.5
Grade 6	46.1	64.1
Grade 7	46.1	64.1
Grade 8	46.1	64.1

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

short term, the index makes it easier for Rhode Island schools to meet their targets, although the effect of the index diminishes as the targets approach the 100% proficiency requirement dictated under NCLB for 2014.[9]

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% rela-

tive 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

---

[9] In six of the states studied (Massachusetts, Minnesota, Vermont, New Hampshire, and Wisconsin, as well as Rhode Island), an index is used that gives full credit to students who achieve proficient (or better) and partial credit to students performing at lower levels. Consequently, the resultant score in states using this "hybrid" model is always higher than the actual proficiency percentage (giving students partial credit for achieving lower proficiency levels is obviously better than no credit, at least for the schools' ratings). The index provides a fair amount of help when annual targets are below 50%; however, once targets rise above 75%, the index has far less impact.
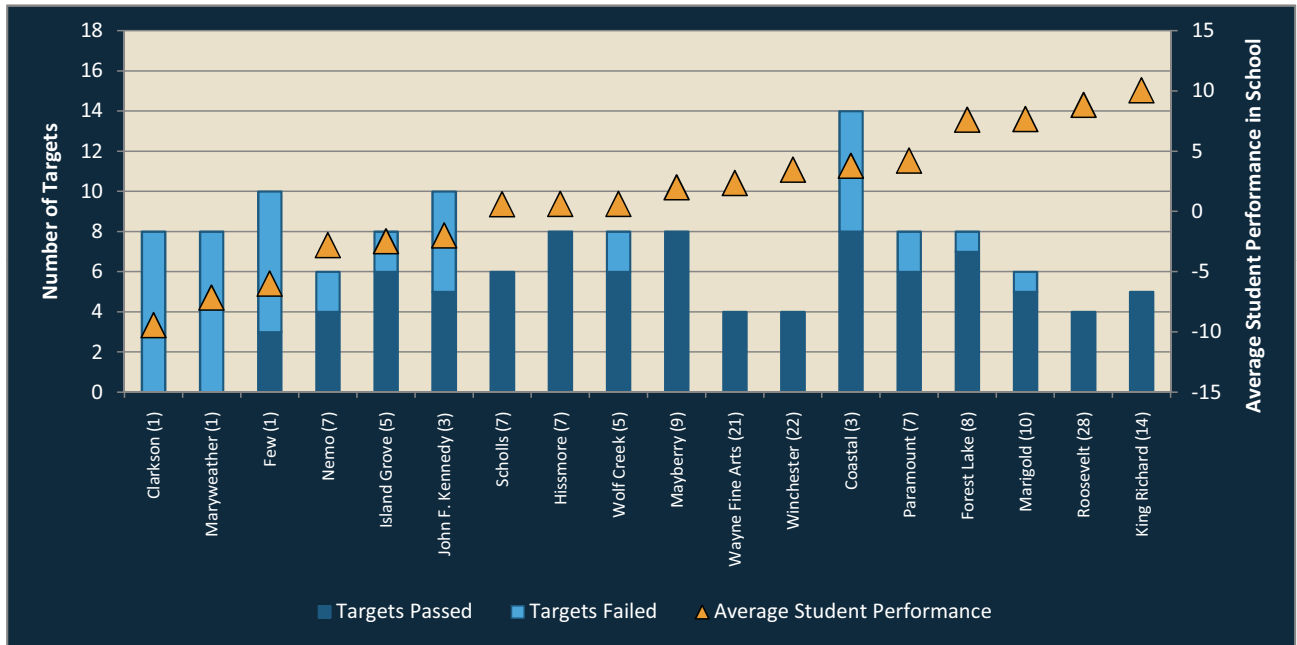
Thomas B. Fordham Institute

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

Note: This figure shows how each elementary school within the sample fares under Rhode Island'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. Marigold Elementary, for example, met five of its six 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.

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 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 Rhode Island's AYP Rules?

