



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 Indiana’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 Indiana’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 Indiana’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 Indiana’s AYP rules for academic year 2007–2008 (shortened to “2008” in this report).

Here are some key findings:

- We estimate that **16 of 18 elementary schools and 16 of 18 middle schools in our sample failed to make AYP** in 2008 under Indiana’s accountability system. (This high failure rate is partly explained by our sample, which intentionally includes some schools with a relatively large population of low-income students.²)

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 Indiana Statewide Testing for Educational Progress Plus.

² Low-income students are those who receive a free or reduced-price lunch.

³ It’s important to note that students in subgroups not meeting the minimum *n* sizes are still included for accountability purposes in the overall student calculations; they simply are not treated as their own subgroup.

- Looking across the 28 state accountability systems examined in the study (see Figure 1), only two states passed fewer of the sample elementary schools than Indiana. Indiana **tied with 5 other states with just 2 elementary schools making AYP.**
- Many of the schools in our sample that failed to make AYP in Indiana met 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.³
- In Indiana, schools with fewer subgroups attained AYP more easily than schools with more subgroups, even when their average student performance was much lower. In other words, schools with greater diversity and size face greater challenges in making AYP. This is the case in other states as well.
- A strong predictor of whether or not a school will make AYP under Indiana’s system is whether it has enough SWDs or English language learners to qualify

Indiana has fewer schools making AYP than in many other states in our study. This is particularly interesting because Indiana’s definitions of proficiency in reading and math generally ranked below the average compared with the standards set by the other states. However, Indiana’s annual targets in reading (the percentage of students in various subgroups that have to meet proficiency) are relatively difficult to achieve. Specifically, 72.4 percent of a given population in any school would have to be proficient on the state reading exam for the school to make AYP in 2008. In addition, Indiana’s minimum subgroup size is smaller than most other states’, meaning that schools in Indiana are accountable for more subgroups than similar schools in other states.

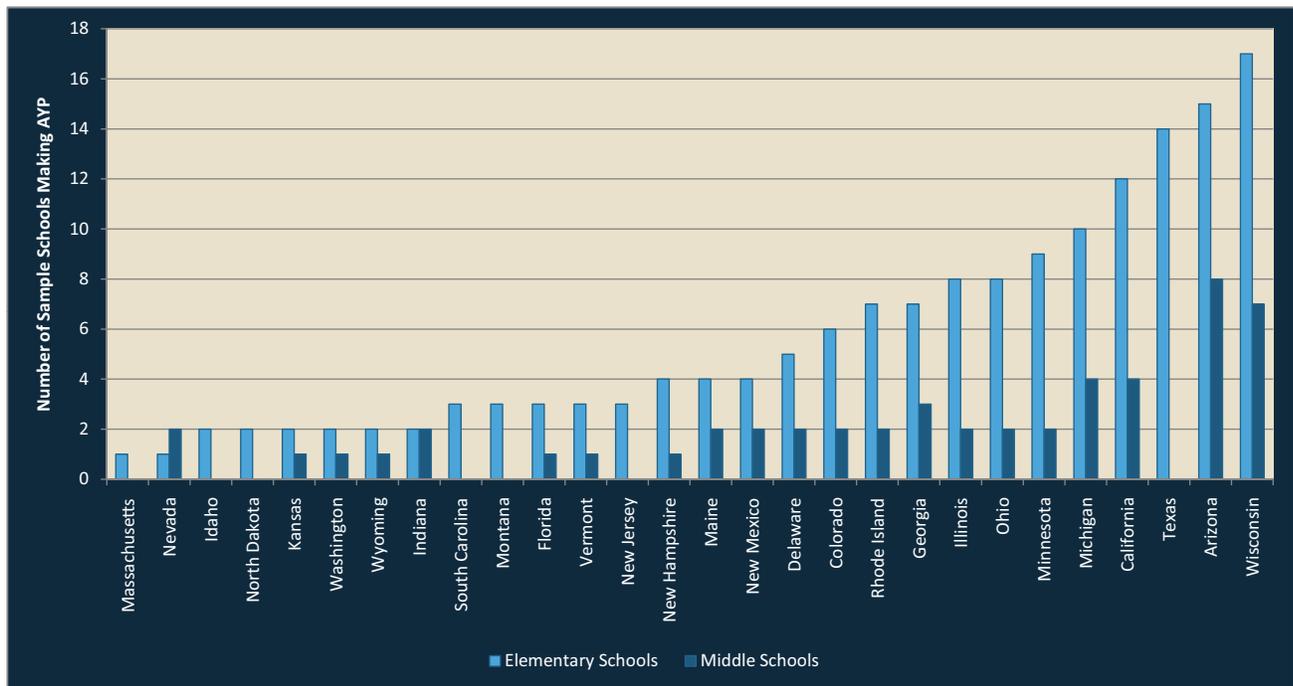


Figure 1. Number of sample schools making AYP by state

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

as a separate subgroup. Every school with limited English proficient (LEP)⁴ subgroups and SWDs failed to make AYP, in part because these students did not meet the state’s targets in reading and/or math.⁵

Introduction

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

⁴ Note that we use “LEP students” and “English language learners” interchangeably to refer to students in the same subgroup.

