

# WHAT GETS MEASURED GETS DONE: IMPROVING PERFORMANCE MEASURES IN THE NEXT GENERATION OF ACCOUNTABILITY UNDER ESSA



By Douglas N. Harris and Lihan Liu, *Tulane University*

## Overview

Over the past three decades, state governments have increasingly held schools accountable for their performance, especially student achievement levels. Many states, as a result of the new federal ESSA law, are also revising the ways they measure and use school performance to assign school performance ratings, such as A-F letter grades.

Most of this brief focuses on the question: How would school ratings change if states measured school performance differently? In particular, how much would school performance ratings change if we added measures, like college entry, that are strong predictors of students' long-term life success? Also, how much would school performance ratings change if we focused not on outcome levels but on schools' contributions to those outcomes, sometimes called "value-added"? We address these questions using data from Louisiana and find:

**Corresponding technical paper:** Douglas N. Harris and Lihan Liu. "What Gets Measured Gets Done: Multiple Measures, Value-Added, and the Next Generation of Accountability under ESSA." Education Research Alliance for New Orleans, 2018.

- If policymakers measured high school performance not only with test scores and graduation levels but also with college entry levels, then our analysis suggests that 28.6% of high schools in Louisiana would receive different performance ratings (e.g., moving from a letter grade of F to D).
- If school performance measures were based on a 50-50 mix of achievement levels and achievement value-added, instead of levels alone, then 24.2% of elementary schools and 32.9% of high schools in Louisiana would change performance categories.
- Value-added can also be used to evaluate school performance on outcomes other than achievement. If high school performance measures were based only on graduation, 22.1% of Louisiana high schools would change performance categories if performance was measured by a mix of graduation levels and value-added, instead of graduation levels alone. If high school performance was evaluated solely on college entry, 30.7% of Louisiana high schools would change categories if performance was measured by a mix of college entry levels and value-added, instead of college entry levels alone.

---

## BACKGROUND: NCLB, ESSA, AND SCHOOL ACCOUNTABILITY

Accountability for student outcomes represents arguably the most important education policy trend of the past two decades. Many states instituted accountability plans during the 1990s, and in 2001, Congress passed President George W. Bush's signature proposal, *No Child Left Behind* (NCLB). Among other things, the law subjected schools in all states to a gradually intensifying cascade of interventions if those schools were not making Adequate Yearly Progress toward the goal of 100% proficiency. While the focus on test scores has continued in law and in practice, the *Every Student Succeeds Act* (ESSA), which was signed into law in 2015, has eliminated the 100% proficient goal and given states more flexibility over how performance is assessed.

Throughout a quarter-century of state and federal changes in test-based accountability, at least two problems have partially persisted: First, accountability policies have focused narrowly on student test scores. Second, states have focused on achievement levels (or achievement "status"). Many NCLB advocates consider the focus on test scores to be appropriate because the main purpose of schooling is to produce academic achievement, and research suggests the United States is falling behind other countries on standardized tests, which measure skills that contribute to economic growth. Moreover, some argue that schools pursue too many goals—they are "a mile wide and an inch deep"—and should focus more on the basics. The debate here is a fundamental and philosophical one; what we value should drive what outcomes we use to measure school performance.

*“... what we value should drive what outcomes we use to measure school performance.”*

Concerns about these core provisions of NCLB are partly what led to the new law. ESSA reflects a view that schools should pursue a larger number of goals because we expect them to address a wide variety of students' needs. Also, other academic and school-age outcomes—particularly years of education, including college—are at least as predictive of students' long-term outcomes as test scores. Limiting school performance measures to just test levels is likely to lead schools to focus less on high school graduation and other

- For the above analyses, the results in New Orleans' elementary schools mostly mirror the results across Louisiana. However, New Orleans' high schools are more likely to change performance categories than Louisiana high schools when using value-added to high school graduation and college entry.
- We estimate the practical impact of shifting toward value-added by simulating a state policy of closing low-performing schools for four consecutive years, similar to the policy used in New Orleans during 2009-2014. Switching from test score and high school graduation rate levels-only to equal weight on levels and value-added when choosing which schools to close would increase annual student achievement levels for the bottom fifth of all schools statewide by about 0.4 percentiles and increase the statewide high school graduation rate by 0.4 percentage points. When we include college entry alongside test scores and high school graduation, switching from levels-only to a mix of levels and value-added would increase the statewide college entry rate by 0.4 percentage points.

Since the choice of performance measures is important, we also ask an additional question: As part of their ESSA plans, how many states are planning to add college outcomes and value-added measures to their performance metrics?