Figure 3 illustrates the AYP performance of the sample elementary schools under Rhode Island's 2008 AYP rules. **Seven elementary schools (Scholls, Hissmore, Mayberry, Wayne Fine Arts, Winchester, Roosevelt, and King Richard) made AYP and 11 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. For the most part, the schools that made AYP are those with relatively few qualifying subgroups—and thus the fewest targets to meet. For example, Scholls passed but has only six targets (two in reading and math for its overall school populations, two for its low income subgroups, and two for its white subgroups).

Figure 4 illustrates the AYP performance of the sample middle schools under the 2008 Rhode Island AYP rules. **Out of 18 middle schools in our sample, only 2 made AYP**—one low-performance school (Pogesto) and one high-performance school (Walter Jones), both of which
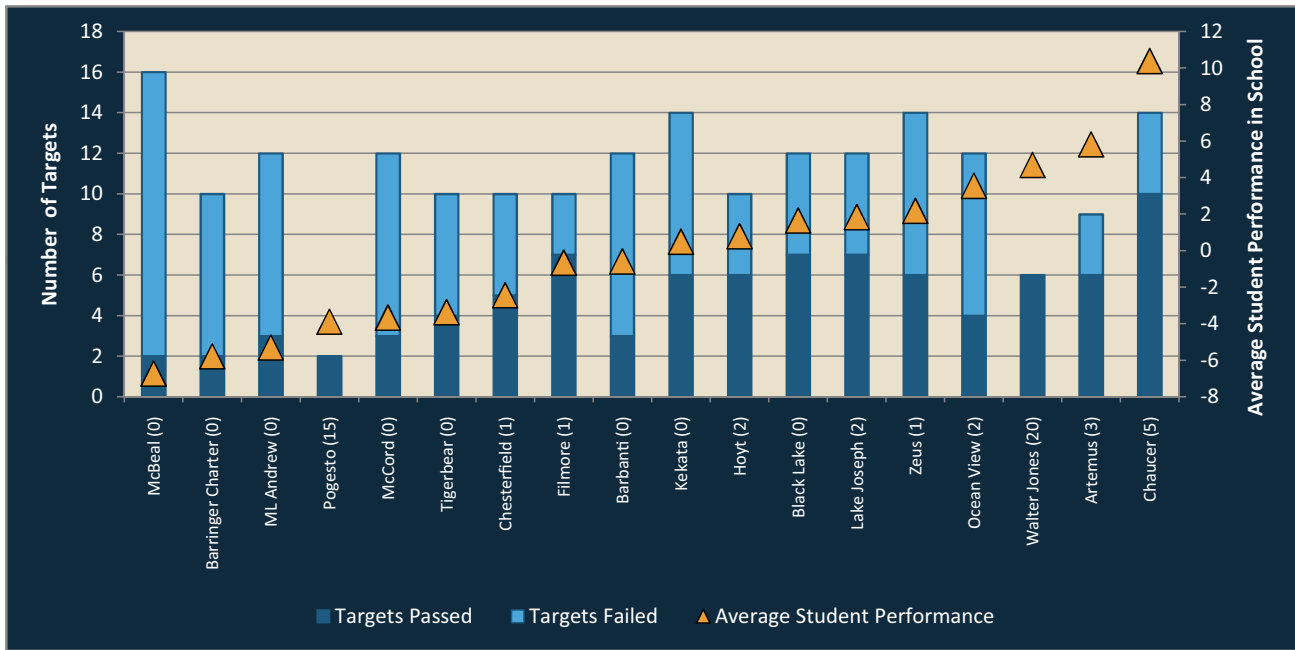
Figure 4. AYP performance of the middle school sample under Rhode Island's 2008 AYP rules

Note: This figure shows how each middle school within the sample would have fared under Rhode Island's AYP rules (as described in Table 1). The bars show the number of targets that each school had 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 AMO for even a single subgroup did not make AYP, so any light blue means that the school failed to make AYP. Artemus, for example, met 6 of its 9 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.