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 Indiana Statewide Testing for Educational Progress-Plus, 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.

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 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 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 Indiana Statewide Testing for Educational Progress-Plus (ISTEP+) are taken from *The Proficiency Illusion* (as shown in Figure 2), which found that Indiana’s definitions of proficiency generally ranked slightly below the average compared with the standards set by the other 25 states in that study. These cut scores were used to estimate whether students would have scored as proficient or better on the Indiana 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 Indiana 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 Indiana’s 2008 AYP rules.)

Table 1 shows the pertinent Indiana AYP rules that were applied to elementary and middle schools in this study. **Indiana’s minimum subgroup size is 30, which is smaller than most other states we examined.**⁷

Although most states examined also apply confidence intervals (or margins of statistical error) to their measure-

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

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

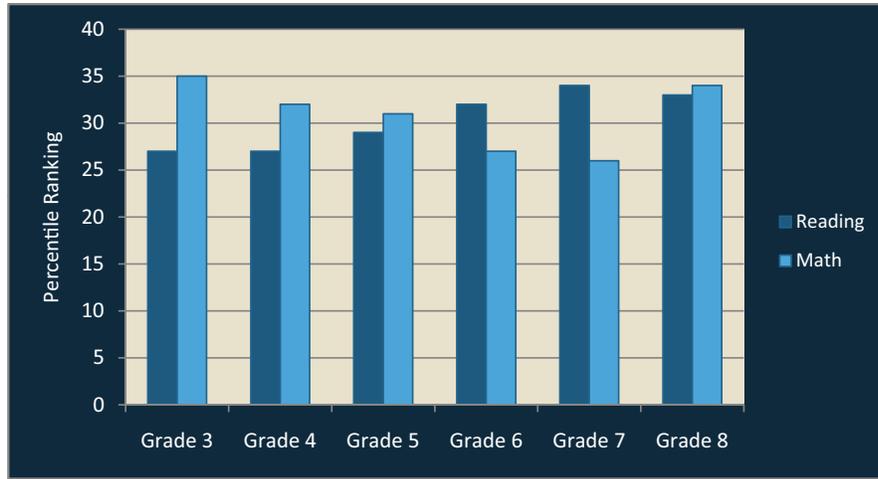


Figure 2. Indiana reading and math cut score estimates, expressed as percentile ranks (2006)

Note: This figure illustrates the difficulty of Indiana'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 Indiana's cut scores are at or below the 35th percentile.

Table 1. Indiana AYP rules for 2008

Subgroup minimum <i>n</i>	Race/ethnicity: 30	
	SWDs: 30	
	Low-income students: 30	
	LEP students: 30	
CI	Applied to proficiency rate calculations?	
	Yes; 99% CI	
AMOs	Baseline proficiency levels as of 2002 (%)	2008 targets (%)
READING/LANGUAGE ARTS		
Grade 3	58.8	72.4
Grade 4	58.8	72.4
Grade 5	58.8	72.4
Grade 6	58.8	72.4
Grade 7	58.8	72.4
Grade 8	58.8	72.4
MATH		
Grade 3	57.1	71.4
Grade 4	57.1	71.4
Grade 5	57.1	71.4
Grade 6	57.1	71.4
Grade 7	57.1	71.4
Grade 8	57.1	71.4

