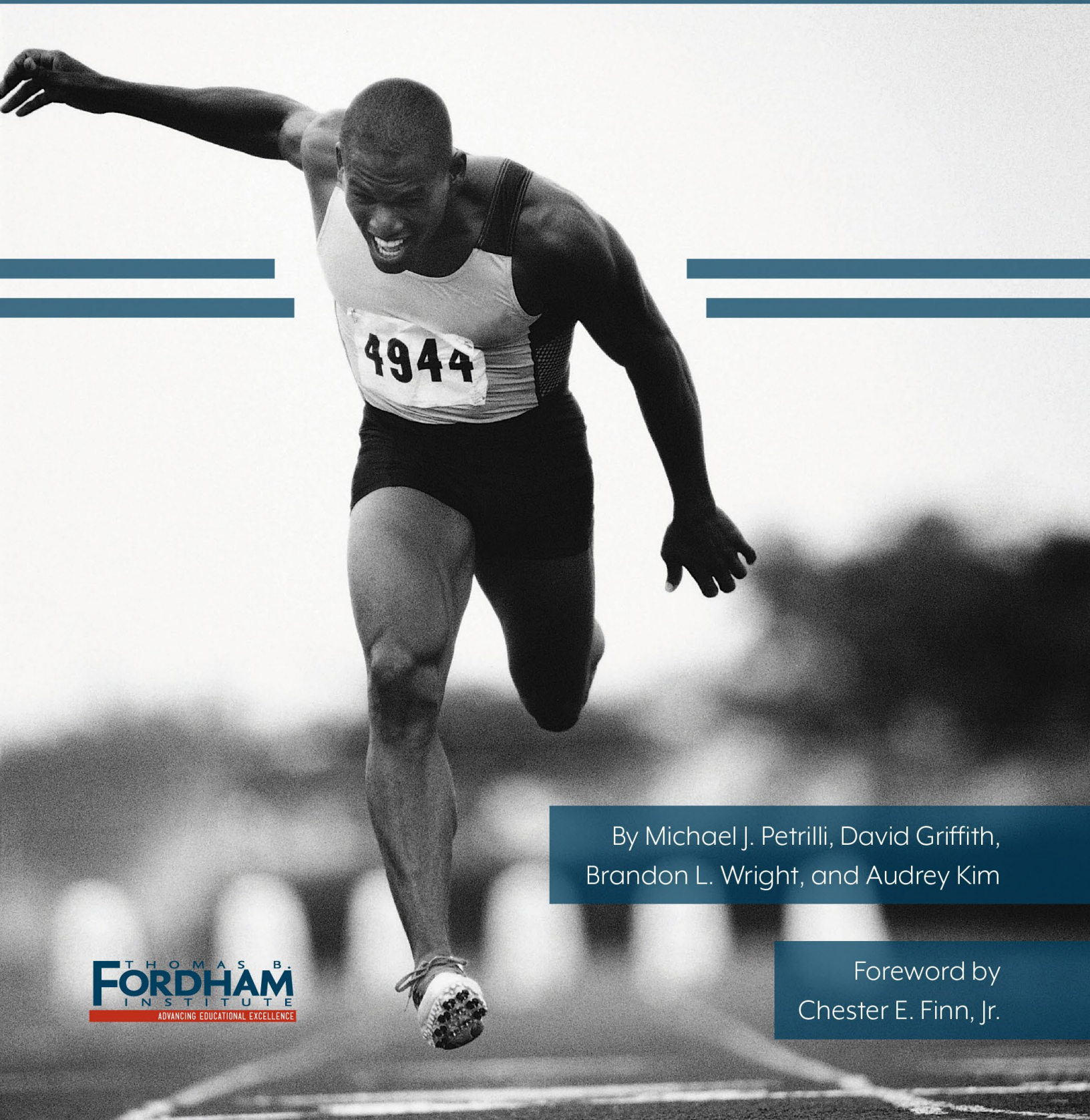
AUGUST 2016

HIGH STAKES FOR HIGH ACHIEVERS:

STATE ACCOUNTABILITY IN THE AGE OF ESSA



By Michael J. Petrilli, David Griffith,
Brandon L. Wright, and Audrey Kim

Foreword by
Chester E. Finn, Jr.

THOMAS B. FORDHAM INSTITUTE
ADVANCING EDUCATIONAL EXCELLENCE

The Thomas B. Fordham Institute is the nation's leader in advancing educational excellence for every child through quality research, analysis, and commentary, as well as on-the-ground action and advocacy in Ohio. It is affiliated with the Thomas B. Fordham Foundation, and this publication is a joint project of the Foundation and the Institute. For further information, please visit our website at www.edexcellence.net or write to the Institute at 1016 16th St. NW, 8th Floor, Washington, D.C. 20036. The Institute is neither connected with nor sponsored by Fordham University.

Contents

Foreword

By Chester E. Finn, Jr.

Accountability has been a central theme of U.S. education reform for almost two decades, driven by the unchallenged central finding of James Coleman's seminal 1966 study: Although some programs are demonstrably more effective than others, there's no direct link between what goes into a school by way of resources and what comes out by way of student learning. Sage policy makers have recognized that instead of trying to micromanage school and district "inputs," they should clearly state the results they want their educational institutions to produce, assess how satisfactorily those results are being achieved, and then hold schools and school systems to account, with rewards of various sorts for success and interventions of various sorts in the event of institutional failure.

This strategy has worked fairly well. In particular, after years of stagnation in the late 1980s and early 1990s, achievement began to rise again in the late 1990s—particularly in the earlier grades and most notably in math—as states set new academic standards, started testing their students regularly, and installed their own versions of "consequential accountability" systems. Once No Child Left Behind (NCLB) made this reform regime inescapable, "late adopter" states—those jurisdictions that hadn't already moved in this direction on their own—also started to see gains. Rigorous studies have shown that accountability deserves at least some of the credit for these improvements, which is not too surprising, considering that just about every person and institution does a little better at any number of undertakings when consequences follow from success and failure.[1]

So far so good. Yet we must not gloss over critical details. Early proponents of accountability in public education tended to speak in generalities; it was said, for example, that we needed to hold schools accountable for "raising student achievement." But whose achievement? All students? In which subjects? Measured how?