- Only 24 states are planning to use value-added or a similar measure according to state ESSA plans, and only 8 of these states are planning to give value-added a weight of 40% or higher in their overall performance measures.
- While 18 states are planning to include college "readiness," only 3 states plan to use actual post-secondary outcomes as school performance measures.

If states seek to hold schools accountable for what they can control, and for those outcomes that are most predictive of students' long-term success, then most states' ESSA plans are still placing too little emphasis on value-added measures and outcomes like college entry.

What gets measured gets done. This is evident in our simulations of school closure and takeover, but other research also clearly shows that the measures matter in more subtle and indirect ways, such as when parents collect information to choose schools. If we can improve school performance measures, then our analysis shows that we can improve actual student outcomes.

---

student outcomes that affect human capital and social welfare. The focus on test scores also creates incentives to distort the scores. For these reasons, calls for “multiple measures” have been widespread, and ESSA encourages states to use measures beyond just test scores.

Whichever student outcomes are chosen, focusing on the levels of such outcomes has the effect of holding schools accountable for factors outside their control. Students start school at different baseline levels of academic ability, motivation, and family support. Moreover, these student characteristics are unequally distributed across schools. This creates a “starting gate inequality” that rewards schools that happen to serve advantaged students and punishes schools that serve students most in need.

One solution to this problem is to focus on value-added to student outcomes (i.e., to take into account students’ predicted outcomes when evaluating school success). Value-added eliminates the starting gate inequality through statistical adjustments and focuses performance measures on student growth, which is a more valid indicator of what schools actually contribute to student learning. It is also consistent with the cardinal rule of accountability—that we should hold schools accountable for what they can control. Research increasingly shows that value-added measures meet this rule better than outcome levels.

“ *Value-added eliminates the starting gate inequality through statistical adjustments and focuses performance measures on student growth, which is a more valid indicator of what schools actually contribute to student learning.* ”

Some critics have countered that value-added measures are unreliable and volatile, though we show below that this can be addressed by adjusting the value-added measures for schools with

few students and by averaging across years. Another potential concern is that value-added measures do not require schools with lower-scoring students to catch up with everyone else, perhaps giving the impression that there are different expectations for different students. This represents a partial misunderstanding of value-added and performance measurement generally, however, because value-added can be used without changing standards for any group of students. A third potential concern is the statistical complexity of some types of value-added measures, raising the issue of transparency, though there are simple ways to calculate value-added (for example, simple growth measures) and the intuition behind value-added is relatively straightforward even if the calculations are sometimes not.

While all measures have their limits, there is now little debate that including value-added measures, along with outcome levels, would improve the validity of school performance measures. However, our analysis of state ESSA plans shows that only 24 states are using true value-added or growth measures. This includes Louisiana, whose data are the basis for our later analysis. Only 8 of those states, however, give value-added a weight of 40% or higher in their overall performance measures. (See the technical report for the lists of specific states.) While any cutoff like 40% is somewhat arbitrary, this tells us whether the states that are considering value-added are giving it a substantial weight.

Most states plan to use “growth-to-target” type measures, either for the general student population or only for English Language Learners. In other words, they address the question, are students on track to reach some specific goal such as proficiency? This is a useful question, especially for parents who want to understand the paths their children are on and perhaps provide extra help. But helping parents understand what their children need is not the same as measuring school performance. The problem is that students who start far behind have to grow faster, meaning growth-to-target measures are extremely similar to achievement levels, which unfairly punish and reward certain schools.

Our analysis of the ESSA plans also shows that only three states—Connecticut, Rhode Island, and Vermont—mention plans to use actual college outcomes in measuring school performance (another 18 mention “college readiness”). This is noteworthy because college entry occurs only a few months after high school graduation, so it is

still partly within the control of schools, and because college entry is a stronger predictor of later life outcomes than test scores.

School performance measures remain an important issue given ESSA's continued calls for states to intervene in low-performing schools. The issue of school performance measurement is particularly salient in districts, such as New Orleans, that have large percentages of charter schools operating under performance-based contracts. In these cases, the main decision the government can make is whether to allow charter schools to continue operating. School performance measures play a large role in these decisions.

“*The issue of school performance measurement is particularly salient in districts, such as New Orleans, that have large percentages of charter schools operating under performance-based contracts.*”

In the remainder of this report, we use data from Louisiana to understand the potential impact of how we measure school performance.

## HOW DID WE CARRY OUT THE ANALYSIS?

Most of the data used in the analysis were provided by the Louisiana Department of Education (LDOE) and track enrollment, achievement, graduation, college entry, and other measures for essentially all individual students in Louisiana's publicly funded schools for the years 2006-2014.