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

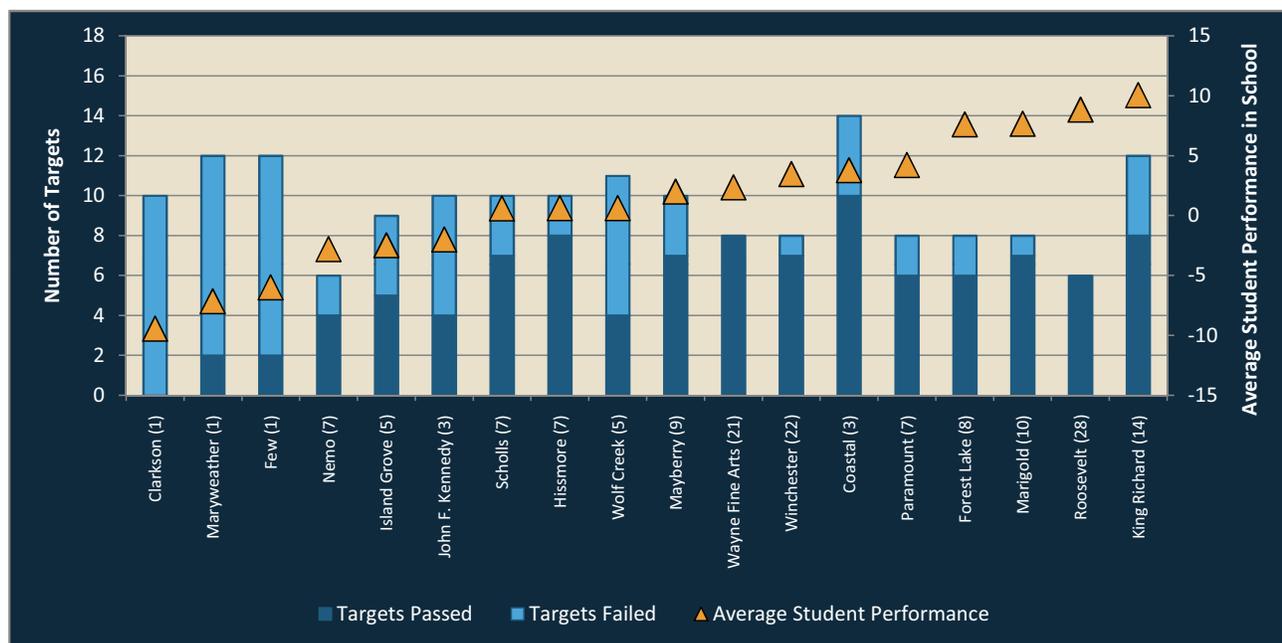


Figure 3. AYP performance of the elementary school sample under the Indiana 2008 AYP rules

Note: This figure indicates how each of the elementary schools within the sample fared under Indiana'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 make 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 didn't make AYP, so any light blue means the school failed to make AYP. Marigold 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), 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.

ments of student proficiency rates, **Indiana's 99% confidence interval gives schools greater leniency than the 95% confidence interval used by most other states.** So, for instance, although schools are supposed to get 72.4% of their students to the proficient level on the state reading test (and 72.4% of their students in each subgroup), applying the confidence interval means that the real target can actually be lower, particularly with smaller groups.⁸

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

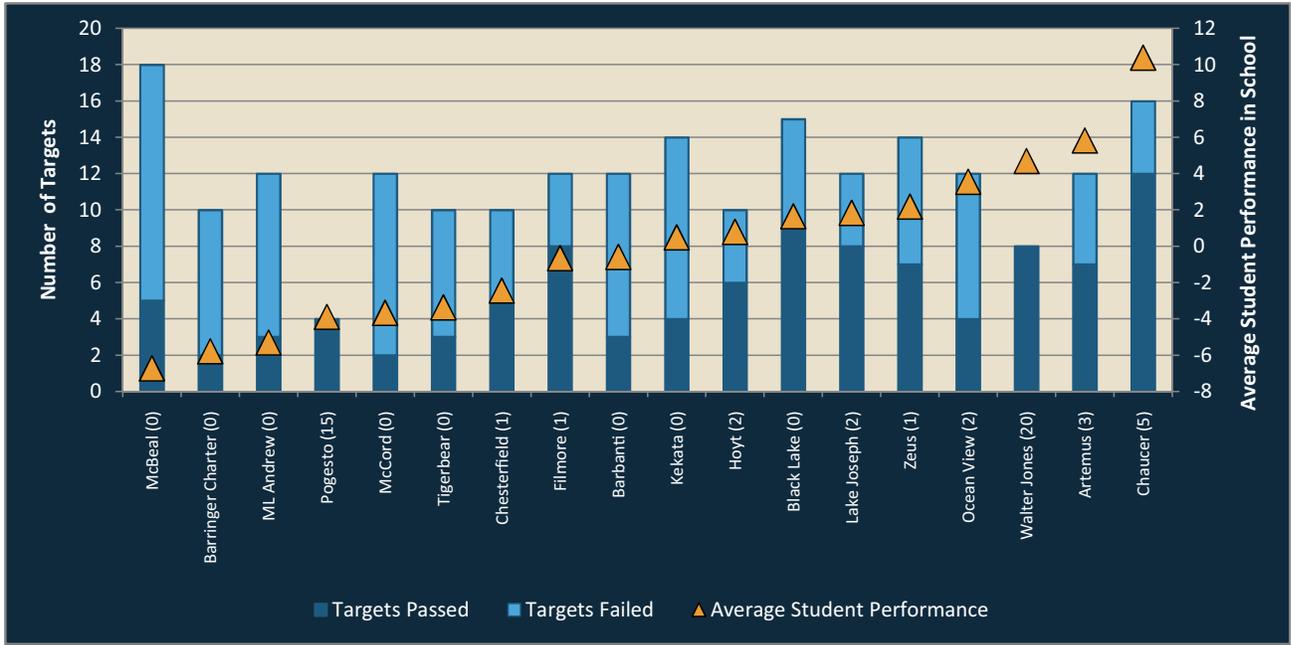


Figure 4. AYP performance of the middle school sample under the Indiana 2008 AYP rules