NCLB provided its own answers to these questions. Schools would be held to account for getting increasing proportions of their students, and increasing proportions of key subgroups, to "proficiency" in reading and math. States would define "proficiency" as they saw fit, but they would eventually need to sanction any school that didn't raise all of its students to that level.

Faced with these requirements, most states did the rational thing and set the proficiency bar low.[2] And that move, combined with NCLB's mandatory cascade of sanctions, created a powerful incentive for schools to pay close attention to students below the proficiency bar. Conversely, there was absolutely no incentive to worry about the achievement of those who had already reached, or were likely to reach, that bar. To put it bluntly, NCLB did some good for America's struggling pupils, but for high achievers, it mostly just hit the education pause button.

Research has demonstrated that students just below the bar were most likely to make large gains in the NCLB era, while high achievers made lesser gains.[3] Those most victimized by this regime were high-achieving poor and minority students—kids who were dependent on the school system to cultivate their potential and accelerate their achievement.[4] (Equally able youngsters from middle-class circumstances have other people and educational resources to keep them moving forward.)

The good news is that accountability works: Districts, schools, and educators do respond to its incentives and disincentives. The bad news is that kids can get left high and dry when policy makers incent schools to ignore their needs.

WHY FOCUS ON HIGH ACHIEVERS?

Many education reformers look at results for the National Assessment of Educational Progress (NAEP) and other macro-measures and see some positive trend lines in recent decades. Gaps are indeed closing, especially between low- and high-achieving students. Isn't that what we want?

Yes, of course—up to a point. Historically our K–12 system has done the greatest harm to our lowest-performing students, who tend to come from poor and minority families. Therefore, using accountability (as well as school choice and other strategies) to improve matters for disadvantaged youngsters has been and should remain a policy focus.

But it should not be the only focus. The policy challenge going forward is to devise accountability systems that deal with the ceiling as well as the floor. This is partly about fairness. It's wrong for any child to miss out on academic challenges at school, and we should do everything we can to develop the full potential of all our students, including high achievers. We must also remember, though, that the country's future economic competitiveness, scientific leadership, and national security depend on how successfully we maximize the learning of our ablest children. If we want tomorrow's scientists, entrepreneurs, and inventors to "look like America," our schools need to take special pains with the education of high-ability kids from disadvantaged circumstances. They too should have the chance to realize the American Dream.

There's a political argument, too: How can we expect parents to support public education when many of their children aren't a priority for the schools they attend?

And there's a powerful case to be made for accelerating social mobility by educating high-ability, low-income children. These are the poor kids, many of them from minority groups, who have the best chance to succeed in selective universities, become leaders in their communities, and climb the ladder to the middle class. Yet they are also the kids most dependent on the education system to recognize and draw out their potential. Research from Fordham, the Jack Kent Cooke Foundation, and elsewhere shows that these low-income "high flyers" are likeliest to "lose altitude" as they make their way through school.[5] The result is an "excellence gap" rivaling the "achievement gaps" that have been our policy preoccupation.

NCLB-style accountability is partly to blame for that. After all, low-income high achievers are likely to attend high-poverty schools, which face the greatest pressure to raise the test scores of their lowest-performing students and neglect their top pupils. They're also schools that typically face a host of other challenges.

Going forward, policy makers who care about their low-income high achievers should take full advantage of their newfound authority under the Every Student Succeeds Act (ESSA) to ensure that their schools have ample incentives to educate those children, and all children, to the max.

Mindful of both the challenges the country faces and the new opportunity state leaders have to set matters right, the analysis that follows does two things. First, it advances specific ideas for how state accountability systems can be designed to demand strong performance and growth from high-achieving students while meeting the requirements set forth in ESSA. Second, it rates current (or proposed) accountability systems in the fifty states and the District of Columbia based

on how well they draw attention to high achievers. The evidence, regrettably, is that few of them are doing it well. Which is to say, the problem is sizable, but the opportunity to solve it is at hand.

In an unusual move for Fordham, our own talented research and policy team completed this analysis in-house. Kudos are owed to co-authors David Griffith, Audrey Kim, Mike Petrilli, and Brandon Wright for rolling up their sleeves and seeing this project through to completion. This quartet was responsible for all phases of the study: developing the metric, collecting and analyzing the data, and summarizing the findings.

More than most Fordham publications, this one is motivated by an explicit desire to influence policy makers in the short term. We're mindful that much of what we unearthed about state accountability systems could be out of date within a year's time. But that same year offers state leaders a rare opportunity to do things differently and better. Many issues will be debated as states design their new accountability systems. Our hope is that the educational needs of high-achieving students get the attention they deserve—and that they didn't get in the NCLB era.

Let us say to educators and policy makers who are already retooling their state accountability systems: Those children are counting on you. Their futures depend in no small part on the decisions you are making.

ACKNOWLEDGMENTS

This report was made possible through the generous support of the Lynde and Harry Bradley Foundation, Bloomberg Philanthropies, and our sister organization, the Thomas B. Fordham Foundation. We are grateful to interns Daniel Cohen, Andrew Scanlan, Kate Stringer, and Darien Wynn for their research assistance, to Alyssa Schwenk for her role in disseminating the final product, to Jonathan Lutton for the beautiful layout, and to Kevin Mahnken for copy editing. We also thank the many individuals who helped ensure that the information contained in this report was as timely and accurate as possible, including our local respondents and reviewers. We are particularly appreciative of the officials in state departments of education who took the time to review drafts and verify that we had obtained the most current version of their accountability systems. Any errors are ours alone.

Executive Summary

In this report, we examine the extent to which states' current (or planned) accountability systems for elementary and middle schools attend to the needs of high-achieving students, and how these systems might be redesigned under the Every Student Succeeds Act to better serve all students. (A forthcoming analysis will examine accountability for high schools.)

