

RATING THE RATINGS

**ANALYZING THE FIRST 17 ESSA
ACCOUNTABILITY PLANS**

BY BRANDON L. WRIGHT
AND MICHAEL J. PETRILLI



The Thomas B. Fordham Institute promotes educational excellence for every child in America via quality research, analysis, and commentary, as well as advocacy and charter school authorizing 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.
The Institute is neither connected with nor sponsored by Fordham University.

CONTENTS

Report Contents

Executive Summary	4
Introduction	6
Methods	9
National Results	12
Conclusion	16
Profiles	17
Glossary	34
Endnotes	35

Index of Profiles

Arizona	17
Colorado	18
Connecticut	19
Delaware	20
District of Columbia	21
Illinois	22
Louisiana	23
Maine	24
Massachusetts	25
Michigan	26
Nevada	27
New Jersey	28
New Mexico	29
North Dakota	30
Oregon	31
Tennessee	32
Vermont	33

EXECUTIVE SUMMARY

The federal Every Student Succeeds Act (ESSA) grants states more authority over their school accountability systems than did No Child Left Behind (NCLB).¹ Hence, states have an opportunity to design improved school rating systems. To date, sixteen states plus the District of Columbia have submitted their plans to the U.S. Department of Education.²

In our view, three of the most important improvements that states can make are to ensure that their accountability systems:

1. Assign annual ratings to schools that are clear and intuitive for parents, educators, and the public;
2. Encourage schools to focus on all students, not just their low performers; and
3. Fairly measure and judge all schools, including those with high rates of poverty.

Based on these three objectives, we rate states' planned accountability systems using the rubric in Table ES-1 and the most recent publicly available information.

Table ES-1. Rubric for rating state accountability systems

Assigns to schools annual ratings that are clear and intuitive			Encourages schools to focus on all students, not just low performers			Measures all schools fairly, including those with high rates of poverty		
Weak	Medium	Strong	Weak	Medium	Strong	Weak	Medium	Strong

Table ES-2 shows the results for the seventeen plans that have been submitted to the U.S. Department of Education, sixteen of which have enough information for us to rate. Individual one-page profiles explain each state's ratings in more detail. According to our analysis, seven of these sixteen proposed school rating systems are either good or great, defined as receiving at least two strong grades and one medium. That so many of the early submitters have designed systems that are big improvements over NCLB is excellent news indeed.

The seventeen jurisdictions that submitted their plans by the first deadline volunteered to be guinea pigs—or, if you like, sacrificial lambs. There is little doubt that the other thirty-four states are watching closely, both to see models they might emulate and to learn how the U.S. Department of Education reacts to what has been proposed.

Although many states included elements in their school rating systems that we don't love, we are encouraged that nearly half of the proposals we were able to assess are either good or great by our reckoning, and that only one misses the mark entirely. This is significantly better than the status quo that reigned during the NCLB era.

Will the remaining states do even better? We see no reason that they cannot decide to rate their schools in a clear and intuitive way; to signal that schools should focus on all kids, not just low performers; and to ensure that high-poverty schools are treated fairly. We'll be back in the autumn to find out how they do.

EXECUTIVE SUMMARY

Table ES-2. Results for states that have submitted plans to the U.S. Department of Education

	Clear Labels	Focus on All Students	Fair to All Schools
Arizona	Strong	Strong	Strong
Colorado	Strong	Strong	Strong
Connecticut	Strong	Strong	Medium
Delaware	Medium	Weak	Medium
District of Columbia	Strong	Strong	Medium
Illinois	Strong	Strong	Strong
Louisiana	Strong	Strong	Weak
Maine	Medium	Medium	Medium
Massachusetts	Strong	Strong	Weak
Michigan	N/A	N/A	N/A
Nevada	Strong	Weak	Medium
New Jersey	Medium	Medium	Medium
New Mexico	Strong	Medium	Medium
North Dakota	Weak	Weak	Weak
Oregon	Weak	Medium	Medium
Tennessee	Strong	Medium	Strong
Vermont	Strong	Strong	Medium

Acknowledgments

This report was made possible through the generous support of our sister organization, the Thomas B. Fordham Foundation. We are grateful to Chester E. Finn, Jr., Amber M. Northern, Morgan Polikoff, and David Griffith for their feedback; to Alyssa Schwenk for her role in disseminating the final product; to Jonathan Lutton for the report's layout; to Tiyanna Stewart for her general support; and to Pamela Tatz for copyediting.

We are particularly appreciative of the officials in state departments of education who took the time to review our assessments of their accountability systems.

Any errors are ours alone.

INTRODUCTION

Critics aren't wrong when they say that much of what's in state plans to implement ESSA is meaningless blather.³ This is, at its heart, an elaborate compliance exercise, and most of the pretty words that states put to paper will probably not amount to much. That is why many of the recent and forthcoming reviews of these ESSA plans will be somewhat misguided in their focus on various bromides and promises.

However, there is one part of states' ESSA plans that will likely matter a great deal: the design of state accountability systems—and, in particular, the ratings or labels states place on their public schools, their various components, and the methodologies used to develop them. Rigorous and well-respected studies from the NCLB era demonstrated that such ratings can and do drive behavior in schools, for better or worse.⁴

Famously, ESSA grants states greater authority over their school accountability regimes than did NCLB. Hence states have an opportunity to design significantly improved school ratings (see *A note on accountability and state autonomy*). In our view, three of the most important improvements that states can make are to ensure that their accountability systems (1) assign to schools annual ratings that are intuitive for parents, educators, and the public; (2) encourage schools to focus on all of their students, not just low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty. Let us explain why.

A note on accountability and state autonomy

This analysis prescribes a trio of objectives that we believe states ought to fulfill when designing their school accountability systems. Nevertheless, we do not believe that there's any one "best way" to construct these frameworks. ESSA gives substantial authority back to the states, where it belongs, both rationally and constitutionally. States ought to take advantage of this newfound freedom to design systems that best fit the needs of their students, families, and communities.

No matter the route state policymakers choose, however, there are tradeoffs. Focusing on all students—as we strongly advise—means that schools and teachers may pay a little less attention to low performers alone, for example. We understand the risk, but we also strongly believe that it's a mistake to write off higher-achieving students, especially those growing up in poverty. This is a normative value, and we don't assume that everyone will share it. That's the case with every principle that informs our analysis and recommendations.

1

Intuitive ratings

For more than two decades now, school ratings have been at the heart of state accountability systems—and for good reason. Easy-to-understand labels, such as A–F letter grades, provide clear signals to parents, citizens, and educators about the quality of a school and can nudge systems toward improvement. "Dashboards of data" are great complements to this, especially when teams sit down to determine how to take a school to the next level, but they are no replacement. Furthermore, there's simply no excuse for states to assign labels that are impossible to parse, which strikes us as an Orwellian approach to keep interested parties in the dark about school quality.

INTRODUCTION

Therefore, states ought to assign annual ratings to schools and ensure that these ratings are clear and intuitive for educators, parents, and the general public.

2

A focus on all students

NCLB meant well (as did many state accountability systems that were put in place before 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 pretty much ignoring everyone else. This approach led schools to focus on “bubble kids,” those just below or above the proficiency cutoff, to the detriment of other students.⁵ Among those neglected were both high achievers—those already over the proficiency threshold—and exceptionally low achievers for whom proficiency (even when none too demanding) seemed beyond the best efforts of teachers and schools.

We understand the impulse to make low-performing students a priority. Many U.S. schools need to do far better by them, and before NCLB, their needs were often ignored. But they aren’t the only children who matter, and acting as if they are signals to schools that students who are already proficient don’t deserve to have their education maximized. Such neglect is inequitable. It’s also bad for social mobility and harmful to the country’s long-term prosperity. The students most harmed by this are able pupils from disadvantaged backgrounds.⁶ They depend far more than upper-middle-class students on the public education system to support them. So if they don’t receive the attention that they—like all children—deserve, many will fall by the wayside, destined by circumstances beyond their control never to realize their full potential.⁷

The country also needs such children to be highly educated to ensure its long-term competitiveness, security, and innovation. These boys and girls hold great promise for making major advances in science, technology, medicine, the humanities, and much more. America’s economic vitality depends heavily on the quality and productivity of our human capital and its capacity for innovation.⁸

3

Fairness for high-poverty schools

Finally, we think it’s imperative that state ratings be fair to high-poverty schools. Under NCLB accountability in many states, nearly every school serving a high proportion of low-income students was eventually designated as failing.⁹ Although it’s no secret that too many high-poverty schools are in fact ineffective, it’s absurd to conclude that that’s the case with nearly all of them.

This happened because most of the NCLB-era measures of school performance—especially proficiency and graduation rates—were strongly correlated with prior achievement and pupil demographics. Such metrics are more reflective of the students that a school serves—and what they have or haven’t learned before stepping foot in a given school—than the effectiveness of their instruction.¹⁰

Thankfully, ESSA allows states to move on and to focus a school’s metrics more on what’s under the control of educators: how much students learn while in their classrooms. States that embrace this approach should find that at least some of their high-poverty schools earn good to excellent ratings—because they deserve them. If not, the rating system is likely broken.