Note: This figure shows how each of the middle schools within the sample fared under Indiana’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 AMO for even a single subgroup didn’t make AYP, so any light blue means the school failed to make AYP. Chaucer Middle School, for example, meets 12 of its 16 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.

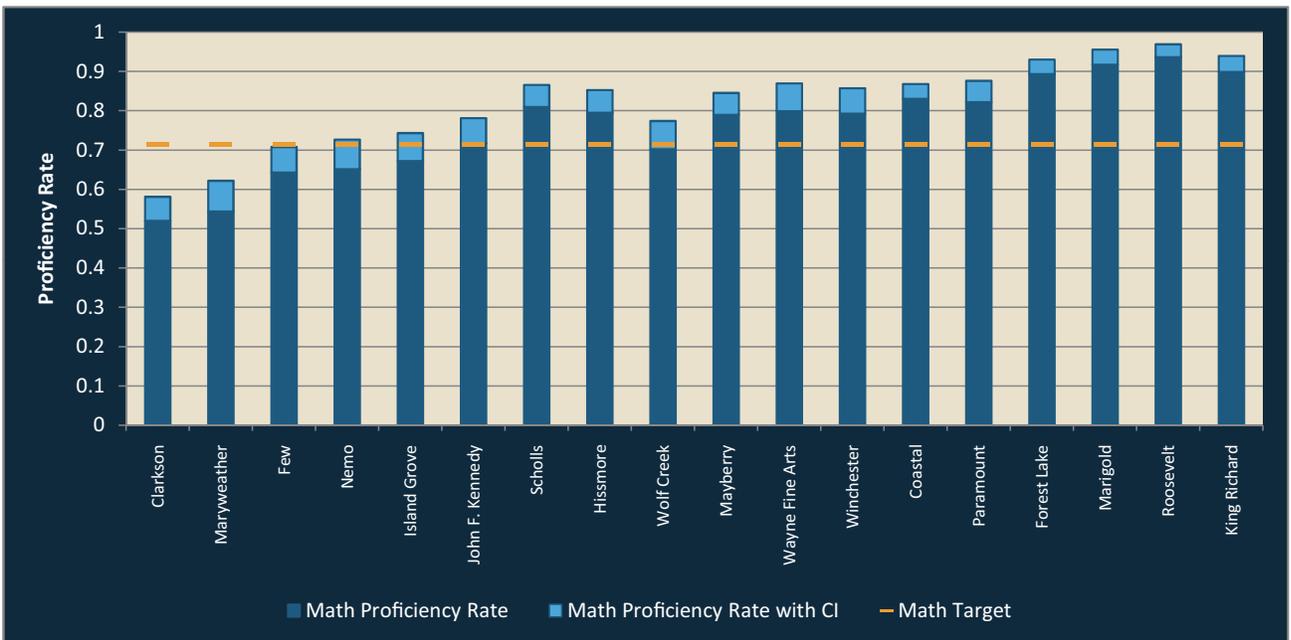


Figure 5. Impact of the confidence interval on elementary school math proficiency rates

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 elementary schools (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.