When considering elementary student achievement, we examine scores from state standardized tests (LEAP and iLEAP) that are given in the spring to all students enrolled in grades 3-8. Since the testing system covers both elementary and middle school grades, we combine elementary and middle schools together throughout the analysis.

For high school student achievement, we analyze scores from the Graduate Exit Exam (GEE) and, in the later years of our study, End-of-Course tests (EOC), which students are required to take in order to graduate from high school. The testing regime shifted in the middle of the period we are studying, which is why there are two types of tests.

For this study, students are considered to be high school “graduates” if they complete some type of degree or credential (such as GED) on-time (within four years of the first entry into ninth grade). Data on students' college enrollment came from the National Student Clearinghouse, an organization that tracks enrollment in almost all colleges throughout the country. We count any type of college enrollment, including part-time and full-time enrollment in two- and four-year colleges.

With the above data, we calculate a single overall performance measure for each school and year. The test score levels measure represents a simple average of the end-of-year scores across subjects. When we add high school graduation, we create a composite measure that gives each included measure the same weight. For example, with two measures, each measure is weighted 1/2, and with three measures, each measure is weighted 1/3. (Since the various measures are on different scales, all the measures are first placed on “common scale” based on the statewide school average and variance.)

The analysis focuses on how changing the performance measures impacts the set of schools in each performance category. We place schools into five performance categories, mirroring the policy in Louisiana, which uses an A-F rating system (excluding letter grade E). We also place the same share of schools in each category as Louisiana had in 2014. As in most states, the lowest category—with a grade of F—has the fewest schools (8% percent of the total). This approach for creating both the performance measure and the performance category is deliberately much simpler than most state accountability systems (e.g., we ignore subgroup scores), allowing us to focus on the changes in the performance measures that are the subject of the analysis.

Value-added statistical methods are now widely accepted by researchers when analyzing test scores. We can describe the calculations in two connected steps: (1) predict student outcomes using prior student test scores, student demographics, participation in special education and English language learning, student mobility,

and other measures; and (2) subtract the predicted outcome from the actual outcome to obtain an estimate of schools' value-added. When students in a school do better than predicted, then that school has above average value-added.

One of the main concerns with value-added is that it is unreliable, that it changes a great deal from one year to the next, due to what statisticians call random error. To address this concern, all the analyses of value-added reported below use a four-year average of value-added, which reduces volatility.

Researchers have not focused nearly as much attention on applying value-added methods to outcomes such as high school graduation and college entry as they have on applying the methods to test scores. Unlike test scores, graduation and college entry only occur once, rather than every year. This makes estimating value-added more difficult for these outcomes because we have no prior-year measure to use in the first step of prediction, but we can predict high school graduation and college entry based on eighth grade test scores and student demographics using the same basic logic as value-added to test scores.

### HOW MUCH DO SCHOOL PERFORMANCE RATINGS CHANGE WHEN WE ADD STUDENT OUTCOMES?

Federal law requires that states use both test scores and high school graduation rates as high school performance measures, therefore we start by examining the effect of adding graduation levels to a system that already uses test score levels. We then go a step further and add college entry.

The accompanying technical report provides correlations between all the measures and shows, for example, that the relationship between test scores and graduation is strong and positive, but weaker than the relationship between high school graduation and college entry. The closer link between graduation and college entry is likely because high school graduation is a prerequisite to attend college. Students can have low test scores and still graduate, but they cannot easily go to college without graduating high school. This is important because the practical impact of adding measures depends on the correlation among the measures and the number of measures used. Other things equal, going from one to two measures will have a larger impact on school performance measures than going from four to five measures

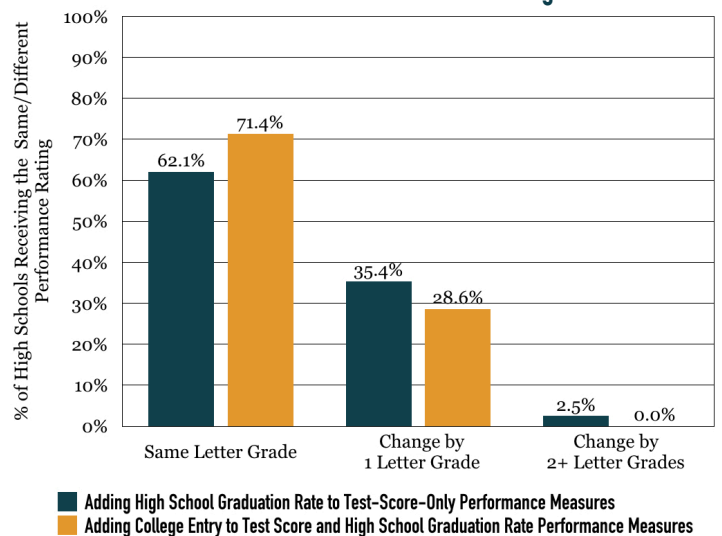
(see more on this in the technical report). Because all student outcomes tend to be correlated with one another, each additional measure adds progressively less information.