In our view, states can and should take four steps to ensure that the needs of high achievers are prioritized under ESSA:

1. **For the first academic indicator required by ESSA ("academic achievement"), reward schools for getting more students to an "advanced" level.**

2. **For the second academic indicator expected by ESSA ("student growth"), rate schools using a "true growth model," i.e., one that looks at the progress of individual students at all achievement levels and not just those who are low-performing or below the "proficient" line.**

3. **Include "gifted students" (or "high-achieving students") as a subgroup, and report their results separately.**

4. **When determining summative school ratings, make "growth for all students" count for at least half of the rating.**

Based on these four design features, we rate states' current (or planned) accountability systems using the rubric below and the most recent publicly available information.

TABLE ES-1: RUBRIC FOR RATING STATE ACCOUNTABILITY SYSTEMS

INDICATOR	*RATING*
1. Does the state rate schools' "academic achievement" using a model that gives additional credit for students achieving at an "advanced" level?	Yes / No
2. Does the state rate schools' growth using a model that looks at the progress of all individual students, not just those below the "proficient" line?	Yes / No
3. Does the state's accountability system include "gifted students," "high-achieving students," or the like as a subgroup and report their results separately?	Yes / No
4. When calculating summative school ratings, does "growth for all students" count for at least half of the rating?	Yes / No / NA*
Total number of stars possible	A maximum of 3 or 4 stars

** State doesn't calculate summative school ratings*

This rubric is the basis for two sets of ratings: one for the thirty-nine states (plus the District of Columbia) that calculate summative school ratings (or intend to) and one for the eleven states that don't take this step (or don't plan to).

TABLE ES-2: RESULTS FOR STATES WITH NO SUMMATIVE SCHOOL RATINGS

★★★	Ohio
★★☆	South Carolina
★☆☆	Illinois, Kansas, New Jersey, Tennessee
☆☆☆	California, Maryland, Montana, New York, North Dakota

TABLE ES-3: RESULTS FOR STATES WITH SUMMATIVE SCHOOL RATINGS

★★★★	(None)
★★★☆	Arkansas, Oregon
★★☆☆	Colorado, Connecticut, Georgia, Idaho, Kentucky, Massachusetts, Missouri, Nebraska, Nevada, North Carolina, Pennsylvania, Rhode Island, Wisconsin, Wyoming
★☆☆☆	Alabama, Alaska, Arizona, Delaware, District of Columbia, Florida, Hawaii, Indiana, Iowa, Louisiana, Maine, Minnesota, Mississippi, New Hampshire, New Mexico, Texas, Utah, Washington, West Virginia
☆☆☆☆	Michigan, Oklahoma, South Dakota, Vermont, Virginia

As these ratings suggest, the overwhelming majority of current (and planned) state accountability systems provide schools with few incentives to focus on their high-achieving students. In fact, our analysis indicates that just four states—**Arkansas, Ohio, Oregon, and South Carolina**—have truly praiseworthy systems when it comes to focusing attention on these students.

Our results also highlight the specific areas where states need to improve:

» Only four states (Arkansas, Colorado, Idaho, and Oregon) base at least half of a school's rating on "growth for all students," and seven states and the District of Columbia assign no weight to this measure. (Eleven states don't calculate summative school ratings.) Given that student growth is the best way to evaluate schools' impact on student achievement—and the best way to signal that all kids matter—this finding is extremely alarming.

» Just five states (Nevada, North Carolina, Ohio, Oregon, and Wyoming) include high-achieving or gifted students as a subgroup and separately report their results.

» Fourteen states and the District of Columbia rate (or plan to rate) schools' achievement using a model that gives extra credit for students who achieve at an "advanced" level, such as a performance index.

Unfortunately, draft regulations published by the Department of Education appear to disallow such indices, and those fourteen states may be required to resume measuring academic achievement via proficiency rates alone. That's a shame, as research suggests that measuring school quality via proficiency rates is a deeply flawed approach that encourages principals and teachers to narrowly focus attention on students performing just above or below the proficiency line.[6]

For this reason, we have one major recommendation for the Department of Education:

ALLOW STATES TO RATE ACADEMIC ACHIEVEMENT USING A PERFORMANCE INDEX.

Such an allowance is both consistent with ESSA and in the best interests of students. Rather than once again encouraging schools to focus on "bubble kids" as they did under NCLB, the department's final regulations should allow—or, better yet, encourage—performance metrics that account for the achievement of all students.

Introduction

The Every Student Succeeds Act (ESSA) grants states more authority over their school accountability systems than its predecessor, No Child Left Behind (NCLB). Consequently, states now have an opportunity to design school rating systems that improve upon the NCLB model, especially when it comes to high achievers.

NCLB meant well (as did many state accountability systems that preceded it), but it had a pernicious flaw. Namely, it created strong incentives for schools to focus all their energy on helping low-performing students get over a modest "proficiency" bar, while ignoring the educational needs of high achievers, who were likely to pass state reading and math tests regardless of what happened in the classroom. This may be why the United States has seen significant achievement growth for its lowest-performing students over the last twenty years but smaller gains for its top students.

Starting in 2011, former Secretary of Education Arne Duncan offered waivers to states that wanted the flexibility to redesign their accountability systems. In particular, states were allowed to incorporate the use of real student growth measures into their school determinations. This was important for a variety of reasons. First, growth measures do a better job of capturing schools' impact on student achievement than proficiency rates, which are strongly correlated with student demographics, family circumstance, and prior achievement. And just as significantly, well-designed growth measures can eliminate the temptation for schools to ignore their high achievers.

ESSA maintains NCLB's requirement that states assess students annually in grades 3–8 and once in high school, as well as the mandate that states adopt accountability systems that lead to ratings for schools. These systems must include four types of indicators: academic achievement; another academic indicator, which can include student growth for elementary and middle schools; growth toward English proficiency for English language learners; and at least one other valid, reliable indicator of school quality or student success. Each of the academic indicators (1–3) must carry "substantial" weight and, in the aggregate, count "much more" than the fourth.

Here we examine whether each state's accountability system prioritizes high achievers. We do not examine the quality of their standards, tests, or sanctions for low performance. (See *Important Issues Beyond the Scope of this Analysis*.)

This analysis also illustrates how states can seize the opportunity under ESSA to redesign their accountability systems to prioritize high achievers.

This last point is especially important because many state accountability systems are currently in flux. In part, that's because of recent changes allowed by ESEA waivers, as well as the coming changes driven by ESSA implementation. But it's also because states across the country recently moved to new, tougher assessments linked to their new, tougher standards.