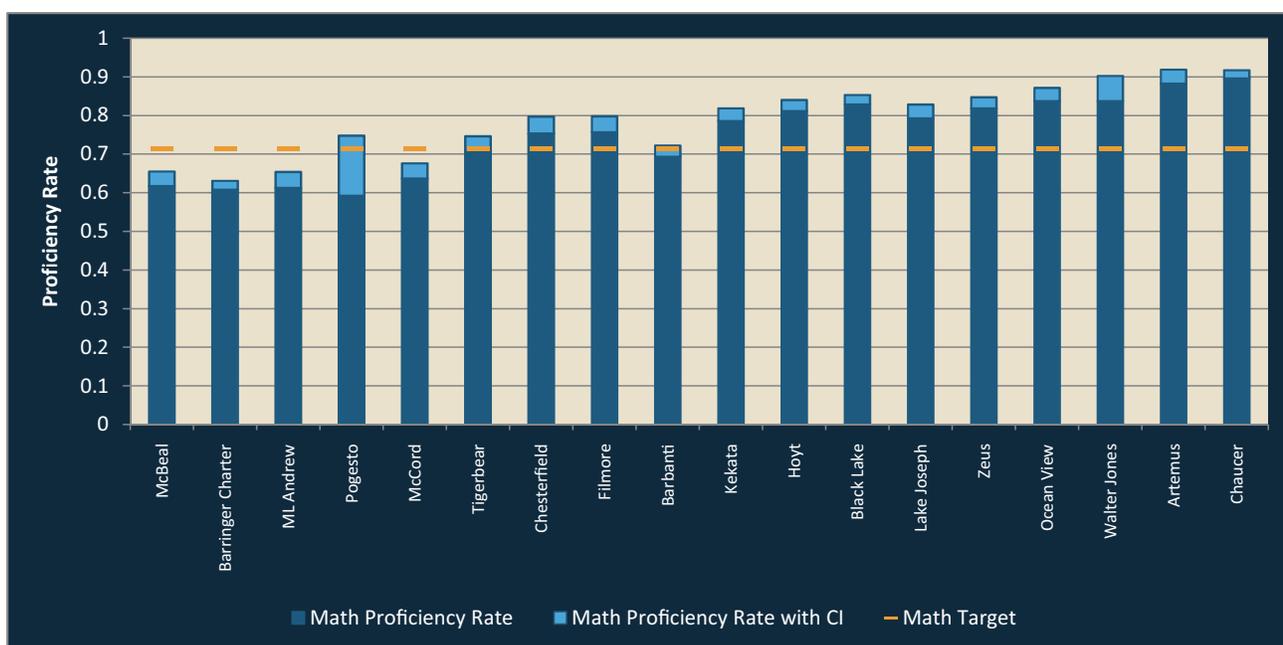


Figure 6. Impact of the confidence interval on middle school math proficiency rates

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.

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 Indiana’s AYP Rules?

Figure 3 illustrates the AYP performance of the sample elementary schools under Indiana’s 2008 AYP rules. **Only 2 elementary schools out of 18 made AYP.** 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 two schools that made AYP are in the right half of the figure, meaning that the highest average performing students were found at these schools.

But among the schools in the right half of the figure, the ones that made AYP are those with relatively few qualifying subgroups—and thus the fewest targets to meet

(since each subgroup has its own separate targets). For example, Wayne Fine Arts and Roosevelt made AYP, but had only eight and six targets each, respectively. Each school must make AYP for its overall student population in reading and math (two targets), for its low-income students (two targets), and for its white population (two more targets). Wayne Fine Arts also has to make AYP for its African American population (two targets).

Figure 4 illustrates the AYP performance of the sample middle schools under the 2008 Indiana AYP rules. **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 have relatively few qualifying subgroups.

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 are increased by the application of the confidence interval. The orange lines show the annual measurable objective (or annual