*“ Because all student outcomes tend to be correlated with one another, each additional measure adds progressively less information. ”*

Every time we add a measure, however, the weights of all the other measures change. For example, in going from one to two measures, the weight goes from 1.0 for the single measure to some split, such as 1/2, for each of two measures, and so on for more measures. This means that even if we add a measure that is perfectly correlated with an already-included measure, the change in the weights alone would still change the performance measures and ratings for some schools.

Figure 1 shows what happens to school performance ratings when we move from measuring performance by test score levels alone to measuring performance by test score levels and high school graduation rates.

**Figure 1. Effect of Adding High School Graduation and College Entry to Performance Measures on the Percentage of High Schools Receiving the Same/Different Performance Rating**

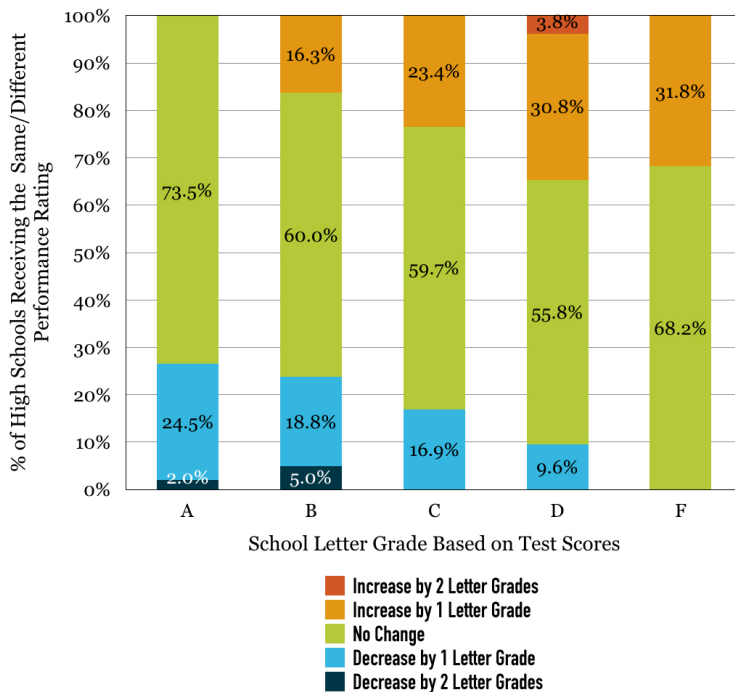


The blue bars show that 62.1% of schools end up with the same performance rating or grade when we add graduation levels to test score levels. The remaining schools change grades, though only 2.5% change more than one letter grade (e.g., from A to C). All states have added high school graduation as a factor when measuring high school performance because of federal requirements.

The orange bars in Figure 1 show results comparing performance measures using test scores and high school graduation with a measure that adds college entry (the three measures are equally weighted 1/3 each). The effect of adding this third measure is somewhat less than adding the second one, with 71.4% receiving the same grade. This pattern illustrates the idea that adding more measures that are correlated with one another has a progressively smaller effect on performance ratings.

School accountability also affects different types of schools in different ways. For example, schools with F grades are subject to intervention in Louisiana, due partly to federal requirements that started under NCLB and persist under ESSA. Therefore, Figure 2 below illustrates changes in performance ratings by initial letter grade.

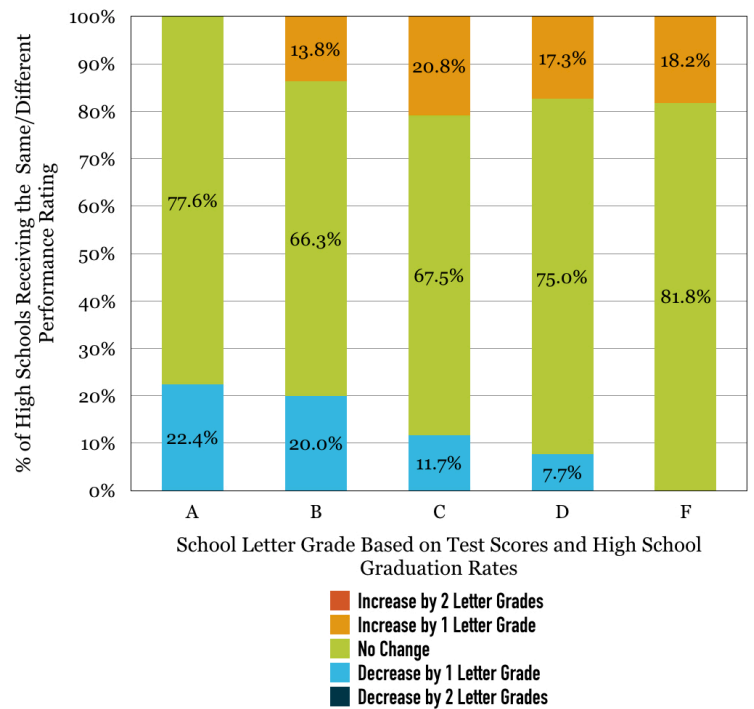
**Figure 2. Effect of Adding High School Graduation Rates to Performance Measures on the Percentage of High Schools Receiving the Same/Different Performance Rating, by Initial Letter Grade**



One-third of F-graded schools (based solely on test score levels) move to higher ratings when adding high school graduation rates (Figure 2). The percentage of schools changing categories is slightly lower for A schools (26.5%). The percentages of schools changing categories are always larger in the B-D range because, unlike A and F schools, they can move in both directions. D schools, in particular, seem quite sensitive to the addition of graduation rates.

While F schools were somewhat more sensitive than A schools to the addition of graduation rates, A schools are slightly more sensitive than F schools to the addition of college entry rates. Figure 3 shows that adding college entry on top of high school graduation rates and test scores leads 22.4% of A schools to move down a rating, while 18.2% of F schools move up.

**Figure 3. Effect of Adding College Entry Rates to Performance Measures on the Percentage of High Schools Receiving the Same/Different Performance Rating, by Initial Letter Grade**



The above figures show how sensitive school ratings are when high school graduation and college entry rates are added to school performance measures. Our results suggest that a large percentage of schools would be treated differently—by policymakers and parents, for example—if the method of measuring performance changed. The effect of adding college entry, in particular, is important to consider because of how well this predicts long-term life success and because

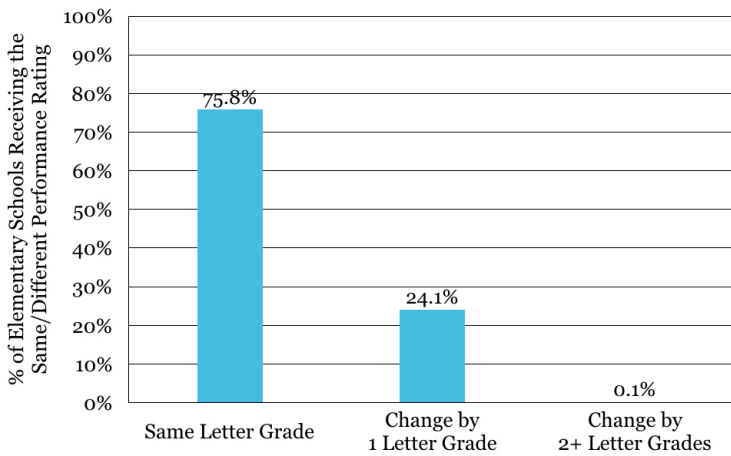
of the fact that job skill demands are increasing. Almost anyone can benefit from some form of college education, from vocational certificates at two-year colleges to bachelor's degrees.

## HOW MUCH DO ELEMENTARY/MIDDLE SCHOOL PERFORMANCE MEASURES CHANGE WHEN WE MOVE TO A MIX OF TEST SCORE LEVELS AND VALUE-ADDED?

One issue with current school accountability is that, in focusing mostly on outcome levels, performance measures do not account for where students start—the factors outside the control of the school. Below, we show how many schools change categories when performance measures use half test score levels and half test score value-added instead of test levels only. We follow this approach because even scholars who recommend using value-added also recommend including some level measures, too.

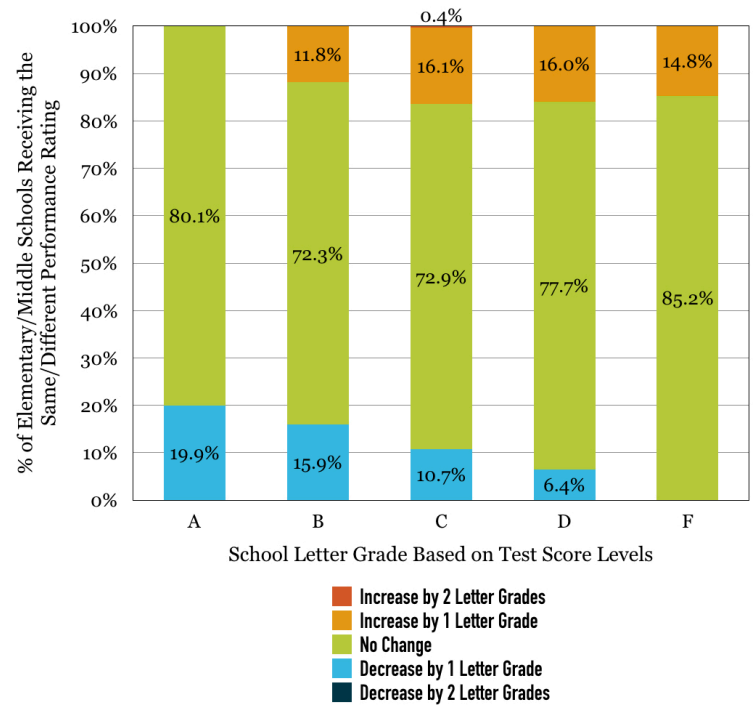
Figure 4 shows that 75.8% of elementary/middle schools in Louisiana receive the same performance rating using test score levels and using a mix of test score levels and value-added. This is higher than what we showed earlier for the percentage of schools that have the same performance rating when adding either high school graduation or college entry rates (see Figure 1).

**Figure 4. Effect of Switching Performance Measures from Test Levels to Half Levels and Half Value-Added on the Percentage of Elementary/Middle Schools Receiving the Same/Different Performance Rating**



The influence of giving equal weight to value-added depends, again, on schools' initial ratings. Figure 5 shows that 19.9% of A schools drop one letter grade. The effect of switching to a mix of levels and value-added is much smaller at the other end of the performance spectrum; 14.8% of F schools end up with higher ratings when performance measures include value-added.

**Figure 5. Effect of Switching Performance Measures from Test Levels to Half Levels and Half Value-Added on the Percentage of Elementary/Middle Schools Receiving the Same/Different Performance Rating, by Initial Performance Category**



The results are fairly similar when we look only at New Orleans, except that the New Orleans' results are a bit more stable: 86.4% of the city's elementary/middle schools stay in the same performance category.

## WHAT HAPPENS WHEN WE APPLY VALUE-ADDED ANALYSIS TO HIGH SCHOOL TEST SCORES, GRADUATION, AND COLLEGE ENTRY?

Most policy discussion about value-added focuses exclusively on student test scores. This is partly because value-added is similar to "growth," and with test scores, we can track students' scores from one year to the next and get a natural measure of student growth. However, we also described value-added earlier as a comparison of actual student outcomes with what we might predict based on students' prior outcomes and background. Value-added therefore can be applied to any student outcome, not just test scores.

In addition to high school test scores, we calculate school value-added for high school graduation and college entry where the predicted outcomes are based on free and reduced price lunch

eligibility, participation in programs such as special education, and test scores prior to high school. In general, we find that the relationship between outcome levels and value-added is closer with high school graduation and college entry rates than with test scores.

We compare schools' ratings based on test scores, graduation, and college entry levels with their ratings based on a mix of levels and value-added (Figure 6). Our results suggest that 32.9% of high schools change performance categories when switching from test levels to a mix of test levels and test value-added. The percentage of high schools changing categories is 22.1% when switching from high school graduation levels to a mix of graduation levels and graduation value-added, and 30.7% when switching from college entry levels to a mix of college entry levels and value-added. These results reinforce that shifting toward value-added has a significant impact on school performance measures, regardless of what outcome is used.

**Figure 6. Summary of Changes in School Ratings when Switching Performance Measures from Outcome Levels Only to a 50-50 Mix of Outcome Levels and Value-Added**

Change in Performance Rating Definition	% of Louisiana Schools that Change Ratings	% of New Orleans Schools that Change Ratings
<b>Elementary/Middle Schools</b>		
Switching from Test Levels Only to 50-50 Mix of Levels and Value-Added	24.2%	13.6%
<b>High Schools</b>		
Switching from Test Levels Only to 50-50 Mix of Levels and Value-Added	32.9%	0%
Switching from High School Graduation Levels Only to 50-50 Mix of Levels and Value-Added	22.1%	56.6%
Switching from College Entry Levels Only to 50-50 Mix of Levels and Value-Added	30.7%	66.6%

Figure 6 shows that, on a statewide level, the effect of switching from levels toward value-added is somewhat smaller than the effect of adding additional student outcomes (see Figure 1). It also shows that the effects are highly uneven across districts. As the last column shows, New Orleans' schools are much less sensitive to the use of value-added with test scores (for both elementary/middle and high schools), but they are more sensitive to the use of value-added with high school graduation and college entry.

There are two main reasons why New Orleans' high schools are generally more sensitive than the state as a whole to the use of value-added with high school graduation and college entry. The first is that New Orleans has more selective high schools than other districts, which creates large differences across schools in their predicted student outcomes. This reinforces the basic logic behind using value-added measures. When evaluating schools primarily on outcome levels, selective schools are rewarded simply for being selective, and the remaining schools are punished for serving students with low initial outcome levels. Shifting toward value-added therefore has a bigger impact on performance measures when some selective schools are present.

A second factor is the "ceiling effect," which refers to the general idea that some measures have a maximum. When schools are near the maximum level, it is more difficult for them to show value-added. High school and college entry rates are good examples because they can go no higher than 100%. (There are ways to address this, as discussed in the technical paper.) This issue is also related to high school selectivity, as selective schools are more likely to be near the ceiling and have low value-added. Since the ceiling effect mainly applies to high school graduation and college entry rates (and other rates such as daily attendance), this also helps explain why we see a different pattern with those two high school measures. This does not mean that value-added should not be used with rates, but only that policymakers should be careful with applying value-added to rates.

## HOW WOULD SWITCHING TO VALUE-ADDED AFFECT STATE INTERVENTIONS IN LOW-PERFORMING SCHOOLS?

The analysis above focuses on how changing school performance measures can affect schools' performance ratings. This section shows that aggressively intervening in low-performing schools, where performance is measured by outcome levels instead of levels and value-added, sacrifices student outcomes. Motivated by a similar policy in New Orleans, we simulate the effect of switching performance measures from test scores and graduation levels to mix of levels and value-added for test scores and graduation on actual school performance under a school closure/takeover policy. In this hypothetical policy, the state takes over the bottom five percent of schools every year, for each of four years, and replaces them with new schools (or move students to other existing schools) that are at the state average. A similar policy, without explicit



---

closure percentages, was implemented in New Orleans from 2009-2014, as part of the city's intense school reform effort. This policy is also similar in spirit to a fully implemented version of NCLB or ESSA, both of which have required state intervention in the bottom five percent of schools.

The figures above provide some indication of what would happen if policymakers intervened in low-performing schools based only on outcome levels. Given that some schools placed in the F category based on test levels are not really the least effective in generating achievement—see the percentage of schools switching categories in Figure 5—giving 1/2 weight to value-added in school performance measures will tend to target state-mandated interventions to the schools that are actually less effective at improving student scores. This means that giving some attention to value-added will, over the long run, improve school quality and student outcomes.

Our analysis shows that taking over the bottom five percent of elementary schools on the basis of a mix of test levels and value-added, as opposed to levels-only, would increase the average actual performance of the original bottom 20% of schools by about 0.4 percentiles (e.g., going from the 50th to the 50.4th percentile). We focus on the bottom 20% because these are the only schools that could be affected by such a policy.

For high schools, we start with an accountability system that includes only test scores and high school graduation (equally weighted). Switching from 100% levels on both measures to a 50-50 split of levels and value-added therefore means we have four equally weighted measures: test score levels, test score value-added, high school graduation levels, and high school graduation value-added (weighted 1/4 each). Moving to this new composite measure would increase actual school performance in the bottom 20% of schools by 0.4 percentiles and increase those schools' graduation rates by 0.4 percentage points.

Finally, in an accountability system that also includes college entry, switching from levels only to a mix of levels and value-added increases college entry in the bottom 20% by about 0.4 percentage points (the same as the high school graduation rate above).

The simulation shows how switching from a levels-only performance measures to a mixture of levels and value-added would affect actual school performance—the actual school quality that students experience.

## CONCLUSION

This brief focuses on two important principles for the design of school performance measures. The first is that we should hold schools accountable for factors they can control, using value-added. The second is that we should focus on student outcomes that are predictive of long-term life success, with measures such as college entry. These are not the only relevant principles, but they are important.

Our analysis measures the practical impact of applying these principles. The size and significance of these effects is a matter of judgment. Clearly, changing performance measures by themselves would not dramatically change actual school performance. On the other hand, the fact that the policy design seems to sacrifice student outcomes at all is noteworthy. Also, we are only able to capture the effect of closure/takeover, one of many mechanisms through which performance measures affect actual school performance. Our other research shows school performance ratings send signals to parents. While families consider a range of factors when choosing schools, school performance ratings are clearly among them. Schools with higher ratings will attract more students than other schools, and this, by itself, increases student outcomes.