Figure 5. Impact of the confidence interval on elementary school mathematics proficiency rates under Rhode Island's 2008 AYP rules

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 one of the sample elementary schools (Few) was 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.
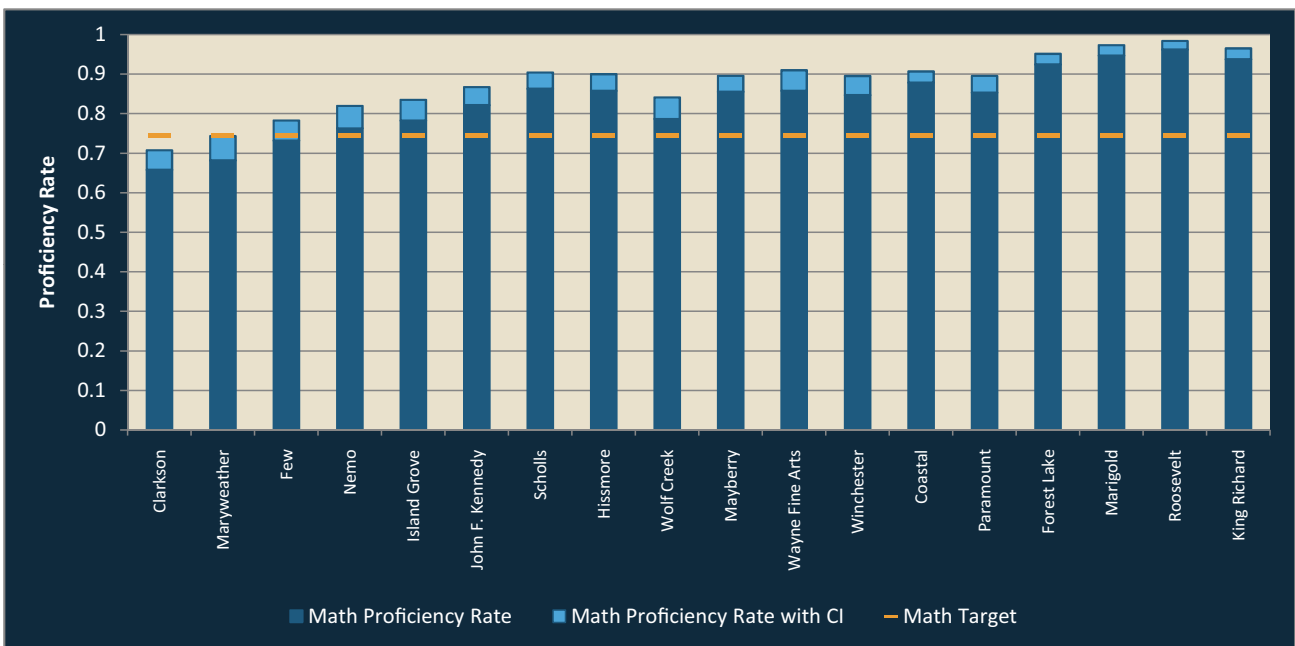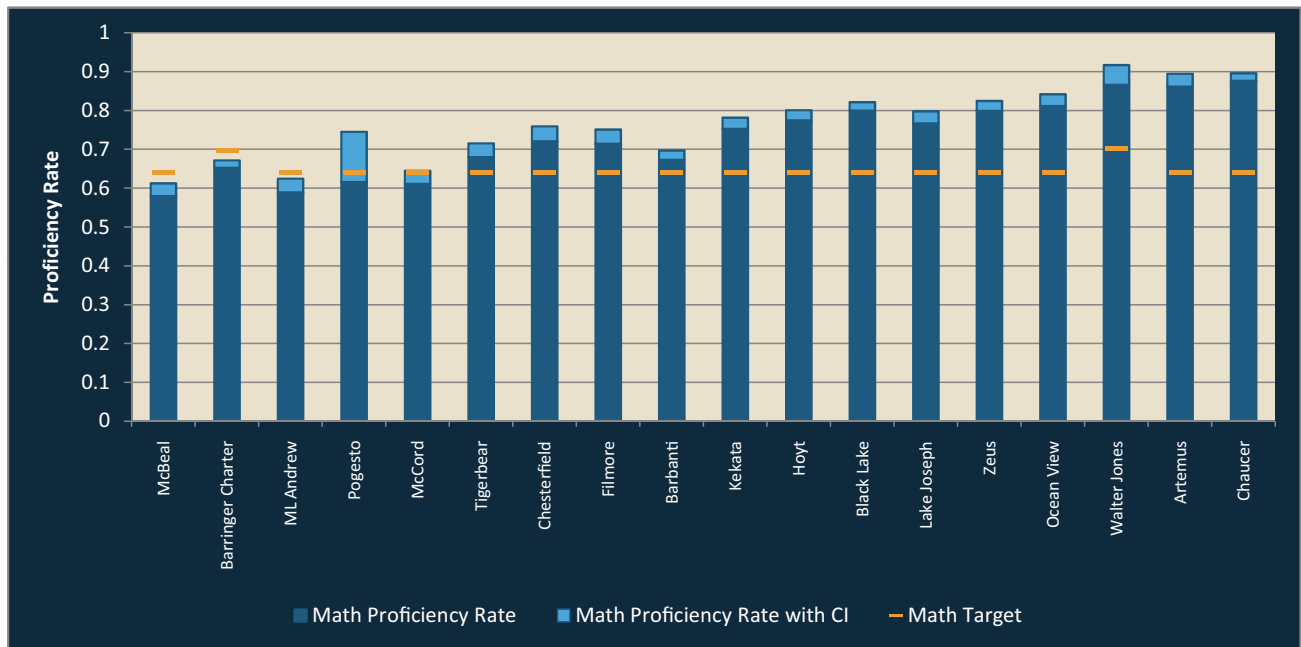
Figure 6. Impact of the confidence interval on middle school mathematics proficiency rates under Rhode Island's 2008 AYP rules