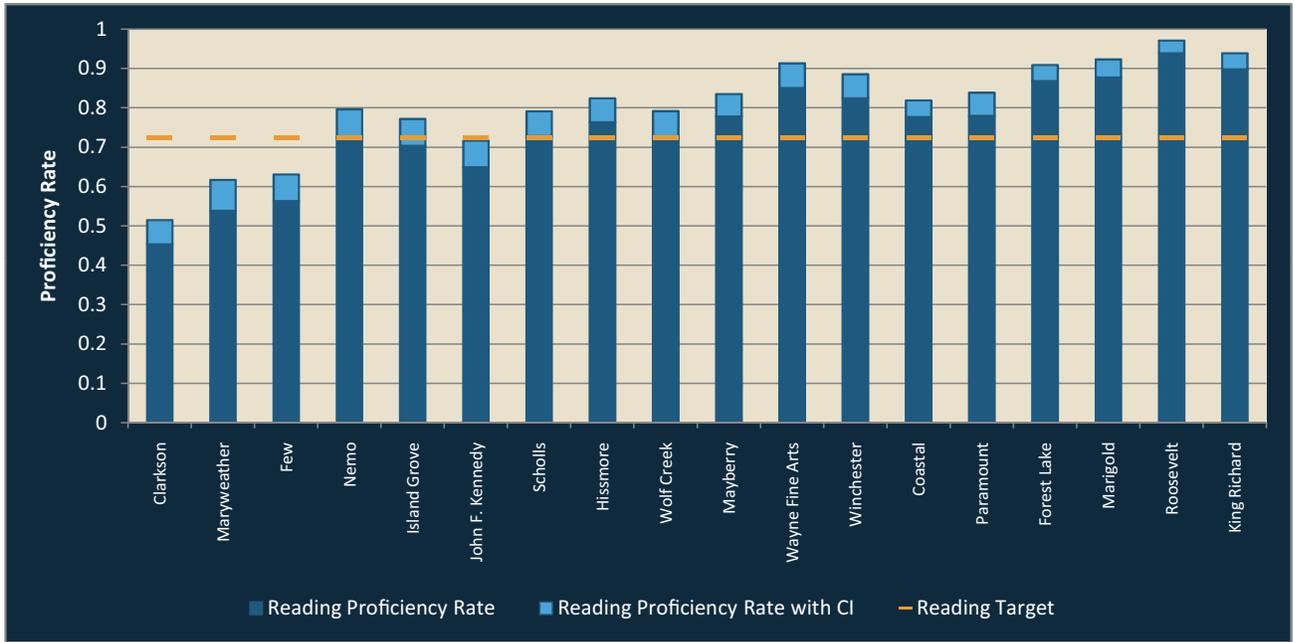


Figure 7. Impact of the confidence interval on elementary school reading proficiency rates

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 (Island Grove) 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.

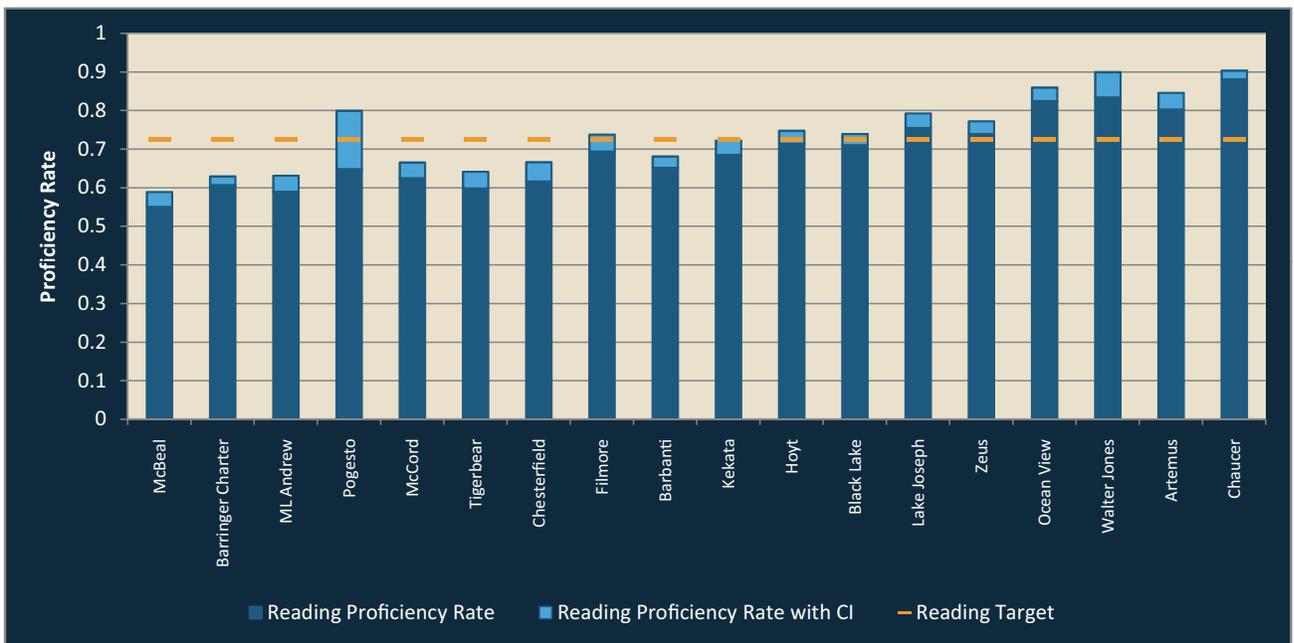


Figure 8. Impact of the confidence interval on middle school reading proficiency rates

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 (Pogesto and Filmore) 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 Indiana AYP