States may think we're being premature in evaluating their systems during this time of massive change. Please understand that our primary objective is to identify the design features of an accountability system that does right by high achievers—which we hope will become the prevailing model now that ESEA is reauthorized and states' testing regimes are becoming stable once again.

Our focus here is on rating systems for elementary and middle schools. A separate report will examine the same issues for high school accountability.

IMPORTANT ISSUES BEYOND THE SCOPE OF THIS ANALYSIS

In addition to browsing through this report, we encourage readers to spend time with the Jack Kent Cooke Foundation's fifty-state report card on closing the excellence gap, which offers a comprehensive look at the variety of state policies that can support high-achieving students.[7] After all, the four design features examined here do not encompass everything that states could be doing to encourage schools to serve their high-achieving students well. Nor does our analysis capture all of the critical elements of a state accountability system as they pertain to high-achieving students. Most notably, we do not consider the content standards and tests that states have adopted, both of which are worth some discussion.

The foundation of any well-designed accountability system is a set of clear, demanding academic standards like the Common Core State Standards for English and math, which are still in place in more than forty states (despite the political backlash against them). As readers likely know, the Fordham Institute has been a staunch defender of these standards, which we've found to be stronger—in substance, in rigor, and in clarity—than what three-quarters of the states had in place before their adoption, and on par with the rest. Yet we've also warned that they should not be used as an excuse to eliminate services for the nation's academic superstars. (See our white paper, written by Jonathan Plucker, *Common Core and America's High-Achieving Students*.) While the Common Core standards aim higher than most of the expectations that came before them, they still don't aim high enough for the country's top students. No standards could. Consequently, we've excluded an evaluation of state content standards from this analysis.

The quality of state assessments matters enormously too. And here we wish we could collect data, especially about the capacity of state tests to accurately measure the performance and growth of students who are well above grade level (i.e., whether they contain enough hard questions to capture growth at the high end). Unfortunately, to our knowledge, such data do not exist. Furthermore, a provision of NCLB requiring that all students take the "same tests" was interpreted by both the George W. Bush and Barack Obama administrations as requiring "on-grade-level" testing, effectively prohibiting states from building tests that were accurate for students well above (or below) grade level.

Though the intent of that decision was pure—it prevents states from setting lower expectations for, and administering easier tests to, low-performing kids—it has curtailed the use of computer-adaptive testing and other strategies for accurately measuring performance at the top of the achievement distribution. Consequently, even the new Smarter Balanced assessments, which are computer-adaptive, have been unable to precisely measure the achievement of students well above grade level.

Thankfully, ESSA eliminates this federal hurdle by giving explicit congressional approval to truly adaptive testing (both above and below grade level) as long as students are tested on grade-level items as well. We hope that Smarter Balanced states move expeditiously to take advantage of this new flexibility—and that other states also transition to adaptive tests.

Methods

In our view, states can and should take four steps to ensure that the needs of high achievers are prioritized under ESSA:

1. **For the first academic indicator required by ESSA ("academic achievement"), give schools incentives for getting more students to an "advanced" level.** Under ESSA, states will continue to track the percentage of students who attain proficiency on state tests. They should also give schools incentives for getting students to an advanced level (such as Level Four on Smarter Balanced or Level Five on PARCC). For example, they might create an achievement index that gives schools partial credit for getting students to "basic," full credit for getting students to "proficient," and additional credit for getting students to "advanced." (It's not entirely clear from the Department of Education's proposed regulations whether this will be allowed, though we don't see anything in the law prohibiting it.)

2. **For the second academic indicator expected by ESSA (student growth), rate schools using a "true growth model," i.e., one that looks at the progress of individual students at all achievement levels and not just those who are low-performing or below the "proficient" line.** Regrettably, some states still don't consider individual student growth, or else they use a "growth-to-proficiency system" that continues to encourage schools to ignore the needs of students above (or far above) the proficient level. Using true growth models—such as "value added" or the "growth percentile method"—for all students is much preferred.

3. **Include "gifted students" (or "high-achieving students") as a subgroup in the state's accountability system and report results for them separately.** States can signal that high achievers matter by making them a visible, trackable "subgroup," akin to special education students or English language learners, and publishing school ratings for their progress and/or achievement. (Obviously, it makes little sense to simply report that high achievers are high-achieving. But whether they are making strong growth is quite relevant. Alternatively, states might publish results for students labeled as "gifted," though that opens up a can of worms about how that label is applied.)

4. **When determining summative school ratings, make growth—across the achievement spectrum—count the most.** Finally, the Department of Education's proposed regulations require states to combine multiple factors into school ratings, probably through an index. Each of the three academic indicators (achievement, growth, and progress toward English proficiency) must carry "substantial" weight. But in our view, states should (and, under ESSA, are free to) make growth matter the most (50 percent or more of a school's total score). Otherwise, schools will continue to face an incentive to ignore their high performers. (States that don't combine their indicators into a summative school rating receive a "Not Applicable" here.)

Based on the four design features listed above, we rated the school accountability systems in the fifty states and the District of Columbia using the rubric shown below and the most recent publicly available information. (See *Data Collection*.) In particular, we looked at report cards for middle schools, as well as state documents explaining the nitty-gritty of how school grades are (or will be) calculated.[8]

TABLE 1: RUBRIC FOR RATING STATE ACCOUNTABILITY SYSTEMS

INDICATOR	RATING
1. Does the state rate schools' "academic achievement" using a model that gives additional credit for students achieving at an "advanced" level?	Yes / No
2. Does the state rate schools' growth using a model that looks at the progress of all individual students, not just those below the "proficient" line?	Yes / No
3. Does the state's accountability system include "gifted students," "high-achieving students," or the like as a subgroup and report their results separately?	Yes / No
4. When calculating summative school ratings, does "growth for all students" count for at least half of the rating?	Yes / No / NA*
Total number of stars possible	A maximum of 3 or 4 stars

** State doesn't calculate summative school ratings*

DATA COLLECTION

The data in this report reflect information that was publicly available as of July 22, 2016. To collect this information, we scanned state department of education websites for accountability-related documents (such as guides to school rating systems) and inspected school report cards to see what information states reported. For the sake of transparency, we include screenshots of some these documents in the exhibits of the state profiles. To ensure that the information we collected was as up-to-date as possible, we gave state officials the opportunity to review their state's profile before publication.