Organization of this report

States must describe their proposed accountability systems in “consolidated state plans” that they submit to the U.S. Department of Education for review.¹¹ The first deadline was in April 2017, when sixteen states and the District of Columbia submitted their plans.¹² The remaining thirty-four states, plus the territories, will submit theirs by the second deadline in September 2017.

This report examines how well the first seventeen plans fulfill the three objectives we delineate above. It does not examine the quality of states’ standards, tests, or interventions in low-performing schools.

The data in this report reflect information that was publicly available as of July 16, 2017, compiled from state plans as published on the U.S. Department of Education website.¹³

For simplicity’s sake, our focus here is on state rating systems for elementary and middle schools.

METHODS

As explained above, we believe that a strong state accountability system (1) assigns to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourages schools to focus on all students; and (3) measures and judges all schools fairly, including those with high rates of poverty.

To determine whether the proposed accountability systems described in the first seventeen plans submitted to U.S. Department of Education accomplish these three objectives, we evaluate them in the following way, using the most recent publicly available information. A state can receive a designation of “Strong,” “Medium,” or “Weak” on each objective.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Annual school ratings should immediately convey to all observers how well a given school is performing. Straightforward approaches like A–F grades and five-star systems do this well and are therefore ideal. Text labels that are easy to understand have some merit, but these often fail to communicate how much better or worse a given school could do (it’s unclear whether “highly effective,” for example, would be the best possible rating in a state that uses that label). And systems that offer myriad data points with no bottom line (for example, “data dashboards”) or that employ murky text labels do neither. Therefore, this analysis assesses states in the following manner:

- **Strong:** A–F, five stars, or the equivalent
- **Medium:** Text labels that are easy to understand
- **Weak:** Myriad data points with no bottom line or text labels that aren’t clear regarding a school’s quality

2

Does the rating system encourage a focus on all students?

There are two primary ways that state accountability systems can encourage schools to focus on all students and not just low-performers. First, they can measure achievement in a way that gives schools credit for raising the performance of students along the entire performance spectrum, by using a performance index or average scale scores (see *Glossary* for an explanation of scale scores).¹⁴ For example, a state might create a performance index that gives schools partial credit for getting students to a basic level of achievement, full credit for getting students to proficiency, and additional credit for getting students to an advanced level.

Second, when calculating annual school ratings, growth of all students (as opposed to growth to proficiency, growth for low performers, etc.) from one year to the next can be given substantial weight (see *Glossary* for an explanation of different growth measures). Again, this puts the focus on all kids, instead of just some. Hence state plans are assessed as follows:

- **Strong:** At least 50 percent of schools’ annual ratings are made up of one or more of the following academic indicators: (1) performance index; (2) average scale scores; and (3) growth for all students.

- **Medium:** Between 33 and 50 percent of schools' annual ratings are made up of one or more of the following academic indicators: (1) performance index; (2) average scale scores; and (3) growth for all students.
- **Weak:** Less than 33 percent of schools' annual ratings are made up of one or more of the following academic indicators: (1) performance index; (2) average scale scores; and (3) growth for all students.

3

Is the rating system fair to all schools, including those with high rates of poverty?

Across all seventeen plans that have been submitted to the U.S. Department of Education under ESSA, indicators of student achievement and student growth are the two biggest components of accountability systems. States are required to measure achievement, but they should not overweight it when calculating annual school ratings, even if they're using a performance index or scale scores. That's because all achievement measures are strongly correlated with prior achievement—and given that low-income students tend to enter school far behind their peers, high-poverty schools are likely to fare poorly under such measures, no matter how good the school and its teachers are.¹⁵

Growth measures, however, quantify changes in achievement over time, independent of whether students start as high or low performers; hence they're less correlated with poverty (see *Glossary* for an explanation of various types of growth measures).¹⁶ Annual school ratings should, above all, accurately assess the true performance of schools, and that can't be done unless it's possible for high-performing, high-poverty schools to actually earn positive ratings (see *Growth measures as a proxy for fairness for high-poverty schools*). Growth measures should therefore constitute the majority of summative ratings, and state plans are thus evaluated in the following manner:

- **Strong:** Academic growth of any kind constitutes at least 50 percent of schools' annual ratings.
- **Medium:** Academic growth of any kind constitutes between 33 and 50 percent of schools' annual ratings.
- **Weak:** Academic growth of any kind constitutes less than 33 percent of schools' annual ratings.

Using these criteria and the rubric in Table 1, we rate the proposed accountability systems described in the first seventeen plans submitted to the U.S. Department of Education.

Table 1. Rubric for rating state accountability systems

Assigns to schools annual ratings that are clear and intuitive			Encourages schools to focus on all students, not just low performers			Measures all schools fairly, including those with high rates of poverty		
Weak	Medium	Strong	Weak	Medium	Strong	Weak	Medium	Strong

Growth measures as a proxy for fairness to high-poverty schools

The optimal way to determine whether a given state's accountability system fairly measures the performance of high-poverty schools would be to measure rigorously how closely correlated with poverty each indicator is that the state proposes to use. One could then combine those correlation coefficients to determine how strongly correlated that state's total accountability framework is with poverty—and, therefore, how fairly it measures high-poverty schools.

Unfortunately, no such correlation measurements exist for many of the indicators that states have proposed; and even if they did, states' explanations of their indicators in their consolidated state plans are often not detailed enough to accurately ascertain whether any extant research can be justifiably applied.

Nevertheless, we do know a few things. First, status measures in general—and proficiency rates in particular—are strongly correlated with poverty.¹⁷ Second, growth measures are much less correlated with poverty than are achievement measures.¹⁸ And third, achievement and growth are by far the two biggest components of state accountability frameworks for K–8 schools.¹⁹ Therefore, the more states focus on growth instead of achievement, the fairer they will be to high-poverty schools.

NATIONAL RESULTS

Table 2 shows the results for all seventeen state plans that have been submitted to the U.S. Department of Education. According to our analysis, seven of the seventeen plans propose school rating systems that are either good or great. That so many of the early submitters have designed systems that are big improvements over NCLB is excellent news indeed.

Table 2. Results for states that have submitted plans to the U.S. Department of Education

	Clear Labels	Focus on All Students	Fair to All Schools
Arizona	Strong	Strong	Strong
Colorado	Strong	Strong	Strong
Connecticut	Strong	Strong	Medium
Delaware	Medium	Weak	Medium
District of Columbia	Strong	Strong	Medium
Illinois	Strong	Strong	Strong
Louisiana	Strong	Strong	Weak
Maine	Medium	Medium	Medium
Massachusetts	Strong	Strong	Weak
Michigan	N/A	N/A	N/A
Nevada	Strong	Weak	Medium
New Jersey	Medium	Medium	Medium
New Mexico	Strong	Medium	Medium
North Dakota	Weak	Weak	Weak
Oregon	Weak	Medium	Medium
Tennessee	Strong	Medium	Strong
Vermont	Strong	Strong	Medium

The accountability systems in three states—Arizona, Colorado, and Illinois—are the best thus far, receiving **strong** grades across the board. They propose to use ratings that clearly and intuitively convey to all observers how well a given school is performing. They signal that all students matter by ensuring that at least 50 percent of schools' annual ratings are comprised of measures of growth for all students and/or measures of achievement that look beyond proficiency rates. And they're fair to all schools—including those with high rates of poverty—by virtue of making growth measures of any kind constitute at least half of schools' summative ratings.

The accountability systems proposed by four others—Connecticut, the District of Columbia, Tennessee, and Vermont—are also good, each receiving two **strong** grades and one **medium**. Those states provide educators, parents, and the public clear annual ratings for schools, with either A–F grades, five-star systems, or user-friendly numerical systems. In three plans—Connecticut, D.C., and Vermont—measures that encourage schools to heed the education needs of all students constitute 50 percent or more of schools' annual ratings. And Tennessee received a **strong** for its fairness to high-poverty schools.

Louisiana's and Massachusetts's plans are also laudable in multiple ways, receiving two **strong** grades and one **weak**. They both propose to use clear and intuitive annual school ratings and encourage schools to focus on students across the achievement spectrum. Yet a low emphasis on growth means that neither is sufficiently fair to high-poverty schools.

New Mexico earned one **strong** grade for its proposal to use clear and intuitive annual school ratings, but **medium** grades for its focus on all students and its fairness to high-poverty schools. It still uses proficiency rates, and growth counts for 42 percent of schools' annual ratings.

Nevada also earned one **strong** grade for its proposal to use clear and intuitive annual school ratings. But it received a **medium** grade for its fairness to high-poverty schools and a **weak** grade for its focus on all students, due to a reliance on proficiency rates and a lack of emphasis on student growth.

Four more—**Delaware, Maine, New Jersey, and Oregon**—earn only **medium** and **weak** grades by failing to sufficiently fulfill all three objectives.

But **North Dakota** is the worst of the seventeen. It relies on proficiency rates, doesn't emphasize student growth, and proposes using a dashboard-like approach for reporting school quality to parents that fails to communicate any bottom line. It therefore received three **weak** grades.