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	52.0%	45.4%	N	N	N	N	N	N	N	N					N	N					10	0	0%	N	1
Maryweather	54.3%	53.9%	N	N	N	N	N	N	N	N					N	N			Y	Y	12	2	17%	N	1
Few	64.3%	56.4%	N	N	N	N	N	N	N	N					Y	N			Y	N	12	2	17%	N	1
Nemo	65.1%	72.6%	Y	Y					N	N									Y	Y	6	4	67%	N	7
Island Grove	67.2%	70.4%	Y	Y				N	N	Y					N	N			Y	Y	9	5	56%	N	4
JFK	71.7%	64.9%	Y	N	N	N			Y	N	N	N							Y	Y	10	4	40%	N	3
Scholls	81.0%	72.9%	Y	Y	N	N			Y	Y	Y	N							Y	Y	10	7	70%	N	7
Hissmore	79.5%	76.3%	Y	Y	N	N			Y	Y	Y	Y							Y	Y	10	8	80%	N	7
Wolf Creek	70.2%	72.1%	Y	Y	N	N			N	N	N				N	N			Y	Y	11	4	36%	N	5
Alice Mayberry	79.0%	77.9%	Y	Y	N	N			Y	Y	Y	N							Y	Y	10	7	70%	N	9
Wayne Fine Arts	79.9%	85.1%	Y	Y					Y	Y	Y	Y							Y	Y	8	8	100%	Y	21
Winchester	79.2%	82.5%	Y	Y	Y	N									Y	Y			Y	Y	8	7	88%	N	22
Coastal	83.0%	77.7%	Y	Y	N	N	Y	N	Y	Y	Y	N			Y	Y			Y	Y	14	10	71%	N	3
Paramount	82.2%	78.0%	Y	Y					Y	N					Y	N			Y	Y	8	6	75%	N	7
Forest Lake	89.3%	86.8%	Y	Y	N	N			Y	Y									Y	Y	8	6	75%	N	8
Marigold	91.7%	87.7%	Y	Y	Y	N			Y	Y									Y	Y	8	7	88%	N	10
Roosevelt	93.6%	93.9%	Y	Y					Y	Y									Y	Y	6	6	100%	Y	28
King Richard	89.9%	89.8%	Y	Y	Y	N	Y	N	Y	N					Y	N			Y	Y	12	8	67%	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 (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.

target) needed to meet AYP. In math, two elementary schools (Nemo and Island Grove) and one middle school (Pogesto) met the overall student population target with the confidence interval, although we know from Figure 3 that Nemo and Island Grove still failed to meet targets for some of their subgroups.

Figures 7 and 8 show the effect of confidence intervals

on the reading proficiency rates for elementary and middle schools, respectively. One elementary school (Island Grove) and two middle schools (Pogesto and Filmore) met the overall reading targets through application of the confidence interval. **Overall, the application of the confidence interval has a moderate effect on whether sample schools met their overall targets in Indiana (or whether they make AYP).**⁹

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 through 8. However, we chose not to show how the confidence interval impacted subgroup performance because it would have added greatly to the report's complexity and length.

Table 3. Middle school subgroup performance of sample schools under the 2008 Indiana 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	61.7%	55.0%	N	N	N	N	N	N	N	N	N	N	N	Y	Y	N	N	N	Y	Y	Y	18	5	28%	N	0
Barringer Charter	60.8%	60.7%	N	N	N	N			N	N	N	N			Y	Y					10	2	20%	N	0	
ML Andrew	61.3%	59.0%	N	N	N	N			N	N	N	N			Y	N			Y	Y	12	3	25%	N	0	
Pogesto	59.3%	64.8%	Y	Y															Y	Y	4	4	100%	Y	15	
McCord Charter	63.7%	62.5%	N	N	N	N			N	N	N	N			N	N			Y	Y	12	2	17%	N	0	
Tigerbear	70.6%	59.8%	Y	N	N	N			N	N	N	N							Y	Y	10	3	30%	N	0	
Chesterfield	75.3%	61.7%	Y	N	N	N			Y	N	Y	N							Y	Y	10	5	50%	N	1	
Filmore	75.7%	69.4%	Y	Y	N	N	Y	N	Y	N					Y	Y			Y	Y	12	8	67%	N	1	
Barbanti	69.4%	65.2%	Y	N	N	N	N	N	N	N					N	N			Y	Y	12	3	25%	N	0	
Kekata	78.6%	68.5%	Y	N	N	N	N	N	Y	N	N	N			N	N			Y	Y	14	4	29%	N	0	
Hoyt	81.1%	71.4%	Y	Y	N	N			Y	N	Y	N							Y	Y	10	6	60%	N	2	
Black Lake	82.9%	71.0%	Y	Y	N	N	N		Y	N	N	N	Y	Y	Y	Y			Y	Y	15	9	60%	N	0	
Lake Joseph	79.3%	75.5%	Y	Y	N	N	N	N	Y	Y					Y	Y			Y	Y	12	8	67%	N	2	
Zeus	81.8%	73.9%	Y	Y	N	N	N	N	Y	N	Y	Y			N	N			Y	Y	14	7	50%	N	1	
Ocean View	83.7%	82.4%	Y	Y	N	N	N	N	N	N					N	N			Y	Y	12	4	33%	N	2	
Walter Jones	83.7%	83.4%	Y	Y					Y	Y					Y	Y			Y	Y	8	8	100%	Y	20	
Artemus	88.3%	80.3%	Y	Y	N	N			Y	N			Y	Y	N	N			Y	Y	12	7	58%	N	3	
Chaucer	89.6%	88.1%	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.

Where Do Schools Fail?