The task of evaluating state accountability systems is complicated by the fact that so many of them are in flux. Consequently, throughout this report we take the following approach: When a state has publicly committed to changes that satisfy the requirements of one of our indicators, we acknowledge that fact by giving it credit for those changes. However, when a state's intent is ambiguous or unclear, we do not give credit. (Thus, since the process of revising a state's accountability system is often a lengthy one, our scores sometimes reflect a mix of states' current and intended systems.)

Results

Our analysis suggests that the overwhelming majority of current (or planned) state accountability systems provide schools with few incentives to focus on their high-achieving students. However, there is a great deal of variation between states.

To get a more nuanced view, it is helpful to distinguish between states that produce summative ratings of school quality and those that do not. As mentioned in previous sections, states could earn a maximum of either three or four stars depending on whether they combined the indicators by which schools are judged into single grades or ratings. Thus, the thirty-nine states (plus the District of Columbia) that assign such ratings could earn a maximum of four stars, while the ten states that don't assign them could earn a maximum of three.

We present the results for both groups of states below, as well as the results for each individual indicator.

States with no summative school ratings (maximum of three stars)

As shown in Table 2, the states that lack summative school ratings do little to encourage schools to focus on their high achievers, with two exceptions: Ohio, which is the only state to earn three out of three stars (and the only state in either group that earns the maximum number of stars available to it), and South Carolina, which is the only state to earn two out of three stars.

TABLE 2: RESULTS FOR STATES WITH NO SUMMATIVE SCHOOL RATINGS

★★★	Ohio
★★☆	South Carolina
★☆☆	Illinois, Kansas, New Jersey, Tennessee
☆☆☆	California, Maryland, Montana, New York, North Dakota

We view Ohio's accountability system as the best in the country for high achievers: It gives schools additional credit for students who achieve at an advanced level; it rates schools' growth using a model that looks at the progress of all students, not just those below proficient; and it includes "gifted" students as a subgroup and reports their results separately. South Carolina's system, which shares all the characteristics of Ohio's except the mandate for a high-achiever subgroup, is also quite good.

Less impressive, however, are the four states in this group that earn one of three stars, which do little to incentivize schools to focus on their brightest students. And worse still are the five states that earn zero stars—California, Maryland, Montana, New York, and North Dakota—by doing nothing to encourage schools on this front. Besides failing to reward advanced achievement and separately report growth for high achievers, these states fail to rate school-level growth altogether.

STATES WITH SUMMATIVE SCHOOL RATINGS (MAXIMUM OF FOUR STARS)

As shown in Table 3, of the thirty-nine states (and the District of Columbia) that assign summative school ratings, none earn the maximum of four stars. And only two—Arkansas and Oregon—earn three stars, and might be considered leaders when it comes to encouraging a focus on high achievers.

TABLE 3: RESULTS FOR STATES WITH SUMMATIVE SCHOOL RATINGS

★★★★	(None)
★★★☆	Arkansas, Oregon
★★☆☆	Colorado, Connecticut, Georgia, Idaho, Kentucky, Massachusetts, Missouri, Nebraska, Nevada, North Carolina, Pennsylvania, Rhode Island, Wisconsin, Wyoming
★☆☆☆	Alabama, Alaska, Arizona, Delaware, District of Columbia, Florida, Hawaii, Indiana, Iowa, Louisiana, Maine, Minnesota, Mississippi, New Hampshire, New Mexico, Texas, Utah, Washington, West Virginia
☆☆☆☆	Michigan, Oklahoma, South Dakota, Vermont, Virginia

Both Arkansas and Oregon use growth models that include high achievers and make "growth for all students" count for half of schools' summative ratings. Still, both states' accountability systems could be improved. For example, Oregon doesn't give additional credit for students who achieve at an "advanced" level, and Arkansas doesn't include "talented and gifted" students as a subgroup or separately report their results.

Similarly, most of the fourteen states that earn two stars out of four include high achievers in their growth models but fall short in other ways. For example, most don't assign much weight to growth or give schools extra credit for students who achieve at an advanced level.

That observation also applies to the eighteen states (plus the District of Columbia) that earn just one star (usually for including high-achieving students in their growth model). These states do a poor job of encouraging schools to focus on their high achievers, and often discourage such a focus.

Finally, five states earn zero stars out of four, meaning they explicitly or implicitly discourage schools from focusing on their brightest students. For example, many base school achievement ratings entirely on proficiency rates, with no additional credit for advanced achievement.

In short, despite ample opportunity to do so over the past few years, most states have largely failed to move beyond the flawed approach to accountability embodied in No Child Left Behind, which placed undue emphasis on proficiency at the expense of students who had already exceeded that standard.