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 one of the sample middle schools (Pogesto) was 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.

have the lowest number of qualifying subgroups (and hence, targets to meet).

Figures 5 and 6 indicate the degree to which schools' math proficiency rates are aided by the confidence interval for elementary and middle schools, respectively. On these figures, the dark blue bars show the actual proficiency rates at each school, and the light blue bars show the degree to which these proficiency rates were increased by applying the confidence interval. The orange lines show the annual measurable objective needed to meet AYP. These figures show that one of the sample elementary schools (Few) and one middle school (Pogesto) are assisted by the confidence intervals, although we know from Figure 3 that Few Elementary still failed to make AYP because of subgroup performance. The effect of confidence intervals on schools' reading proficiency rates for elementary and middle schools is much the same (not shown).

In reading (not shown), four elementary schools (Nemo, Island Grove, Scholls, and Wolf Creek) and one middle school (Pogesto) are able to meet the overall target with the confidence interval, although we know from Figures 3 and 4 that most of these schools still failed to make AYP since they didn't meet targets for subgroups. **Overall, the application of the confidence interval appears to have modest impact on AYP decisions.**[10]

Where do schools fail?

Figures 3 and 4 illustrate that schools with low or middling 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 failed or passed in which school. Tables 2 and 3 list information on individual subgroup performance for elementary and middle schools, respectively.

[10] 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 may be 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 2. Elementary school subgroup performance of sample schools under the 2008 Rhode Island 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	65.8%	62.0%	N	N			N	N	N	N					N	N					8	0	0%	N	1
Maryweather	68.2%	66.4%	N	N			N	N	N	N					N	N					8	0	0%	N	1
Few	73.2%	69.7%	Y	N	N	N	N	N	Y	N					Y	N					10	3	30%	N	1
Nemo	76.3%	81.5%	Y	Y					N	N									Y	Y	6	4	67%	N	7
Island Grove	78.3%	79.9%	Y	Y					Y	N					Y	N			Y	Y	8	6	75%	N	4
JFK	82.2%	78.0%	Y	N	N	N			Y	N	Y	N							Y	Y	10	5	50%	N	3
Scholls	86.3%	81.7%	Y	Y					Y	Y									Y	Y	6	6	100%	Y	7
Hissmore	85.7%	84.3%	Y	Y					Y	Y	Y	Y							Y	Y	8	8	100%	Y	7
Wolf Creek	78.7%	81.2%	Y	Y					Y	N					Y	N			Y	Y	8	6	75%	N	5
Alice Mayberry	85.5%	86.2%	Y	Y					Y	Y	Y	Y							Y	Y	8	8	100%	Y	9
Wayne Fine Arts	85.8%	92.4%	Y	Y															Y	Y	4	4	100%	Y	21
Winchester	84.7%	88.3%	Y	Y															Y	Y	4	4	100%	Y	22
Coastal	87.9%	83.4%	Y	Y	N	N	Y	N	Y	N	Y	N			Y	N			Y	Y	14	8	57%	N	3
Paramount	85.3%	84.6%	Y	Y					Y	N					Y	N			Y	Y	8	6	75%	N	7
Forest Lake	92.4%	92.0%	Y	Y	Y	N			Y	Y									Y	Y	8	7	88%	N	8
Marigold	94.7%	91.2%	Y	Y					Y	N									Y	Y	6	5	83%	N	10
Roosevelt	96.2%	96.2%	Y	Y															Y	Y	4	4	100%	Y	28
King Richard	93.8%	93.5%	Y	Y	Y														Y	Y	5	5	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.