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

Table 4. Summary of subgroup performance of sample elementary schools under the 2008 Indiana 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	13	10	13
Students with limited English proficiency	5	3	7
Low-income students	17	6	8
African-American students	6	1	4
Asian/Pacific Islander students	0	0	0
Hispanic students	9	4	7
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 Indiana 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	16	16
Students with limited English proficiency	9	8	8
Low-income students	17	7	14
African-American students	11	7	9
Asian/Pacific Islander students	4	0	0
Hispanic students	14	7	8
American Indian/Alaska Native students	1	1	0
White students	17	0	0

- Four elementary schools (Clarkson, Maryweather, Few and JFK) failed to meet reading targets for their overall school population and three of these elementary schools (Clarkson, Maryweather, and Few) also failed to meet targets in math.
- Four middle schools (McBeal, Barringer, ML Andrew, and McCord) failed to meet both reading and math targets for overall populations, and four mid-

dle schools (Tigerbear, Chesterfield, Barbanti, and Kekata) failed to meet overall reading targets.

- Three of the 16 failing elementary schools (Hissmore, Winchester, and Forest Lake) did not make AYP because of a single subgroup (SWDs).

Tables 4 and 5 summarize subgroup performance for elementary and middle schools, respectively. First, the per-

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

	Elementary Schools		Middle Schools	
	Made AYP	Failed to make AYP	Made AYP	Failed to make AYP
Number of schools in sample	2	16	2	16
Average student body size	243	312	124	951
Average % low income	18	50	42	45
Average % nonwhite	25	43	27	46
Average performance [†]	5.61	0.68	0.40	-0.11
Average % growth [‡]	100	117	109	97
Average number of targets to meet	7	10	6	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.

formance of SWDs is proving most challenging for schools under Indiana's system, particularly in middle schools, where this subgroup tends to have enough students to meet the state's minimum *n* of 30. In fact, all but three elementary schools and all of the middle schools in the study with qualifying SWD subgroups failed to make targets in math and all schools with such subgroups failed in reading. Students with LEP are also struggling to meet the state's targets; all schools with a large enough LEP population to qualify as a separate subgroup failed to meet 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 Indiana's NCLB accountability system is, in most respects, behaving like those in other states. For example, Roosevelt and Wayne Fine Arts are among those schools that made AYP in the greatest number of states—28 and 21, respectively. And these schools made AYP in Indiana, too. Likewise, the elementary and middle schools that failed to make AYP in the greatest number of states also failed to make AYP in Indiana.

But Indiana is also home to a few anomalies. Consider Pogesto Middle School (Figure 4). Even with its relatively low average performance it made AYP in Indiana, but failed to do so in 13 of 28 states. Its AYP success in Indiana is likely attributable to the relatively small number of targets (four) it had to meet (as shown in Table 3). **In addition, Indiana has relatively easy proficiency standards, compared to other states, and a lenient confidence interval.** A second anomaly is apparent with Winchester Elementary, which made AYP in most of the states examined, but failed to make AYP in Indiana because of its SWD subgroup. This may be because Indiana uses a smaller minimum subgroup size than most other states, meaning that **schools in Indiana are accountable for more subgroups than similar schools in other states.**

This is consistent with the patterns shown in Table 6, which compares the sample schools that did and didn't make AYP on a number of academic and demographic dimensions. Within the sample, schools that made AYP do indeed show higher average student performance, but they also differ in the following ways: they have much smaller student populations, fewer subgroups (and thus

fewer targets to meet), and much lower percentages of traditionally academically disadvantaged (e.g., low-income) students. Similarly, middle schools that made AYP have slightly higher performing students, on average, than middle schools that failed to make AYP, but have far smaller total enrollments, smaller nonwhite populations, and fewer subgroups (and thus targets to meet).

Concluding Observations

The 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 Indiana's AYP rules (and AMOs) for 2008. We found that only 2 elementary schools and 2 middle schools—4 out of a sample of 36—made AYP in Indiana. Looking across the 28 state accountability systems examined in the study, this puts Indiana at the low end of the distribution in terms of the numbers of schools making AYP, as shown in Figure 1. **Though Indiana has relatively easy proficiency standards, it also uses fairly ambitious annual targets, and a smaller minimum subgroup size than most other states, meaning that schools in Indiana are accountable for more subgroups than similar schools in other states. All of these factors potentially inhibit the chances of a school making AYP in the Hoosier State.**

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, Indiana'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. Almost all of the sample schools made AYP in California for their student populations as a whole (i.e., without considering subgroup results). 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 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 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? Even if actual participation guidelines for English language learners and SWDs are more generous under the current state assessment system¹⁰ doesn't the massive failure of these students to meet Indiana'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 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.

¹⁰ See footnote 5.

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

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