RESULTS FOR INDIVIDUAL INDICATORS

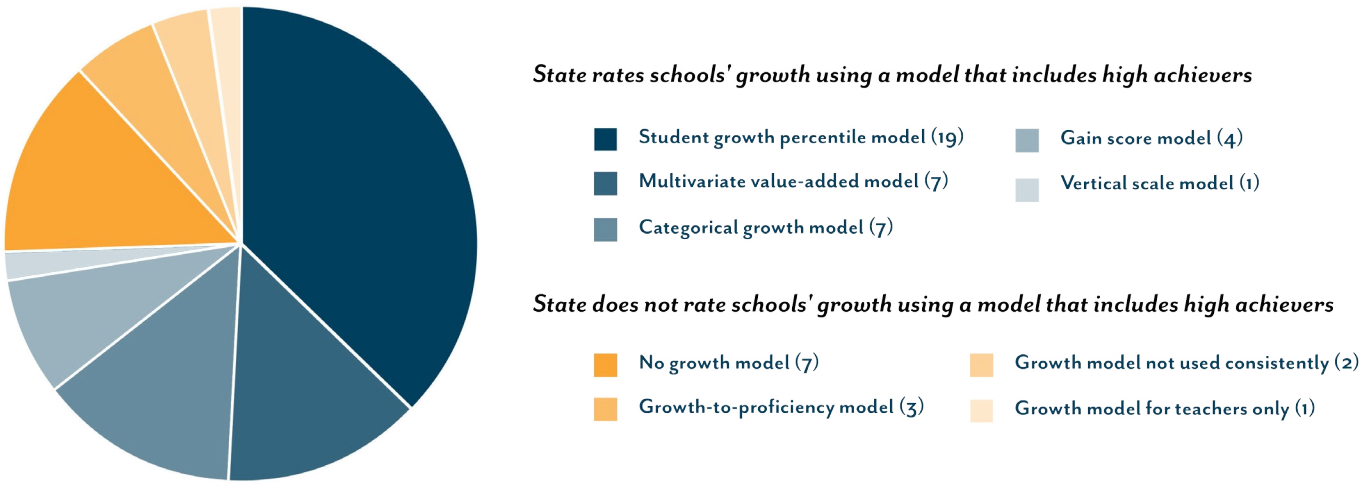
Disaggregating our results by indicator largely confirms our central finding that state accountability systems do little to encourage schools to focus on high achievers. Still, our analysis identifies a few bright spots.

Most states rate schools' growth using a model that includes high-achieving students

Encouragingly, thirty-eight states now rate student growth (at the school level) using a model that includes high achievers, meaning they reward growth beyond proficiency. That number represents real progress from a few years ago, when such an approach was considered unlawful under NCLB. Of those thirty eight states, nineteen use a student growth percentile model, seven use a multivariate value-added model, seven use a categorical growth model, four use a gain score model, and one uses a vertical scale model.[9] (See Figure 1.)

Of the twelve states that don't rate student growth using a model that includes high achievers, three (Louisiana, Oklahoma, and South Dakota) use some form of growth-to-proficiency model, which does nothing to encourage schools to pay attention to students who are already proficient. And two (New York and Virginia) have developed a growth model, but as far as we can tell, don't use it to rate schools' growth.[10] The other seven states, including Alabama, California, Maryland, Michigan, Montana, North Dakota, and Vermont, have yet to develop a growth model (though Alabama and Michigan appear to be moving toward adopting one).

FIGURE 1: MOST STATES RATE SCHOOLS' GROWTH USING A MODEL THAT INCLUDES HIGH ACHIEVERS



Most states don't give additional credit for students who achieve at an advanced level

Fourteen states and the District of Columbia rate (or plan to rate) schools' achievement using a model that gives additional credit for students who achieve at an "advanced" level (meaning that thirty-six states do not). Most of these states use an achievement index that gives schools partial credit for getting students to "basic," full credit for getting students to "proficient," and additional credit for getting students to "advanced" (or something along those lines).[11] Unfortunately, it's unclear from the Department of Education's proposed regulations whether states will be allowed to use such an index as one of their "academic indicators" under ESSA. Obviously we believe that they should be—and that the statute provides plenty of room for such an interpretation.[12] (See *Recommendation for the U.S. Department of Education*.)

Very few states report results for high-achieving students separately

Just five states (Nevada, North Carolina, Ohio, Oregon, and Wyoming) include high-achieving or gifted students as a subgroup and separately report their results at the school level, meaning that parents and other stakeholders in the other forty-five states and the District of Columbia have little information with which to determine how well these students are being served. In a number of states, school report cards include disaggregated results for almost every subgroup that is of interest to policy makers except high achievers, underscoring the degree to which they are not viewed as a priority.

RECOMMENDATION FOR THE U.S. DEPARTMENT OF EDUCATION

As we were repeatedly reminded by state officials while drafting this report, state accountability systems must abide by Uncle Sam's requirements. Thus, the degree to which states can improve these systems in the coming years depends greatly on how the Department of Education views its role under the new law.

In light of these circumstances, we have one major recommendation for the Department of Education:

ALLOW STATES TO RATE ACHIEVEMENT USING A PERFORMANCE INDEX.

ESSA requires the use of an academic achievement indicator that "measures proficiency on the statewide assessments in reading/language arts and mathematics." But there are multiple ways to interpret this. Unfortunately, the department's proposed regulations seem to expect states to use proficiency rates to measure school performance. This is a mistake that will encourage schools to focus on "bubble kids"—those just above or below the proficiency cutoff—exactly as they did under NCLB.

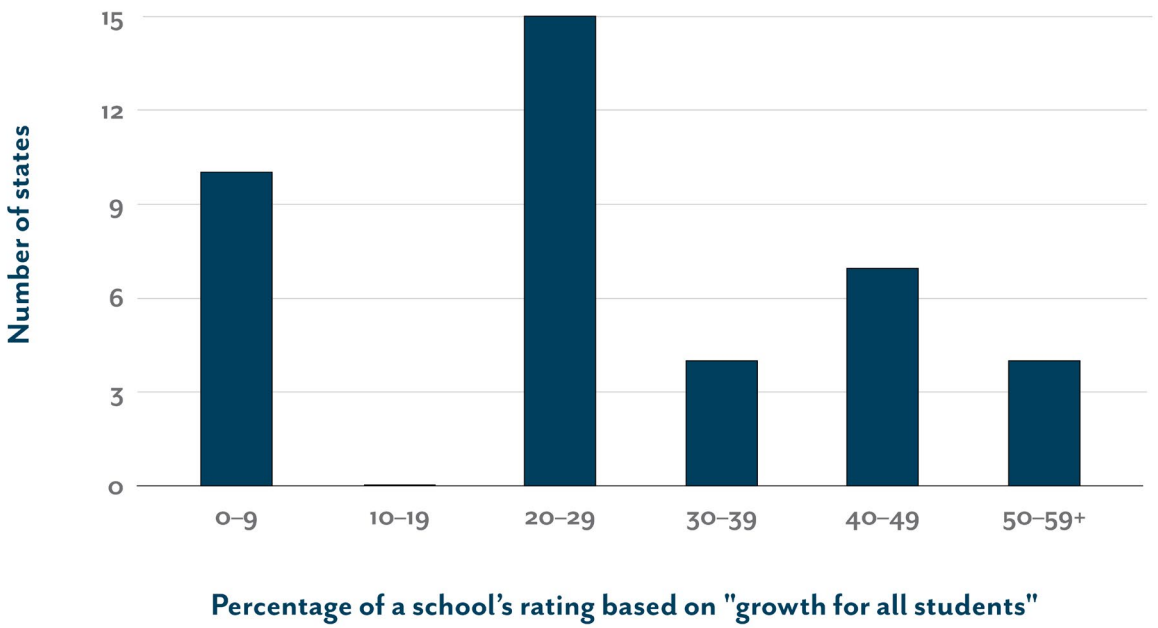
Instead, the department's final regulations should allow or even encourage performance metrics that account for the achievement of all students, using practices such as proficiency indices or average scale scores. Such a regulation would be consistent with ESSA and would encourage schools to focus on all kids—as they should.