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 Rhode Island 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:

- Most schools met their targets for their overall school populations, except for two (Clarkson and Maryweather) that failed to meet the overall reading and math targets, and two others (Few and JFK) that failed to meet their overall reading targets.

- Three middle schools (McBeal, Barringer, and ML Andrew) failed to meet both their reading and math

Table 3. Middle school subgroup performance of sample schools under the 2008 Rhode Island 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	58.0%	65.0%	N	N	N	N	N	N	N	N	N	N			N	N	N	N	Y	Y	16	2	13%	N	0
Barringer Charter	65.3%	73.1%	N	N	N	N					N	N	N	N	Y	Y					10	2	20%	N	0
ML Andrew	58.9%	71.2%	N	N	N	N					N	N	N	N	Y	N			Y	Y	12	3	25%	N	0
Pogesto	61.6%	74.5%	Y	Y																	2	2	100%	Y	15
McCord Charter	61.1%	73.9%	Y	N	N	N					N	N	N	N	N	N			Y	Y	12	3	25%	N	0
Tigerbear	68.0%	68.8%	Y	N	N	N					Y	N	N	N					Y	Y	10	4	40%	N	0
Chesterfield	72.1%	72.9%	Y	N	N	N					Y	N	Y	N					Y	Y	10	5	50%	N	1
Filmore	71.4%	78.9%	Y	Y	N	N					Y	Y			Y	N			Y	Y	10	7	70%	N	1
Barbanti	67.3%	74.1%	Y	N	N	N	N	N	N	N					N	N			Y	Y	12	3	25%	N	0
Kekata	75.3%	76.5%	Y	Y	N	N	N	N	Y	N	Y	N			N	N			Y	Y	14	6	43%	N	0
Hoyt	77.5%	79.3%	Y	Y	N	N			Y	N	Y	N							Y	Y	10	6	60%	N	2
Black Lake	80.0%	78.7%	Y	Y	N	N			Y	N	Y	N			Y	N			Y	Y	12	7	58%	N	0
Lake Joseph	76.7%	82.4%	Y	Y	N	N	N	N	Y	Y					Y	N			Y	Y	12	7	58%	N	2
Zeus	79.9%	80.5%	Y	Y	N	N	N	N	Y	N	Y	N			N	N			Y	Y	14	6	43%	N	1
Ocean View	81.1%	87.3%	Y	Y	N	N	N	N	N	N					N	N			Y	Y	12	4	33%	N	2
Walter Jones	86.6%	88.9%	Y	Y					Y	Y									Y	Y	6	6	100%	Y	20
Artemus	86.2%	85.9%	Y	Y	N				Y	N					Y	N			Y	Y	9	6	67%	N	3
Chaucer	87.7%	91.0%	Y	Y	N	N	N	N	Y	Y			Y	Y	Y	Y			Y	Y	14	10	71%	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.

targets for their overall populations. Four others (McCord, Tigerbear, Chesterfield, and Barbanti) failed to meet their overall reading targets.

- One elementary school (Forest Lake) met its targets for every subgroup except SWDs.

- One elementary school (Marigold) met all its targets except for its low-income subgroup.

- No middle school met its targets for its SWD or students with LEP subgroups.

Tables 4 and 5 summarize the performance of the various subgroups for elementary and middle schools, respectively. Note, first, the performance of SWDs is proving quite challenging for schools under Rhode Island's system, particularly in middle schools, where these subgroups tend to have enough students to meet the state's minimum n of 45. In fact, all but one elementary (King Richard) and every middle school in the study with qualifying SWD subgroups failed to make AYP. Similarly, every single school with a population of students with LEP large enough to qualify as a separate subgroup failed to meet its reading targets for these students

Table 4. Summary of subgroup performance of sample elementary schools under the 2008 Rhode Island 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	5	3	4
Students with limited English proficiency	4	3	4
Low-income students	14	3	10
African-American students	4	0	2
Asian/Pacific Islander students	0	0	0
Hispanic students	7	2	7
American Indian/Alaska Native students	0	0	0
White students	15	0	0

Table 5. Summary of subgroup performance of sample middle schools under the 2008 Rhode Island 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	15	15	16
Students with limited English proficiency	7	7	7
Low-income students	17	6	13
African-American students	10	5	10
Asian/Pacific Islander students	1	0	0
Hispanic students	13	6	11
American Indian/Alaska Native students	1	1	1
White students	16	0	0

and only one such school (Coastal) reached its math target for this subgroup.

Characteristics of Schools that Did and Didn't Make AYP

A close look at Figures 3 and 4 indicates that Rhode Island's NCLB accountability system is, in many 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 Rhode Island, too. Likewise, the elementary and middle schools that failed to make AYP in the greatest number of states also failed to make AYP in Rhode Island.

But Rhode Island is also home to a few anomalies. First,

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

	Elementary Schools		Middle Schools	
	Made AYP	Failed to make AYP	Made AYP	Failed to make AYP
Number of schools in sample	7	11	2	16
Average student body size	268	328	124	951
Average % low income	37	52	42	45
Average % nonwhite	32	47	27	46
Average performance†	3.99	-0.54	0.40	-0.11
Average % growth‡	112	117	109	97
Average number of targets to meet	6	9	4	12

† 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.

consider Mayberry Elementary (see Figure 3). It failed to make AYP in 19 of the 28 states in our sample, yet made AYP in Rhode Island. Examining Table 2 we can see that Mayberry didn't meet the minimum numbers for the LEP or SWD subgroups, which create difficulty for so many other schools within the sample. With fewer accountable subgroups, Mayberry made meet AYP, even when other schools with higher average performance failed. Second, look at Pogesto Middle School (Figure 4). Even with its relatively low average performance, it made AYP in Rhode Island, while failing to do so in 13 of 28 states. Like Mayberry, its AYP success in Rhode Island is likely attributable to the relatively small number of targets (two) it has to meet, as shown in Table 3.

This is consistent with the patterns shown in Table 6, which compares schools making and not making 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 much smaller student populations (especially at the middle school level), fewer subgroups (and thus fewer targets to meet), and lower percentages of non-white 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 Rhode Island's AYP rules and AMOs for 2008. We found that 7 elementary schools and 2 middle schools—9 in all from a total of 36—would have made AYP in Rhode Island. Looking across the 28 state accountability systems examined in the study, this puts Rhode Island in the upper middle of the distribution in terms of the number of schools making AYP (as shown in Figure 1). Rhode Island's proficiency standards (or cut scores) are above average compared to other states; similarly, the state's annual targets for proficiency are fairly high, particularly at the elementary school level. However, Rhode Island uses a proficiency index, which gives partial credit for students achieving "partial proficiency." In the short term, such an index makes it easier for Rhode Island schools to meet their targets, although the effect of the index diminishes as the targets approach the 100% proficiency requirement dictated under NCLB for 2014.

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, Rhode Island'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 particular groups of students, such as low-income or Hispanic students. Most of the sample schools met the Rhode Island math and reading 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 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 making AYP? Does it make sense that having fewer subgroups enhances the likelihood of making AYP? Similarly, does it make sense that subgroup participation differs so much between elementary and middle schools, as it does in Rhode Island? Is it "fair" that, in Rhode Island and in a handful of other states, students are awarded "partial" credit even though they did not achieve proficiency? And equally important, doesn't the failure of English language learners and SWDs[11] to meet Rhode Island'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 students with LEP 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 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 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.

---

[11] See footnote 5.

THOMAS B. FORDHAM INSTITUTE

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.