School performance ratings also send signals to schools about how they are doing. If we send the wrong signals to schools about the success of their various initiatives, they are apt to make poor judgments about what is working and what is not and become frustrated by the unfair representation of their performance. The focus on outcome levels, in particular, gives schools incentives to select the best students and “push out” those who are not doing well. In contrast, value-added rewards schools for helping students no matter where they start.

Improving performance measures in these ways not only improves school quality in several ways, but also comes essentially free of cost, so even a small effect might be worthwhile. These effects of changing performance measures are also meaningful because they can add up over time. For example, even if we closed no additional schools in the simulation, the improvements in school quality would be experienced by future generations of students.

The new federal ESSA law helps address some of the concerns with NCLB, giving states flexibility both in using multiple measures and encouraging the use of value-added. Our hope is that this

analysis will shed light on the remaining weaknesses in state school accountability and highlight new possibilities for the next generation of performance measures. The design of policies affects how well they work, and this is no less true with accountability. What gets measures gets done.

## How is this Research Related to Other ERA-New Orleans Studies?

At ERA-New Orleans, we study the multifaceted reforms put in place in the city after Hurricane Katrina. Accountability based on performance measures, especially test scores, has been a major part of the reform effort because charter schools generally operate under performance-based contracts, and the state has regularly closed low-performing schools.

In our other work, we have addressed different aspects of accountability:

- In the study, *Extreme Measures: When and How School Closures and Charter Takeovers Benefit Students*, Whitney Bross, Douglas Harris, and Lihan Liu studied the effects of school closure and takeover on student outcomes in New Orleans, the policy that motivated the simulation exercise above. Their analysis reinforces the importance of using school value-added, rather than levels, when choosing schools for intervention. The authors find that when students move to schools with higher value-added, students experienced larger outcome gains.
- In *When Tenure Ends: Teacher Turnover in Response to Policy Changes in Louisiana*, Katharine Strunk, Nathan Barrett, and Jane Arnold Lincove studied the effects of Louisiana's policy that essentially ended tenure in the state and created more intense

accountability for teachers. Their main finding was that the policy led more teachers to leave the teaching profession. They are also in the process of extending this work to understand the effect on teachers who had different levels of performance, as measured by teacher value-added. (The concept of teacher value-added is the same as school value-added, but with a focus on the growth of students attributable to an individual teacher.)

- In *How Do Schools Respond to State Policies on Teacher Evaluation?*, Julie Marsh, Susan Bush-Mecenas, Katharine Strunk, Jane Arnold Lincove, and Alice Huguet studied the effects of Louisiana's required teacher evaluation system, which included a combination of value-added measures and classroom observations, a real world application of multiple measures. They found that schools varied widely in their response to the changes in teacher evaluation law. Some schools reacted by reflecting on instructional practice, while others either complied with the law or acted strategically.
- Finally, in ongoing work, we are studying how the distribution of school value-added is changing in New Orleans over time, due to both market accountability and the state intervention described in *Extreme Measures: When and How School Closures and Charter Takeovers Benefit Students*.

---

## About the Education Research Alliance for New Orleans

The mission of the Education Research Alliance for New Orleans (ERA-New Orleans) is to produce rigorous, objective, and useful research to understand the post-Katrina school reforms and their long-term effects on all students. Based at Tulane University, ERA-New Orleans is a partnership between university-based researchers and a broad spectrum of local education groups. Our Advisory Board includes (in alphabetical order): the Louisiana Association of Educators, the Louisiana Association of Public Charter Schools, the Louisiana Federation of Teachers, the Louisiana Recovery School District, New Schools for New Orleans, the Orleans Parish School Board, the Orleans Public Education Network, and the Urban League of Greater New Orleans. For more information, please visit the organization's website.

---

**EducationResearchAllianceNOLA.org**

---

### Contact Information

1555 Poydras Street  
7th Floor, Room # 701  
New Orleans, LA 70112  
(504) 274-3617  
ERANewOrleans@gmail.com

*An Initiative of*



## About the Authors

### Douglas N. Harris

Douglas N. Harris is the founding director of ERA-New Orleans, professor of economics, and the Schleider Foundation Chair in Public Education at Tulane University.

### Lihan Liu

Lihan Liu is a Senior Research Fellow at ERA-New Orleans. She holds a PhD in economics from the University of Wisconsin at Madison.