In general, states that calculate summative school ratings don't assign much weight to "growth for all students"

Of the thirty-nine states (plus D.C.) that calculate summative school ratings, just four (Arkansas, Colorado, Idaho, and Oregon) base at least half of a school's rating on "growth for all students." However, a number of other states approach this standard. "Growth for all students" counts for at least 40 percent of summative school ratings in an additional seven

states, and for between 30 percent and 39 percent in another four.[13] In fifteen states, it counts for between 20 percent and 29 percent of these ratings.[14] And in Rhode Island, it counts for only 9 percent.[15] Eight states (plus D.C.) still assign no weight whatsoever to "growth for all students," though in some cases, they do weight growth for subgroups or other types of growth (such as growth to proficiency). The fact that so many states are basing most or all of their summative school ratings on proficiency rates, which are poor measures of a school's true quality because they are so strongly correlated with student demographics and prior achievement, is difficult to defend. (See Figure 2.)

FIGURE 2: STATES WITH SUMMATIVE SCHOOL RATINGS ASSIGN LITTLE WEIGHT TO "GROWTH FOR ALL STUDENTS"



Twelve states base at least 50 percent of their summative school ratings on growth but base some or all of their growth ratings on growth for low-performing students or other subgroups, as opposed to "growth for all students." For example, Washington bases 60 percent of schools' grades on growth, but just 30 percent on "growth for all students."

Similarly, some states assign significant weight to other growth measures (such as growth to proficiency) that exclude progress for high achievers and thus do not count as "growth for all students." For example, South Dakota bases 40 percent of schools' grades on growth-to-proficiency measures.

Though no doubt well-intentioned, both of these approaches give schools an incentive to ignore their high-achieving students, especially in high-poverty settings where many kids are below grade level. Why not use a growth model that includes all students instead? And why not weight all students' growth equally, or at least make "growth for all students" count for more of a school's summative rating?

TABLE 4: SUMMATIVE RATINGS FOR EACH STATE BY INDICATOR

STATE	GIVE EXTRA CREDIT FOR ADVANCED ACHIEVEMENT	INCLUDE HIGH ACHIEVERS IN GROWTH MODEL	SEPARATELY REPORT GROWTH FOR HIGH ACHIEVERS	MAKE "GROWTH FOR ALL STUDENTS" COUNT FOR AT LEAST HALF OF A SCHOOL'S RATING	RATING
Alabama	★	☆	☆	☆	★☆☆☆
Alaska	☆	★	☆	☆	★☆☆☆
Arizona	☆	★	☆	☆	★☆☆☆
Arkansas	★	★	☆	★	★★★☆
California	☆	☆	☆	NA	☆☆☆
Colorado	☆	★	☆	★	★★☆☆
Connecticut	★	★	☆	☆	★★☆☆
Delaware	☆	★	☆	☆	★☆☆☆
District of Columbia	★	☆	☆	☆	★☆☆☆
Florida	☆	★	☆	☆	★☆☆☆
Georgia	★	★	☆	☆	★★☆☆
Hawaii	☆	★	☆	☆	★☆☆☆
Idaho	☆	★	☆	★	★★☆☆
Illinois	☆	★	☆	NA	★☆☆
Indiana	☆	★	☆	☆	★☆☆☆
Iowa	☆	★	☆	☆	★☆☆☆
Kansas	☆	★	☆	NA	★☆☆
Kentucky	★	★	☆	☆	★★☆☆
Louisiana	★	☆	☆	☆	★☆☆☆
Maine	☆	★	☆	☆	★☆☆☆
Maryland	☆	☆	☆	NA	☆☆☆
Massachusetts	★	★	☆	☆	★★☆☆
Michigan	☆	☆	☆	☆	☆☆☆☆
Minnesota	☆	★	☆	☆	★☆☆☆
Mississippi	☆	★	☆	☆	★☆☆☆
Missouri	★	★	☆	☆	★★☆☆
Montana	☆	☆	☆	NA	☆☆☆
Nebraska	★	★	☆	☆	★★☆☆
Nevada	☆	★	★	☆	★★☆☆
New Hampshire	☆	★	☆	☆	★☆☆☆
New Jersey	☆	★	☆	NA	★☆☆

STATE	GIVE EXTRA CREDIT FOR ADVANCED ACHIEVEMENT	INCLUDE HIGH ACHIEVERS IN GROWTH MODEL	SEPARATELY REPORT GROWTH FOR HIGH ACHIEVERS	MAKE "GROWTH FOR ALL STUDENTS" COUNT FOR AT LEAST HALF OF A SCHOOL'S RATING	RATING
New Mexico	☆	★	☆	☆	★☆☆☆
New York	☆	☆	☆	NA	☆☆☆
North Carolina	☆	★	★	☆	★★☆☆
North Dakota	☆	☆	☆	NA	☆☆☆
Ohio	★	★	★	NA	★★★
Oklahoma	☆	☆	☆	☆	☆☆☆☆
Oregon	☆	★	★	★	★★★☆
Pennsylvania	★	★	☆	☆	★★☆☆
Rhode Island	★	★	☆	☆	★★☆☆
South Carolina	★	★	☆	NA	★★☆
South Dakota	☆	☆	☆	☆	☆☆☆☆
Tennessee	☆	★	☆	NA	★☆☆
Texas	☆	★	☆	☆	★☆☆☆
Utah	☆	★	☆	☆	★☆☆☆
Vermont	☆	☆	☆	☆	☆☆☆☆
Virginia	☆	☆	☆	☆	☆☆☆☆
Washington	☆	★	☆	☆	★☆☆☆
West Virginia	☆	★	☆	☆	★☆☆☆
Wisconsin	★	★	☆	☆	★★☆☆
Wyoming	☆	★	★	☆	★★☆☆

Closing Thoughts

As Uncle Ben famously told Spider-Man, "With great power comes great responsibility." Since the advent of ESEA waivers, and certainly now under ESSA, states have had greater power to fix the flaws inherent in No Child Left Behind and signal to schools that all students—including high achievers—matter.