Additionally, **Michigan** has failed to decide whether it will use annual school ratings and thereby risks making the same mistake as North Dakota. It received marks of **not applicable** across the board.

Individual one-page profiles explain each state's ratings in more detail, and start on page 17.

1

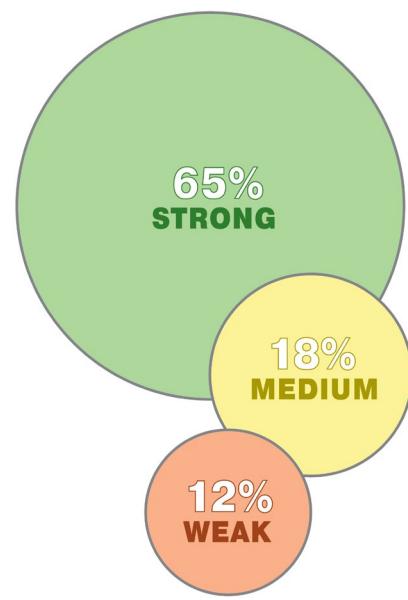
Annual ratings

In most states, the annual school ratings are clear and intuitive for parents, educators, and the public (see Figure 1).

Eleven of the first seventeen states to submit plans—or 65 percent—received a **strong** grade by using clear and intuitive ratings that immediately convey to all observers how well a given school is performing. Three of these exemplary states—**Arizona, Louisiana, and Tennessee**—use letter grades. Another trio—**Colorado, Connecticut, and New Mexico**—rate schools on a hundred-point system. Two more, **D.C. and Nevada**, use a system of up to five stars. And **Illinois, Massachusetts, and Vermont** employ a hybrid system comprising text labels and between four and six possible numerical levels.

Three more states—**Delaware, Maine, and New Jersey**—designed systems of intuitive text labels—that is, they used words that are easy to understand. These have some merit but often fail to communicate how much better or worse a given school could do (it's unclear, for example, whether "highly effective" would be the best possible rating in a state that uses that label). These states, therefore, received **medium** grades.

Figure 1. State grades for the clarity and intuitiveness of their annual school ratings²⁰



NATIONAL RESULTS

Oregon and North Dakota have officially decided against using summative school ratings and were thus the only states to be graded as **weak** in this area.

Michigan, which is currently debating whether to use or forgo annual ratings, received a notation of **not applicable**.

2

Signaling that all students matter, not just low performers

We can report some very good news here: twelve of the sixteen accountability systems that have been proposed and finalized so far range from good to great when it comes to signaling that all students matter—a vast improvement over NCLB (see Figure 2).

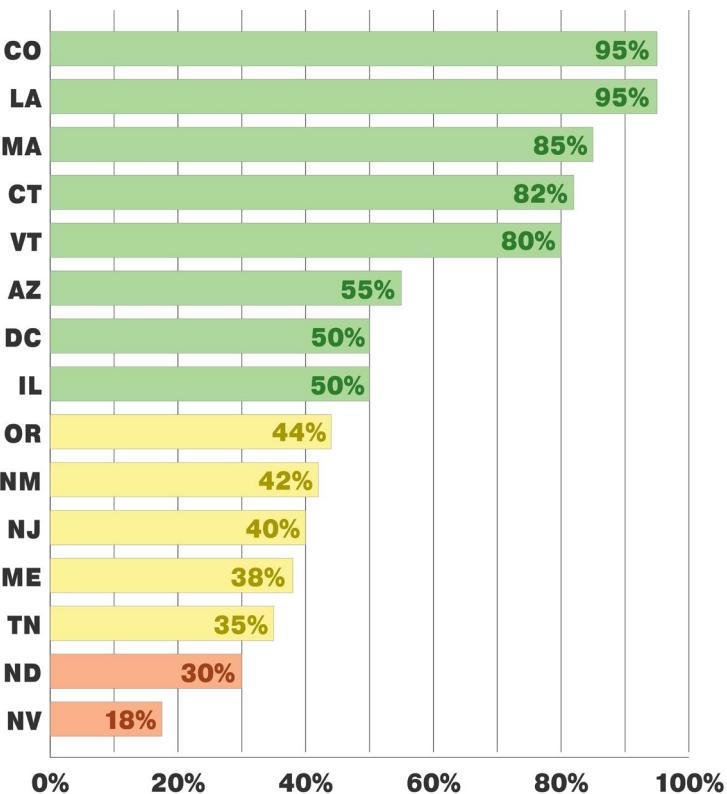
Of these sixteen finalized plans, eight—Arizona, Colorado, Connecticut, D.C., Illinois, Louisiana, Massachusetts, and Vermont—earned a grade of **strong**.²² In each, at least 50 percent of schools' annual ratings are made up of one or more of the following academic indicators: (1) performance index; (2) average scale scores; and (3) growth for all students.

Five states—Maine, New Jersey, New Mexico, Oregon, and Tennessee—received a **medium**. Each of these states carried over NCLB's ill-chosen use of proficiency rates. And growth for all counts for 44 percent of schools' annual ratings in Oregon, 42 percent in New Mexico, 40 percent in New Jersey, 38 percent in Maine, and 35 percent in Tennessee.

Weak grades went to the remaining states—Delaware, Nevada, and North Dakota—for keeping proficiency rates as their measure of achievement and also downplaying growth for all students. In these jurisdictions, one may fairly worry whether everyone above the proficiency line will continue to be ignored.

Michigan received a mark of **not applicable** because it is still debating whether to use or forgo annual ratings and specific weights for the various components of its accountability system.

Figure 2. Percentage of school ratings that comprise performance indices, scale scores, and growth for all²¹



3

Fairness to high-poverty schools

The news here is also good. Most of the proposed accountability systems are much fairer to high-poverty schools than state systems in place under NCLB (see Figure 3).

As previously stated, one of NCLB's most pernicious flaws was how unfairly it treated high-poverty schools. Because proficiency rates are so strongly correlated with prior achievement and so many children growing up in poverty tend to enter school so far behind, under the previous federal law, schools with many low-income pupils had little chance of receiving a positive school rating—regardless of how much its teachers and staff did to boost the learning of those youngsters.

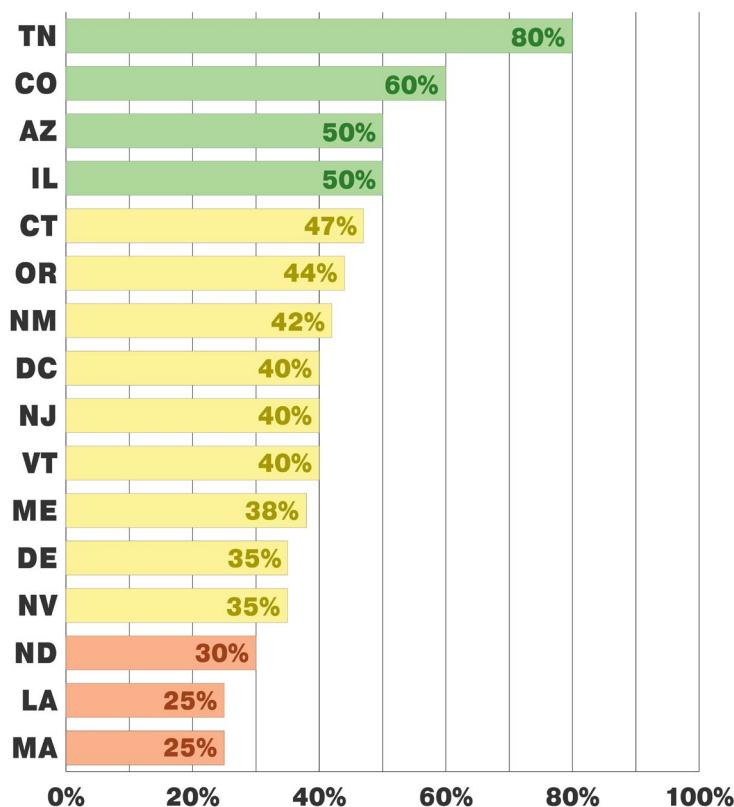
It is therefore encouraging that four of the sixteen states with finalized accountability systems received ratings of **strong** on this indicator. In these jurisdictions, growth of any kind—growth for all, growth to proficiency, growth for a subgroup of students, and so forth—constitutes at least half of schools' annual summative rating. We predict therefore that high-performing, high-poverty schools in **Arizona, Colorado, Illinois, and Tennessee** should be capable of earning positive ratings—a remarkable turnaround from the NCLB era.

Nine others—**Connecticut, Delaware, D.C., Maine, Nevada, New Jersey, New Mexico, Oregon, and Vermont**—received a rating of **medium**. Growth of any kind combines for 47 percent in Connecticut; 44 percent in Oregon; 42 percent in New Mexico; 40 percent in D.C., New Jersey, and Vermont; 38 percent in Maine; and 35 percent in Delaware and Nevada.²³ So compared to the achievement-only structures of these systems under NCLB, their progress is also laudable.

Sadly, the systems in **Louisiana, Massachusetts, and North Dakota** are still lacking in this area, and they all received **weak** grades. The first two count growth of any kind for just 25 percent of schools' annual ratings. In North Dakota, it's 30 percent.

Michigan received a notation of **not applicable**, as it did for the other two objectives, because it is still debating whether to use or forgo annual ratings and specific indicator weights.

Figure 3. Percentage of school ratings that comprises growth of any kind²⁴





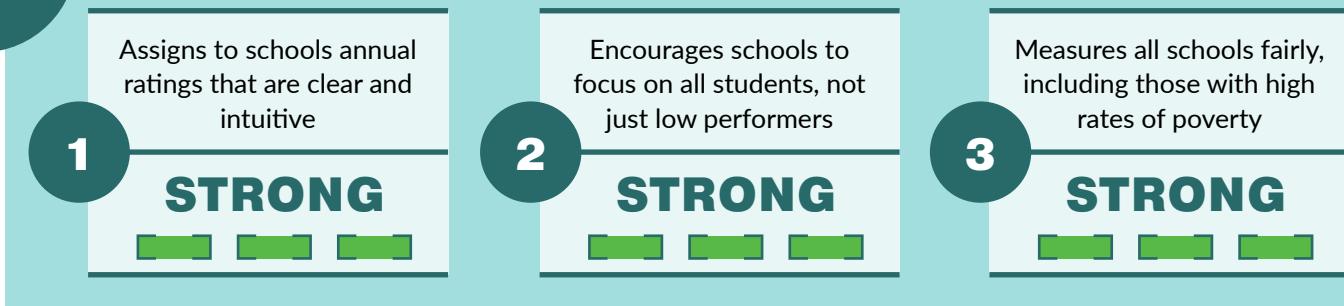
CONCLUSION

The seventeen jurisdictions that submitted their ESSA applications by the first deadline volunteered to be guinea pigs—or, if you like, sacrificial lambs. There is little doubt that the other thirty-four states are watching closely, both to see models they might emulate and to learn how the U.S. Department of Education reacts to what has been proposed.

Although many states included elements in their school rating systems that we don't love, we are encouraged that seven of the proposals are either good or great by our reckoning and that only one misses the mark entirely. This is significantly better than the status quo that reigned during the NCLB era.

Will the remaining states do even better? We see no reason that they cannot decide to rate their schools in a clear and intuitive way; to signal that schools should focus on all kids, not just low performers; and to ensure that high-poverty schools are treated fairly. We'll be back in the autumn to find out how they do.

ARIZONA



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether Arizona's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on May 9, 2017,²⁵ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Arizona's plan is **strong** on this point because it proposes to use an A–F system for schools' annual ratings. This model immediately conveys to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

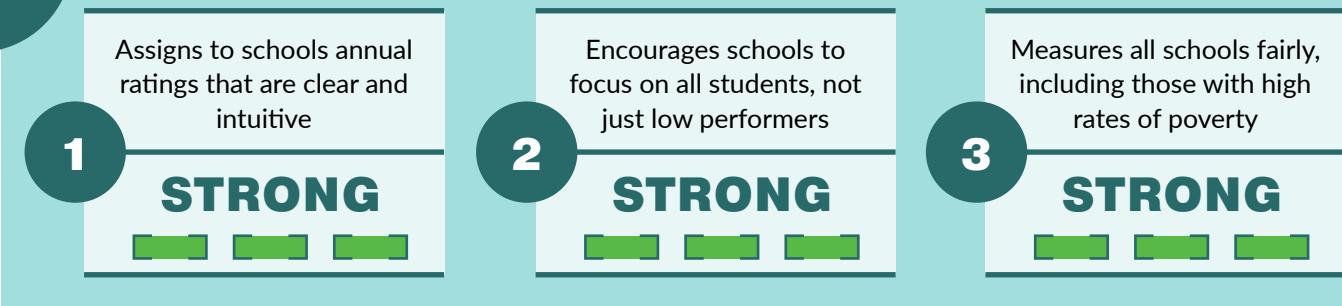
There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. Arizona receives a **strong** because those two components constitute 55 percent of schools' annual ratings. A performance index counts for 30 percent, which encourages schools to look beyond those pupils who are near the cutoff for proficiency. And a measure of growth for all students constitutes another 25 percent of schools' summative ratings, which should also lead schools to heed the educational needs of every child.

3

Is the rating system fair to all schools, including those with high rates of poverty?

Arizona is **strong** here because academic growth will constitute 50 percent of schools' annual ratings—split evenly between growth for all students and growth to proficiency. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, allowing high-poverty schools the opportunity to earn positive ratings.

COLORADO



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether Colorado's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on May 9, 2017,²⁶ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Colorado's plan is **strong** on this point because it proposes to use a hundred-point scale for schools' annual ratings. This model immediately conveys to all observers how well a given school is performing in relation to the state's other schools.

2

Does the rating system encourage schools to focus on all students?

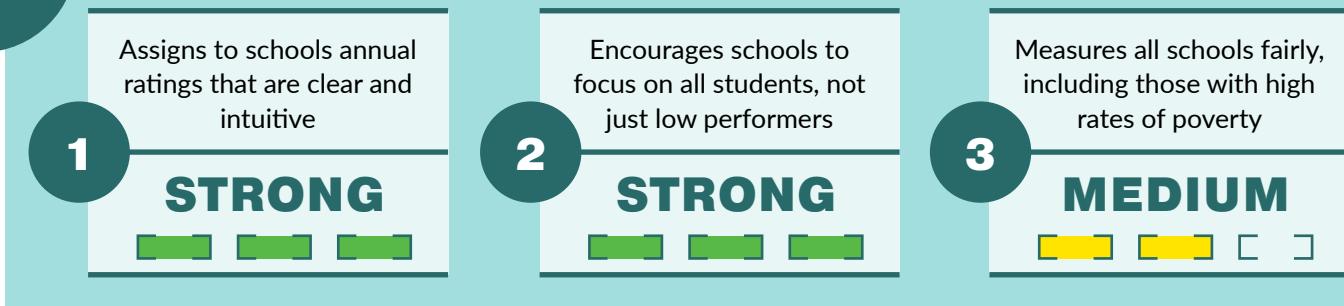
There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. Colorado receives a **strong** rating because those two components constitute 95 percent of schools' annual ratings. Average scale scores count for 35 percent, which encourages schools to look beyond those pupils who are near the cutoff for proficiency.²⁷ And a measure of growth for all students constitutes another 60 percent of schools' summative ratings, which should also lead schools to heed the educational needs of every child.

3

Is the rating system fair to all schools, including those with high rates of poverty?

Colorado is **strong** here because academic growth will constitute 60 percent of schools' annual ratings—all of which is a measure of growth for all students. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

CONNECTICUT



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether Connecticut's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on April 21, 2017,²⁸ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Connecticut's plan is **strong** on this point because it proposes to use a hundred-point scale for schools' annual ratings. This model immediately conveys to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

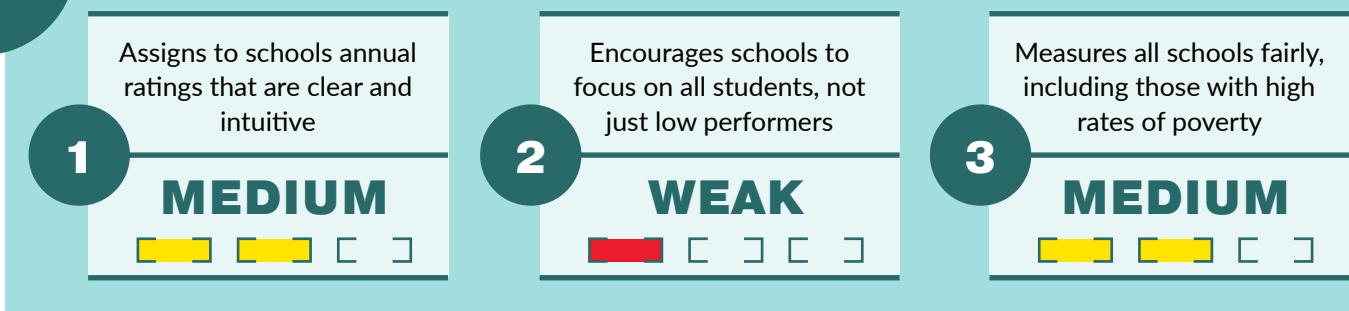
There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. Connecticut receives a **strong** rating because those two components constitute 82 percent of schools' annual ratings. Scale scores count for 35 percent, which encourages schools to look beyond those pupils who are near the cutoff for proficiency.²⁹ And a measure of growth for all students constitutes another 47 percent of schools' summative ratings, which should also lead schools to heed the educational needs of every child.

3

Is the rating system fair to all schools, including those with high rates of poverty?

Connecticut earns a **medium** here because academic growth will constitute 47 percent of schools' annual ratings—all of which is a measure of growth for all students. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

DELAWARE



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether Delaware's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on April 3, 2017,³⁰ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Delaware's plan earns a **medium** on this point because it proposes to use text labels as schools' annual ratings. Although the proposed labels are easy to understand, in isolation each one fails to communicate how much better or worse a given school could do (it's not instantly clear to a parent, for example, whether "exceeds expectations" is Delaware's best possible rating). Thus this model fails to convey immediately to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. Delaware receives a **weak** because it measures achievement with proficiency rates, which may encourage schools to focus on pupils near the proficiency cutoff—and because a measure of growth for all students constitutes less than 33 percent of schools' annual ratings, which is apt to lead schools to disregard the educational needs of higher-achieving children, especially those in high-poverty schools.

3

Is the rating system fair to all schools, including those with high rates of poverty?

Delaware earns a **medium** here because academic growth will constitute 35 percent of schools' annual ratings—comprising overall growth for all students in math and English language arts, growth to proficiency, and growth of the lowest- and highest-achieving student quartiles. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

DISTRICT OF COLUMBIA



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether the District of Columbia's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on May 2, 2017,³¹ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

D.C. is **strong** on this point because it proposes to use a five-star system for schools' annual ratings. This model immediately conveys to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

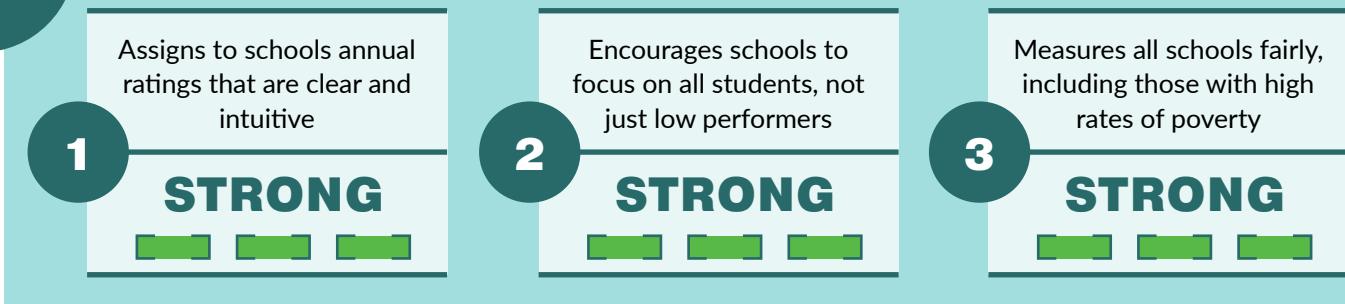
There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. D.C. receives a **strong** because those two components constitute 50 percent of schools' annual ratings. A performance index counts for 30 percent, which encourages schools to look beyond those pupils who are near the cutoff for proficiency. And a measure of growth for all students constitutes another 20 percent of schools' summative ratings, which should also lead schools to heed the educational needs of every child.

3

Is the rating system fair to all schools, including those with high rates of poverty?

D.C. earns a **medium** here because academic growth will constitute 40 percent of schools' annual ratings—split evenly between growth for all students and growth to proficiency. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

ILLINOIS



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether Illinois's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on May 2, 2017,³² as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Illinois's plan is **strong** on this point because it proposes to use a four-tier system for schools' annual ratings. This model immediately conveys to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

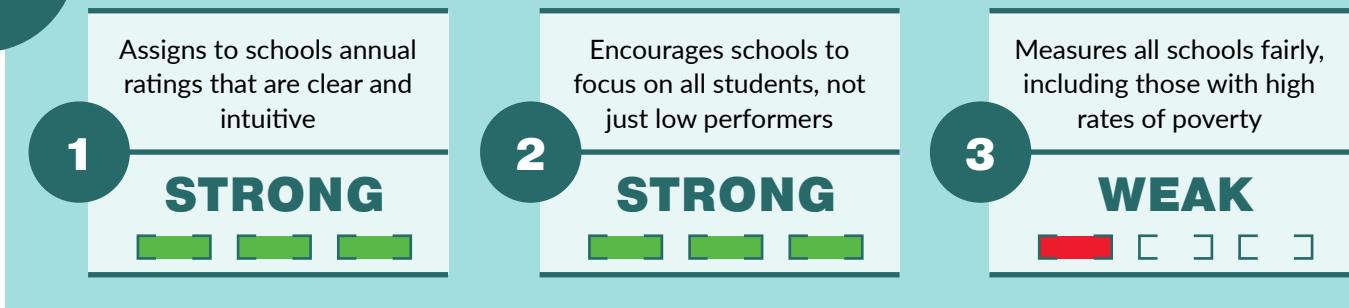
There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. Illinois receives a **strong** because—despite measuring achievement with proficiency rates, which might encourage schools to focus on pupils near the proficiency cutoff—a measure of growth for all students constitutes 50 percent of schools' summative ratings, which should lead schools to heed the educational needs of every child.

3

Is the rating system fair to all schools, including those with high rates of poverty?

Illinois is **strong** here because academic growth will constitute 50 percent of schools' annual ratings—all of which is a measure of growth for all students. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

LOUISIANA



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether Louisiana’s proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on May 3, 2017,³³ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Louisiana’s plan is **strong** on this point because it proposes to use an A–F system for schools’ annual ratings. This model immediately conveys to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

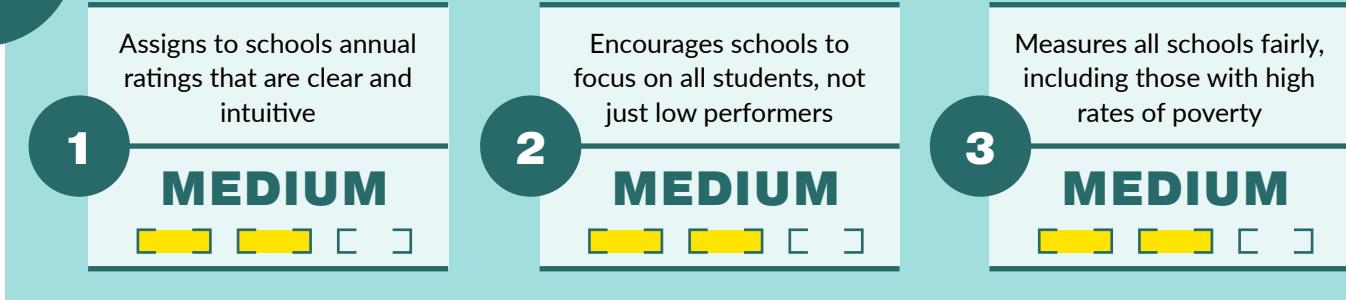
There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. Louisiana receives a **strong** because those two components constitute 95 percent of schools’ annual ratings. A performance index counts for 70 percent, which encourages schools to look beyond those pupils who are near the cutoff for proficiency. And a measure of growth for all students constitutes another 25 percent of schools’ summative ratings, which should also lead schools to heed the educational needs of every child.

3

Is the rating system fair to all schools, including those with high rates of poverty?

Louisiana is **weak** here because academic growth will constitute just 25 percent of schools’ annual ratings—all of which is a measure of growth for all students. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

MAINE



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether Maine’s proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on May 4, 2017,³⁴ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Maine’s plan earns a **medium** on this point because it proposes to use text labels for schools’ annual ratings. Although the proposed labels are easy to understand, in isolation each label fails to communicate how much better or worse a given school could do (it’s not instantly clear, for example, whether “below expectations” is the worst possible rating; in Maine, it’s second worst after “requires review for support”). Thus this model fails to convey immediately to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

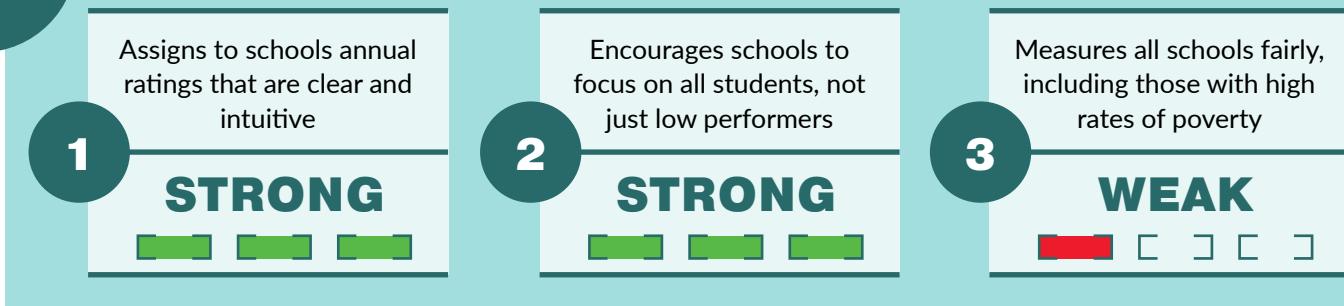
There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. Maine receives a **medium** because—despite measuring achievement proficiency rates, which might encourage schools to focus on pupils near the proficiency cutoff—a measure of growth for all students constitutes 38 percent of schools annual ratings, which should at least partially encourage schools to heed the educational needs of every child.

3

Is the rating system fair to all schools, including those with high rates of poverty?

Maine earns a medium here because academic growth will constitute 38 percent of schools’ annual ratings—all of which is a measure of growth for all students. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

MASSACHUSETTS



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether Massachusetts's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on May 10, 2017,³⁵ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Massachusetts's plan is **strong** on this point because it proposes to use a six-tier system for schools' annual ratings. This model immediately conveys to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. Massachusetts receives a **strong** because those two components constitute 85 percent of schools' annual ratings. Average scale scores count for 60 percent, which encourages schools to look beyond those pupils who are near the cutoff for proficiency.³⁶ And a measure of growth for all students constitutes another 25 percent of schools' summative ratings, which should also lead schools to heed the educational needs of every child.

3

Is the rating system fair to all schools, including those with high rates of poverty?

Massachusetts is **weak** here because academic growth will constitute just 25 percent of schools' annual ratings—all of which is a measure of growth for all students. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

MICHIGAN

Assigns to schools annual ratings that are clear and intuitive	Encourages schools to focus on all students, not just low performers	Measures all schools fairly, including those with high rates of poverty
1 N/A <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	2 N/A <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	3 N/A <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether Michigan's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on May 3, 2017,³⁷ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Michigan receives a mark of **not applicable** because it has not decided whether it will assign summative annual ratings to schools. We encourage policymakers to adopt such ratings—and ensure that they're clear and intuitive for educators, parents, and the general public by using an A–F system, five-star system, or the equivalent.

2

Does the rating system encourage schools to focus on all students?

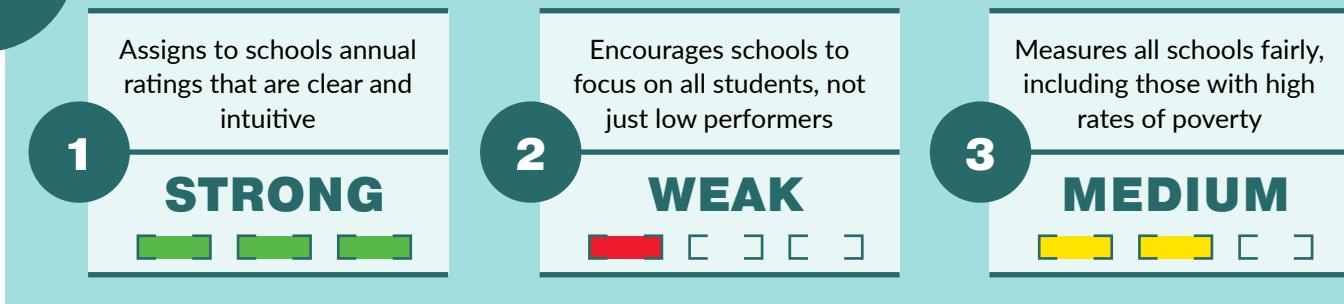
There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. Michigan receives a mark of **not applicable** because it has not decided whether it will assign schools annual ratings. The state uses proficiency rates—a mistake that might encourage schools to focus on pupils just above or below the proficiency cutoff, to the detriment of other pupils. But it also uses a measure of growth for all students, which can encourage schools to heed the educational needs of every child.

3

Is the rating system fair to all schools, including those with high rates of poverty?

Michigan receives a mark of **not applicable** because it has not decided whether it will give schools annual ratings or assign a specific weight to its growth model. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

NEVADA



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether Nevada's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on April 12, 2017,³⁸ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Nevada's plan is **strong** on this point because it proposes to use a five-star system for schools' annual ratings. This model immediately conveys to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

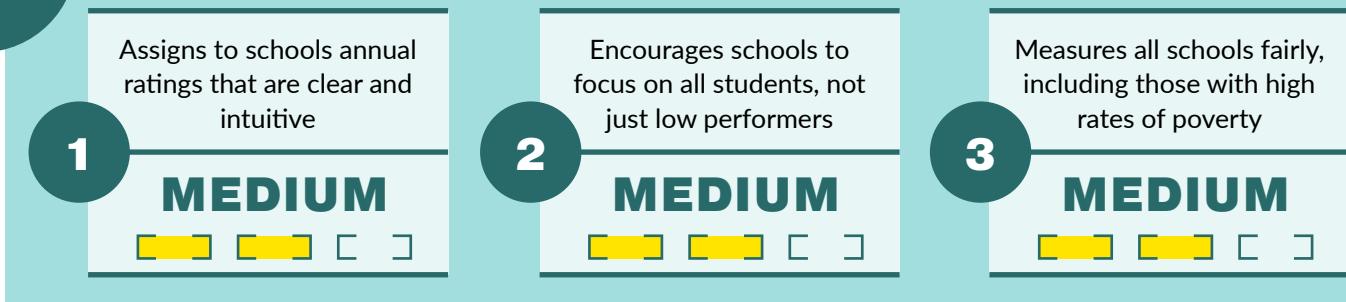
There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. Nevada receives a **weak** because it measures achievement with proficiency rates, which may encourage schools to focus on pupils near the proficiency cutoff—and because a measure of growth for all students constitutes just 17.5 percent of schools' annual ratings, which is apt to lead schools to disregard the educational needs of higher-achieving children, especially those in high-poverty schools.

3

Is the rating system fair to all schools, including those with high rates of poverty?

Nevada is **medium** here because academic growth will constitute 35 percent of schools' annual ratings—split evenly between growth for all students and growth to proficiency. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

NEW JERSEY



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether New Jersey's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on May 3, 2017,³⁹ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

New Jersey's plan earns a **medium** on this point because it proposes to use text labels for schools' annual ratings. Although the proposed labels are easy to understand, in isolation each label fails to communicate how much better or worse a given school could do (it's not instantly clear, for example, how good or bad "meets target" is; in New Jersey it's the middle of three labels, which also include "exceeds target" and "below target"). Thus this model fails to convey immediately to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

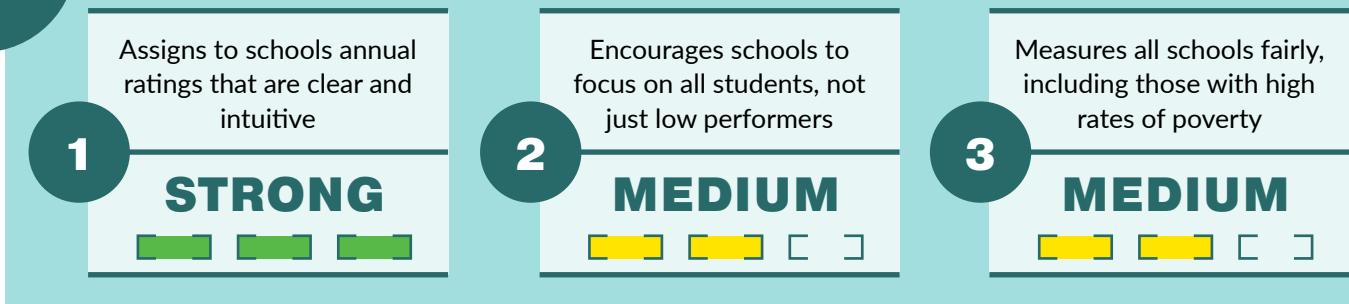
There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. New Jersey receives a **medium** because—despite measuring achievement proficiency rates, which might encourage schools to focus on pupils near the proficiency cutoff—a measure of growth for all students constitutes 40 percent of schools annual ratings, which should at least partially encourage schools to heed the educational needs of every child.

3

Is the rating system fair to all schools, including those with high rates of poverty?

New Jersey is **medium** here because academic growth will constitute 40 percent of schools' annual ratings—all of which is a measure of growth for all students. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

NEW MEXICO



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether New Mexico's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on April 11, 2017,⁴⁰ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

New Mexico's plan is **strong** on this point because it proposes to use a hundred-point system for schools' annual ratings. This model immediately conveys to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

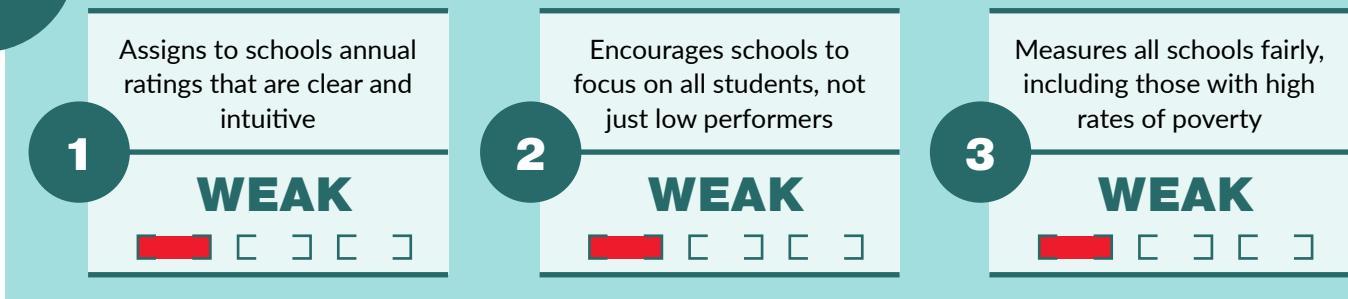
There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. New Mexico receives a **medium** because—despite measuring achievement proficiency rates, which might encourage schools to focus on pupils near the proficiency cutoff—student-level growth for all students constitutes 42 percent of schools annual ratings, which should at least partially encourage schools to heed the educational needs of every child.

3

Is the rating system fair to all schools, including those with high rates of poverty?

New Mexico is **medium** here because student-level growth for all students will constitute 42 percent of schools' annual ratings. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

NORTH DAKOTA



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether North Dakota’s proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on May 5, 2017,⁴¹ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

North Dakota receives a grade of **weak** because it proposes a “dashboard” approach, which comprises myriad data points and no bottom line. This is a mistake because such systems are difficult to understand and do not immediately convey to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

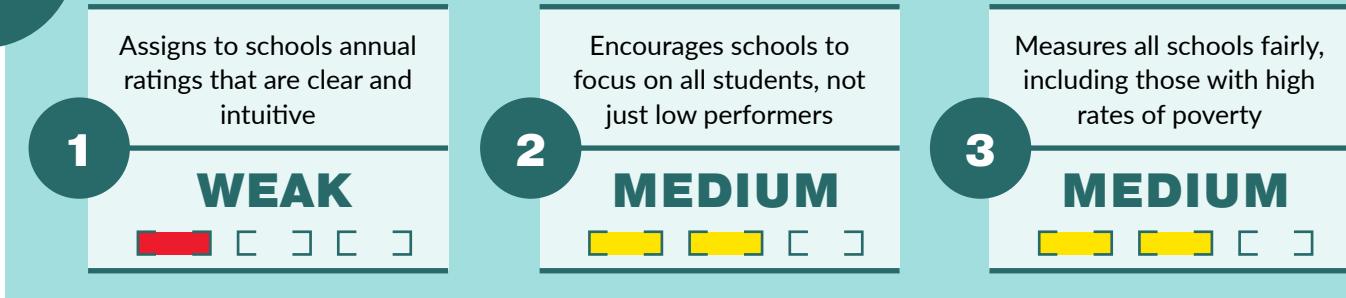
There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. North Dakota receives a grade of **weak** because it measures achievement with proficiency rates, which may encourage schools to focus on pupils near the proficiency cutoff—and because a measure of growth for all students constitutes just 30 percent of schools’ annual ratings, which is apt to lead schools to disregard the educational needs of higher-achieving children, especially those in high-poverty schools.

3

Is the rating system fair to all schools, including those with high rates of poverty?

North Dakota is **weak** here because academic growth will constitute just 30 percent of schools’ annual ratings—all of which is a measure of growth for all students. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

OREGON



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether Oregon's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on May 3, 2017,⁴² as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Oregon's plan is **weak** on this point because it has decided to forgo annual ratings entirely. This is a mistake. For more than two decades, school ratings have been at the heart of state accountability systems—and for good reason. Easy-to-understand labels, such as A–F letter grades, provide clear signals to parents, citizens, and educators about the quality of a school and can nudge systems toward improvement.

2

Does the rating system encourage schools to focus on all students?

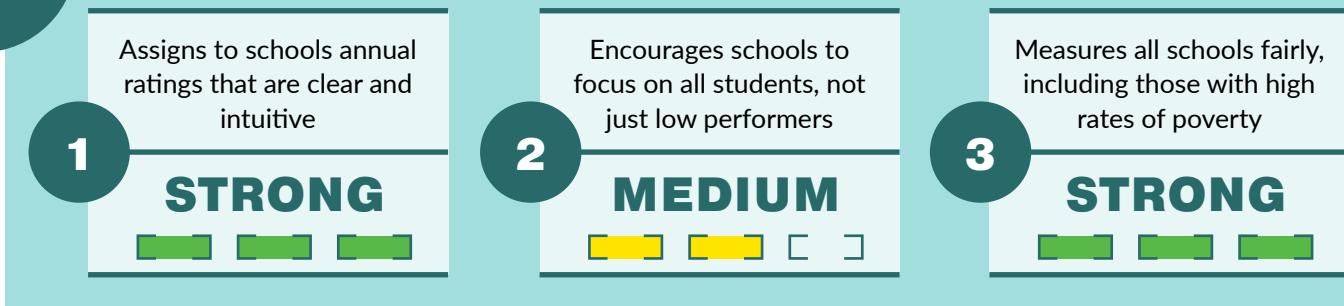
There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. Oregon receives a **medium** because—despite measuring achievement proficiency rates, which might encourage schools to focus on pupils near the proficiency cutoff—a measure of growth for all students constitutes 44 percent of schools annual ratings, which should at least partially encourage schools to heed the educational needs of every child.

3

Is the rating system fair to all schools, including those with high rates of poverty?

Oregon is **medium** here because academic growth will constitute 44 percent of schools' annual ratings—all of which is a measure of growth for all students. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

TENNESSEE



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether Tennessee's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on May 3, 2017,⁴³ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Tennessee's plan is **strong** on this point because it proposes to use an A-F system for schools' annual ratings. This model immediately conveys to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. Tennessee receives a **medium** because—despite measuring achievement proficiency rates, which might encourage schools to focus on pupils near the proficiency cutoff—a measure of growth for all students constitutes 35 percent of schools annual ratings, which should at least partially encourage schools to heed the educational needs of every child.

3

Is the rating system fair to all schools, including those with high rates of poverty?

Tennessee is **strong** here because academic growth will constitute 80 percent of schools' annual ratings—35 percent growth for all students and 45 percent growth to proficiency. Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.

VERMONT



ESSA grants states more authority over their school accountability systems than did NCLB. Three of the most important improvements states can make are to: (1) assign to schools annual ratings that are clear and intuitive for parents, educators, and the public; (2) encourage schools to focus on all students, not just their low performers; and (3) measure and judge all schools fairly, including those with high rates of poverty.

To determine whether Vermont's proposed ESSA accountability system accomplishes these three objectives, this analysis evaluates its state plan, as submitted to the U.S. Department of Education on May 3, 2017,⁴⁴ as explained below.

1

Are the labels or ratings for schools clear and intuitive for parents, educators, and the public?

Vermont's plan is **strong** on this point because it proposes to annually rate schools with a system comprising four numerical levels, accompanied by clear text labels. This model immediately conveys to all observers how well a given school is performing.

2

Does the rating system encourage schools to focus on all students?

There are two primary ways for state accountability systems to encourage schools to focus on all students: (1) use a performance index and/or scale scores in place of proficiency rates when measuring achievement and (2) measure the growth of all students. Vermont receives a grade of **strong** because those two components constitute 80 percent of schools' annual ratings. A performance index counts for half of that, which encourages schools to look beyond those pupils who are near the cutoff for proficiency. And a measure of growth for all students constitutes the other 40 percent of schools' summative ratings, which should also lead schools to heed the educational needs of every child.⁴⁵

3

Is the rating system fair to all schools, including those with high rates of poverty?

Vermont is **medium** here because academic growth will constitute 40 percent of schools' annual ratings—all of which is a measure of growth for all students.⁴⁶ Growth measures gauge changes in pupil achievement over time, independent of prior achievement, and are therefore less correlated with poverty, thus affording high-poverty schools the opportunity to earn positive ratings.



GLOSSARY

Average scale score:

The average score of all students within a school on a state-administered test used for accountability purposes and/or the average scores of student subgroups, such as economically disadvantaged students, African American students, Hispanic students, students with disabilities, and more.⁴⁷

Growth for all:

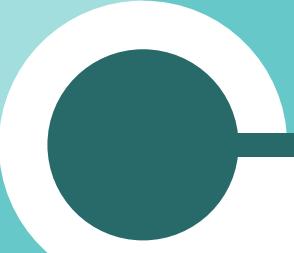
A measure of the academic progress of all students, regardless of their achievement level, based on the results of a state-administered test used for accountability purposes.⁴⁸ There are many ways states do this. Here are four common models, and brief explanations of how they work:⁴⁹

- **Categorical model:** compares the performance categories that students fall into from one year to the next
- **Multivariate value-added model:** estimates a school's contribution to students' academic growth by comparing their actual growth to their expected growth based on prior achievement and other factors
- **Student growth percentile model:** compares students to peers with similar achievement in the previous school year by ranking them based on their year-to-year growth
- **Vertical scale growth model:** tracks student growth within the same subject across grades, despite differences in test content and difficulty

Note, however, that these models only count as measures of growth for all students if states apply them to all students, which isn't always the case. For example, a state might use its model to gauge the progress of a subset set of students, such as low achievers or high achievers.

Growth to standard:

A measure of students' academic progress toward one or more absolute achievement standards, such as "proficient," based on the results of a state-administered test used for accountability purposes. Some growth to standard measures qualify as "growth for all" measures, insofar as they set appropriate goals for students regardless of their achievement level. Alternatively, growth to standard measures can apply to only a subset of students. For example, a "growth to proficiency" measure tracks the growth of students toward proficiency, effectively excluding those who are already achieving at a proficient level.



ENDNOTES

1. See, for example, Caroline Phenicie, "74 Interview: Senator Lamar Alexander on Keeping ESSA From Becoming The Next Obamacare," *The 74* (July 10, 2016), <https://www.the74million.org/article/the-74-interview-lamar-alexander-on-keeping-essa-from-becoming-the-next-obamacare>.
2. *ESSA State Plan Submission*, U.S. Department of Education, retrieved July 6, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/statesubmission.html>.
3. Rick Hess, "Of ESSA Plans and TPS Reports," *Education Week* (May 31, 2017), blogs.edweek.org/edweek/rick_hess_straight_up/2017/05/of_essa_plans_and_tps_reports.html.
4. 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://www.nber.org/papers/w10591.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>.
5. 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.
6. 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>.
7. Chester E. Finn, Jr. and Brandon L. Wright, *Failing Our Brightest Kids: The Global Challenge of Educating High-Ability Students* (Cambridge, MA: Harvard Education Press, 2015).
8. Ibid.
9. See, for example, Aaron Churchill, "Unless they want to flunk virtually all high-poverty schools, policymakers should go for growth," Thomas B. Fordham Institute (May 17, 2017), <https://edexcellence.net/articles/unless-they-want-to-flunk-virtually-all-high-poverty-schools-policymakers-should-go-for>.
10. For many reasons why proficiency rates are problematic, 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>.
11. Secretary DeVos' Letter to States regarding consolidated State plans, U.S. Department of Education, February 10, 2017, <https://www2.ed.gov/policy/elsec/guid/secletter/170210.html>.
12. *ESSA State Plan Submission*, U.S. Department of Education.
13. See *ESSA State Plan Submission*. The evaluation of Michigan's state accountability system is complicated by the fact that its consolidated state plan is not completely finished, despite having been submitted to the Department. Consequently, because this affects our ability to gauge how well it fulfills this report's three objectives, the state's ESSA plan receives a mark of Not Applicable (N/A) in all areas.

14. Note that recent feedback by the Department of Education on Connecticut's ESSA plan suggests that the use of scale scores alone for measuring achievement will not satisfy the law's requirements. See Interim Feedback Letter to Connecticut, U.S. Department of Education, June 30, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/ctprelimdetermltr.pdf>. Instead, any state using scale scores would have to also include a measure of proficiency—by, for example, using simple proficiency rates or a performance index.
15. David Figlio and Susanna Loeb, "School Accountability," in *Handbooks in Economics: Economics of Education*, Volume 3, ed. Eric A. Hanushek, Stephen Machin, and Ludger Woessmann (The Netherlands: North-Holland, 2011), 383–421, https://cepa.stanford.edu/sites/default/files/Accountability_Handbook.pdf.
16. Ibid.
17. Ibid.
18. Ibid.
19. Based on contents of state plans reviewed for this report. See *ESSA State Plan Submission*, U.S. Department of Education.
20. Figure 1 does not display Michigan because the state has not yet decided whether it will use annual school ratings. Therefore, the percentages, which are out of the total number of plans submitted, do not add up to 100 percent.
21. The color of each state's bar corresponds to the grade it received for this objective. Green indicates strong, yellow indicates medium, and red indicates weak. See the section on methods for more information. Figure 2 does not display Delaware or Michigan. Delaware received a weak grade because it uses proficiency rates and growth for all students will constitute less than 33 percent of schools' annual ratings, but the exact percentage has not yet been determined. Michigan has not yet decided whether it will use annual school ratings or specific weights for its indicators.
22. Forty percent is the maximum weight that can be assigned to Vermont's measure of growth for all students. Depending on circumstances, however, that weight can be as low as 35 percent for a given school.
23. Ibid.
24. The color of each state's bar corresponds to the grade it received for this objective. Green indicates strong, yellow indicates medium, and red indicates weak. See the section on methods for more information. Figure 3 does not display Michigan, which has not yet decided whether it will use annual school ratings or specific weights for its indicators.
25. See Arizona's state plan request, submitted to the U.S. Department of Education on May 9, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/azcsa2017.pdf>.
26. See Colorado's state plan request, submitted to the U.S. Department of Education on May 9, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/cocsa2017.pdf>.
27. Note that recent feedback by the Department of Education on Connecticut's ESSA plan suggests that the use of scale scores alone for measuring achievement will not satisfy the law's requirements. See Interim Feedback Letter to Connecticut, U.S. Department of Education, June 30, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/ctprelimdetermltr.pdf>. Instead, any state using scale scores would have to also include a measure of proficiency—by, for example, using simple proficiency rates or a performance index.

- 28.** See Connecticut's state plan request, submitted to the U.S. Department of Education on April 21, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/ctcsa2017.pdf>.
- 29.** Note that recent feedback by the Department of Education on Connecticut's ESSA plan suggests that the use of scale scores alone for measuring achievement will not satisfy the law's requirements. See Interim Feedback Letter to Connecticut, U.S. Department of Education, June 30, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/ctprelimdetermltr.pdf>. Instead, any state using scale scores would have to also include a measure of proficiency—by, for example, using simple proficiency rates or a performance index.
- 30.** See Delaware's state plan request, submitted to the U.S. Department of Education on April 3, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/decsa2017.pdf>. Delaware's weak grade on objective two is based on the state's decision to measure student growth in five ways (as noted in the state's write-up for objective three) that combine to constitute 35 percent of schools' annual ratings, only two of which are measures of growth for all. The state has not yet decided how much weight will be assigned to each of the five growth measures, but because only two of the five are measures of growth for all, we assess the state as counting growth for all students as less than 33 percent of schools' annual ratings. We therefore give the state a weak grade for this objective.
- 31.** See the District of Columbia's state plan request, submitted to the U.S. Department of Education on May 2, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/dccsa2017.pdf>.
- 32.** See Illinois's state plan request, submitted to the U.S. Department of Education on May 2, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/ilcsa2017.pdf>.
- 33.** See Louisiana's state plan request, submitted to the U.S. Department of Education on May 3, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/lacsa2017.pdf>.
- 34.** See Maine's state plan request, submitted to the U.S. Department of Education on May 4, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/mecsa2017.pdf>.
- 35.** See Massachusetts's state plan request, submitted to the U.S. Department of Education on May 10, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/macsa2017.pdf>, and Massachusetts's Additional Information, submitted to the U.S. Department of Education on May 10, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/maadditionalinfo.pdf>.
- 36.** Note that recent feedback by the Department of Education on Connecticut's ESSA plan suggests that the use of scale scores alone for measuring achievement will not satisfy the law's requirements. See Interim Feedback Letter to Connecticut, U.S. Department of Education, June 30, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/ctprelimdetermltr.pdf>. Instead, any state using scale scores would have to also include a measure of proficiency—by, for example, using simple proficiency rates or a performance index.
- 37.** See Michigan's state plan request, submitted to the U.S. Department of Education on May 3, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/micsa2017.pdf>.
- 38.** See Nevada's state plan request, submitted to the U.S. Department of Education on April 12, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/nvcsa2017.docx>.
- 39.** See New Jersey's state plan request, submitted to the U.S. Department of Education on May 3, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/njcsa2017.pdf>.
- 40.** See New Mexico's state plan request, submitted to the U.S. Department of Education on April 11, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/nmcsa2017.pdf>, and New Mexico's state plan appendices, submitted to the U.S. Department of Education on April 11, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/nmappendices17.pdf>.

- 41.** See North Dakota's state plan request, submitted to the U.S. Department of Education on May 5, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/ndcsa2017.pdf>.
- 42.** See Oregon's state plan request, submitted to the U.S. Department of Education on May 3, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/orcsa2017.pdf>.
- 43.** See Tennessee's state plan request, submitted to the U.S. Department of Education on May 3, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/tncsa2017.pdf>. Note that, for objective three, "growth to proficiency" is our term, which we use throughout the report, and define in the *Glossary*. Tennessee calls it "improvement toward proficiency."
- 44.** See Vermont's state plan request, submitted to the U.S. Department of Education on May 3, 2017, <https://www2.ed.gov/admins/lead/account/stateplan17/vtcsa2017.pdf>.
- 45.** Forty percent is the maximum weight that can be assigned to Vermont's measure of growth for all students. Depending on circumstances, however, that weight can be as low as 35 percent for a given school.
- 46.** Ibid.
- 47.** Morgan Polikoff, et al., "A letter to the U.S. Department of Education (updated July 14)," [morganpolikoff.com](https://morganpolikoff.com/2016/07/12/a-letter-to-the-u-s-department-of-education) (July 12, 2016), <https://morganpolikoff.com/2016/07/12/a-letter-to-the-u-s-department-of-education>.
- 48.** For more detailed information of measures of academic growth, see Katherine E. Castellano and Andrew D. Ho, *A Practitioner's Guide to Growth Models*, Council of Chief State School Officers (February 2013), <http://www.ccsso.org/Documents/2013GrowthModels.pdf>.
- 49.** Explanations of these models come from Michael J. Petrilli, David Griffith, Brandon L. Wright, and Audrey Kim, *High Stakes for High Achievers: State Accountability in the Age of ESSA*, Thomas B. Fordham Institute (August 2016), <http://edex.s3-us-west-2.amazonaws.com/08.31%20-%20High%20Stakes%20for%20High%20Achievers%20-%20State%20Accountability%20in%20the%20Age%20of%20ESSA.pdf>.