Admirably, most states have taken advantage of their additional flexibility to adopt robust growth models. But inexplicably, most have failed to put these growth models at the center of their school accountability systems. As a result, they have maintained one of NCLB's biggest problems—a focus on getting kids to "proficient."

States now have a chance to do better. While there may be a temptation for officials to simply tweak the systems that were developed under federal waivers, that would be an enormous mistake and a lost opportunity. Instead, almost every state in the land could dramatically upgrade its system by putting more emphasis on student growth, giving schools credit for getting kids to advanced levels of achievement, and calling attention to the performance of high achievers by treating them as their own subgroup.

High-achieving students—especially those growing up in poverty—need all of the attention they can get. They were an afterthought when No Child Left Behind was crafted fifteen years ago. Let's not make the same mistake again.

Endnotes

1. Eric A. Hanushek and Margaret E. Raymond, *Does School Accountability Lead to Improved Student Performance?* (Washington, D.C.: National Bureau of Economic Research, 2004), http://hanushek.stanford.edu/sites/default/files/publications/hanushek+raymond.2005 jpam 24-2.pdf; and Martin Carnoy and Susanna Loeb, "Does External Accountability Affect Student Outcomes? A Cross-State Analysis," Educational Evaluation and Policy Analysis 24, no. 4 (2002), https://cepa.stanford.edu/sites/default/files/EEPAaccountability.pdf.

2. U.S. Department of Education, National Center for Education Statistics, *Mapping State Proficiency Standards Onto NAEP Scales: Results from the 2013 NAEP Reading and Mathematics Assessments*, NCES 2015-046 (Washington, D.C.: United States Government Printing Office, 2013), http://nces.ed.gov/nationsreportcard/subject/publications/studies/pdf/2015046.pdf.

3. Jennifer Booher-Jennings, Below the Bubble: *"Educational Triage" and the Texas Accountability System* (New York, NY: Columbia University, 2005), http://aer.sagepub.com/content/42/2/231.short; and Dale Ballou and Matthew G. Springer, *Achievement Trade-Offs and No Child Left Behind* (Nashville, TN: Peabody College of Vanderbilt University, 2008), http://www.vanderbilt.edu/schoolchoice/documents/achievement_tradeoffs.pdf.

4. Jonathan Plucker, Jacob Hardesty, and Nathan Burroughs, *Talent on the sidelines: Excellence gaps and America's persistent talent underclass* (Storrs, CT: University of Connecticut, Center for Education Policy Analysis, 2013), http://cepa.uconn.edu/mindthegap.

5. Joshua S. Wyner, John M. Bridgeland, John J. DiIulio, Jr., *Achievement Trap: How America Is Failing Millions of High-Achieving Students from Lower-Income Families* (Washington, D.C.: Jack Kent Cooke Foundation, 2006), http://www.issuelab.org/resource/achievement_trap_how_america_is_failing_millions_of_highachieving_students_from_lowerincome_families; and Robert Theaker, Yun Xiang, Michael Dahlin, John Cronin, Sarah Durant, *Do High Flyers Maintain Their Altitude? Performance Trends of Top Students* (Washington, D.C.: Thomas B. Fordham Institute, 2011), http://edexcellence.net/publications/high-flyers.html.

6. For better ways the Department of Education could address this issue, see Morgan Polikoff et al., "A letter to the U.S. Department of Education (updated July 14)," MorganPolikoff.com (July 12, 2016), https://morganpolikoff.com/2016/07/12/a-letter-to-the-u-s-department-of-education/.

7. See Jonathan Plucker, Jennifer Giancola, Grace Healey, Daniel Arndt, and Chen Wang, *Equal talents, unequal opportunities: A report card on state support for academically talented low-income students* (Lansdowne, VA: Jack Kent Cooke Foundation, 2015), http://www.excellencegap.org/state-report.

8. In most states, the differences between elementary and middle school accountability systems are subtle. Consequently, in order to simplify our analysis, we decided to use middle school systems as a proxy for K–8 accountability in general. High school accountability, of course, involves a number of additional variables (such as graduation rates and college-level coursework). We will tackle that subject in a separate report.

9. Our definitions are taken from "A Practitioner's Guide to Growth Models," Council of Chief State School Officers, 2013, http://www.ccsso.org/Documents/2013GrowthModels.pdf.

10. Virginia calculates value-added for teachers but not schools, while New York uses a mean growth percentile model to identify low-performing schools but doesn't rate (or report) growth for the rest. The District of Columbia also fails to rate schools' growth, even though its primary charter school authorizer (the District of Columbia Public School Charter Board, which oversee 45 percent of the city's schools) does so as part of its accountability system.

11. One exception is Nebraska, which takes an average of students' raw test scores (thus rewarding improvement across the achievement distribution).

12. See, e.g., Morgan Polikoff et al., "A letter to the U.S. Department of Education (updated July 14)."

13. At the middle school level. At the elementary school level, Hawaii and Kentucky also meet this standard.

14. Although technically neither state assigns any weight to "growth for all students," we include Indiana in this group based on the weight it assigns to growth for the highest achieving 75 percent of students. (And we include New Mexico in the previous group based on similar logic.) We also include New Hampshire, where "growth for all students" and "growth for all others" each count for 12 percent of a school's summative rating.

15. In Rhode Island, as well as several other states, the actual percentage depends on the number of subgroups that exist at a given school. In these cases, we went with the lowest possible weight for "growth for all students."

Index of Profiles

Click on any of the following state names to view its rating profile: