

Back to the Basics:

A plan to simplify and balance Ohio's school report cards

By Aaron Churchill

December 2017



Back to the Basics:

A plan to simplify and balance Ohio's school report cards

By Aaron Churchill

December 2017



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

Preface	1
Glossary and acronyms	3
Summary and policy recommendations	5
Analysis of Ohio's report card	11
Overall structure	11
Achievement	15
Progress	16
Graduation	19
Prepared for Success	21
Gap Closing	22
K-3 Literacy	24
In closing: A better balance	27
Appendices	28
Endnotes	31

Preface

Most of us can remember getting report cards as kids. Sometimes the grades would be a validation of a job well done; sometimes they were disappointing, considering all the effort we had made in class. Sometimes—let’s admit it—the grades were low but also fair, given the quality of our work or the lack of effort we had put into it.

Regardless of how we felt at the time, most of us recognize that report cards played an important, if not always pleasant, role in our education. We needed the feedback that they generated in order to know our strengths and weaknesses, and they were important to parents and guardians who could offer help when our grades signaled that greater support was needed.

Likewise, the educational health of our schools hinges in no small part on periodic reviews of how they’re doing. Over the past two decades, states have developed annual report cards intended to offer an objective view of the performance of individual schools and districts. What’s on those report cards is used by families making decisions about which schools to choose for their daughters and sons, by taxpayers seeking evidence that their dollars are being well spent, and by decision-makers tasked with holding schools accountable for student learning and—most importantly—helping them improve when necessary.

In some respects, Ohio’s report cards do a good job informing the public about school quality. Most notably, the user-friendly A-F grading system offers a clear view of performance that’s a long way from yesterday’s bureaucratic jargon (one recalls with chagrin the state’s old “continuous improvement” rating for many schools). For the most part, the grades are properly grounded in hard data taken from state assessments or college entrance exams. Several of the metrics rightfully encourage schools to pay attention to the achievement of all students, rather than focusing solely on children who haven’t yet cleared the “proficient” bar. Especially praiseworthy are the Performance Index—a weighted measure that awards additional credit for high achievers—and value added scores that account for the growth of individual pupils.

Yet Ohio still has ample room for improvement. For starters, Buckeye school report cards have gotten unwieldy in size, containing more than a dozen letter grades with a dizzying array of calculations used to generate these ratings. A second and somewhat related issue is that report cards rely too heavily on “status” measures—e.g., pupil proficiency or graduation rates—that largely correlate with demographics and prior achievement. As a result, Ohio unfairly downgrades high-poverty schools due in part to factors outside of their control. This violates what Douglas Harris, a Tulane University researcher, calls the “cardinal rule of accountability”: holding schools to account for that which they can do something about.

This paper offers suggestions to improve Ohio’s school report cards. Most of them would require legislative action—much of the report card framework is a matter of statute—while others could be handled by the State Board of Education or Ohio Department of Education (ODE). If these proposals were implemented, Ohio would move towards a simpler, fairer, and more balanced approach to grading districts and schools.

Why are we offering these proposals now? Didn't ODE just submit the state's Every Student Succeeds Act (ESSA) plan—including policies related to report cards—to the U.S. Department of Education? It's true that ODE recently did this to comply with the federal law. But ESSA's overarching intent is to devolve authority over education policy back to the states. As Senator Lamar Alexander—one of the chief architects of the law—declared, “The national school board era has come to an end.” State policymakers now have the ability to create report cards that make sense to Ohioans, including the flexibility to make changes to them when necessary. In addition, though ODE's ESSA planning process engaged many stakeholders, this wasn't the appropriate forum to make substantive revisions to school report cards. That's primarily the job of lawmakers, as the accountability framework behind the report card is primarily a matter of state law. If elected officials decide to make changes (within, of course, the basic parameters of ESSA), ODE can submit the requisite paperwork to Washington to amend the state plan.

After report cards were released this fall, media coverage suggested that several state legislators, including the Senate and House education committee chairs, would be open to reexamining this issue. That's good news. We agree that improvements are needed, yet they also need to be made thoughtfully. With these suggested changes, a much improved report card can become a reality.

Acknowledgments

Many members of the Fordham team helped to create this report. Special thanks to Michael J. Petrilli, Chester E. Finn, Jr., Chad L. Aldis, Jamie Davies O'Leary, and Jessica Poiner for their invaluable feedback during the drafting process. Also on the Fordham side, deepest thanks to Madison Yoder and Jeff Murray, who assisted in report production and dissemination. Thanks also to Pavita Singh, who copyedited the report, and Andy Kittles, who designed the layout. All errors are my own.

Aaron Churchill

Ohio Research Director

Thomas B. Fordham Institute

Glossary and acronyms

The present report contains various terms and acronyms that may not be familiar to readers. The following list aims to serve as a reference to the key terms (see also page 12, which contains a broad overview of school report cards).

Achievement is a report card component that indicates student performance on state exams; the term can also be used more broadly to describe student performance at a certain point in time.

Achievement levels refer to the categories of student performance on state exams. Ohio has five levels. From lowest to highest, they are: limited, basic, proficient, accelerated, and advanced. A sixth category—advanced plus—is also included for pupils taking above grade level exams.

ACT and SAT are college entrance exams that students take in high school.

Advanced Placement (AP) and International Baccalaureate (IB) are advanced programs that include exams which offer the potential of college credit based on students' test scores.

Components refer to the six main dimensions of the school report card; subcomponents refer to indicators within a larger component.

Economically disadvantaged (ED) students are generally identified on the basis of eligibility for free or reduced priced lunches (i.e., their family income is less than or equal to 185 percent of the federal poverty level). The percent of ED students in a school or district is an approximation of student poverty; certain higher-income schools may report 100 percent ED on the basis of the Community Eligibility Program, which allows eligible districts to designate all students as ED regardless of household income in order to provide meals at no cost to all students.

End of course exams (EOC) are state assessments given to high school students in English language arts, math, biology, U.S. history, and U.S. government.

Every Student Succeeds Act (ESSA) is the name of a federal law that governs various aspects of K-12 education, including states' school accountability policies; ESSA is a revision of No Child Left Behind or NCLB.

Gap Closing is a report card component that considers subgroup performance on state exams and graduation rates.

Graduation is a report card component that indicates how many students receive a high school diploma within a certain amount of time (i.e., four or five years after entering ninth grade).

Growth measure is a general term used to describe an indicator that depicts changes in student performance over time.

Indicators Met is a subcomponent within the Achievement component that considers schools' proficiency rates; a gifted indicator is also included.

K-3 Literacy is a report card component that indicates literacy improvements among students deemed off track for reading proficiency by the end of third grade.

Performance Index (PI) is a subcomponent within the Achievement component that awards more credit to schools when students score at higher achievement levels.

Prepared for Success (PFS) is a report card component that indicates the readiness of high school graduates for success in college or career.

Proficiency describes pupil achievement in terms of whether they meet (or do not meet) the state's proficiency standard.

Progress is a report card component that indicates student growth on state exams, as measured by value added scores.

Status measure is a general term used to describe indicators of student achievement or educational attainment at a certain point in time (such as proficiency or graduation rates).

Subgroups refer to groups of students who share similar characteristics; under federal law, Ohio identifies subgroups by race/ethnicity, economically disadvantaged status, students with disabilities, and English language learners. Ohio includes two other subgroups in its accountability system: Low-achieving students and students identified as gifted.

Summative rating is a composite rating that combines the various component ratings; also called an "overall" rating.

Value Added is a statistical model that estimates a school's contribution to student growth based on pupils' prior achievement.

Summary and Policy Recommendations

Offering an independent, objective appraisal of the academic performance of Ohio’s 600-plus districts and 3,000 schools, report cards are vitally important to families who make choices about the schools to which they send their children, to taxpayers who provide the resources that pay for those schools, and to policymakers and education leaders who hold schools accountable for student outcomes and help them improve when necessary.

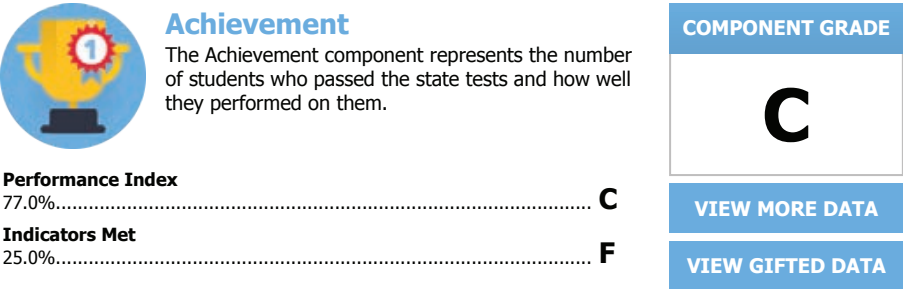
As many Ohioans know, report cards have evolved from the bare-bones version of the early 2000s to a modernized style. Today’s iteration is far superior, as older versions (following federal mandates that are no longer on the books) focused almost exclusively on proficiency rates and vague designations of “adequate yearly progress.” By contrast, current report cards include measures of student academic growth and indicators of post-secondary readiness. They also use intuitive A-F letter grades—a grading system that most people are familiar with. Starting next year, Ohio will add a summative (or overall) rating that combines the various report card components into a single, user-friendly grade.

Yet Ohio’s school accountability policies—and the report cards based upon them—still need work. As the editorial board of the *Record-Courier* (Portage County) recently opined, the report cards “are in danger of becoming irrelevant.”¹ Although concerns

about them are nothing new—school officials, in particular, are apt to grumble when grades are low—we also believe that modifications are warranted. In our view, Ohio’s present school report cards have two key problems: 1) they have become unnecessarily complex and 2) the ratings themselves reflect too strongly the backgrounds of students rather than the effectiveness of their schools. Both problems limit the usefulness of report cards for Ohio parents, community members, school leaders, and policymakers.

Ohio’s report cards now contain six major components—categories such as Achievement, Progress, and Gap Closing—along with eight subcomponents that feed into them. Unlike a student’s report card in which the academic subjects are widely understood, school ratings have terminology and calculations that are likely to be unfamiliar to people who work outside of education. As a result, laypeople may not be able to interpret the ratings. Consider an example: Figure 1 shows a district’s ratings on the Achievement portion of its report card. At face value, the F on Indicators Met raises immediate red flags, yet the C’s on the Performance Index and Achievement components seem to say something slightly different. Unless users can distinguish the meaning of these letter grades, they may not reach a sound conclusion about student performance on state exams. Indeed, recent focus groups and surveys with Ohioans suggest that, though people see the value of report cards, they tend to struggle to interpret the various ratings and data.²

Figure 1: An example of the Achievement portion from an Ohio district’s report card

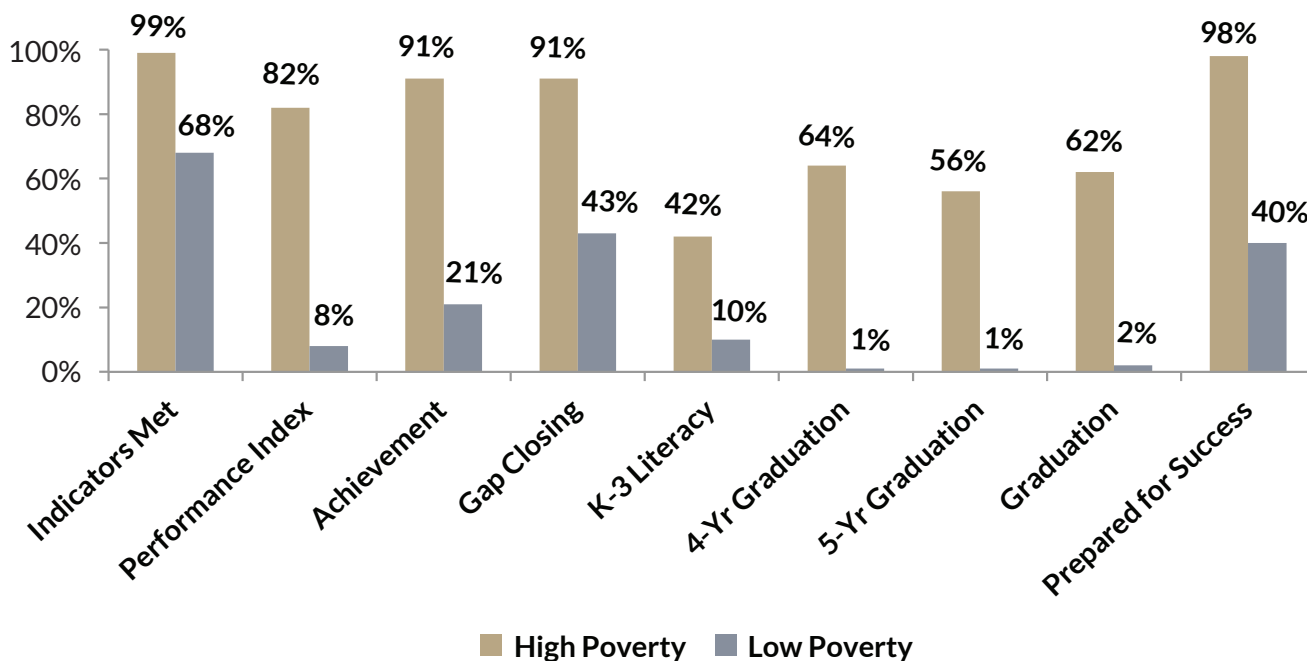


Adding to the challenge is that several components are redundant, signaling very similar information about a school or district. Ohio, for example, assigns not one but *three* ratings based on graduation rates. Duplicative components contribute little to the understanding of school quality but add to the complexity of report cards. Moreover, the large number of ratings also increases the chances that districts and schools receive low ratings: In 2016–17, 94 percent of Ohio districts received at least one D or F on their report card, and 51 percent received five or more of those letter grades. Unless the public can understand the nuances of each rating, this may leave an impression that most Ohio schools are “failing” even when other measures indicate success. Many people, not surprisingly, have determined that these ratings simply don’t meet the

smell test, serving to undermine rather than bolster confidence in report cards.

A second and interrelated problem is that Ohio’s report card system generates staggering numbers of low grades for high-poverty schools, grades that have as much or more to do with achievement gaps than school effectiveness. Last year, for example, 45 percent of economically disadvantaged students reached proficiency on their sixth grade reading exam, while 76 percent of their more-advantaged peers met that mark. Due in part to disparities such as this, high-poverty schools received more than 4,000 D’s or F’s—roughly four per school—on status-based measures that tend to correlate with pupil demographics and prior achievement (versus an average of about one for low-poverty schools).

Figure 2: Percentage of Ohio schools receiving D’s or F’s by poverty level, 2016-17



Note: Figures 2 and 3 define “high-poverty” schools as those reporting 66 to 100 percent economically disadvantaged (ED) students and “low-poverty” as those reporting 0 to 33 percent ED students. Schools between 33 to 66 percent ED are omitted for display purposes (the full results are displayed in Appendix B). For descriptions of these components, see the Glossary (page 3) and Overview of Report Cards (page 12).

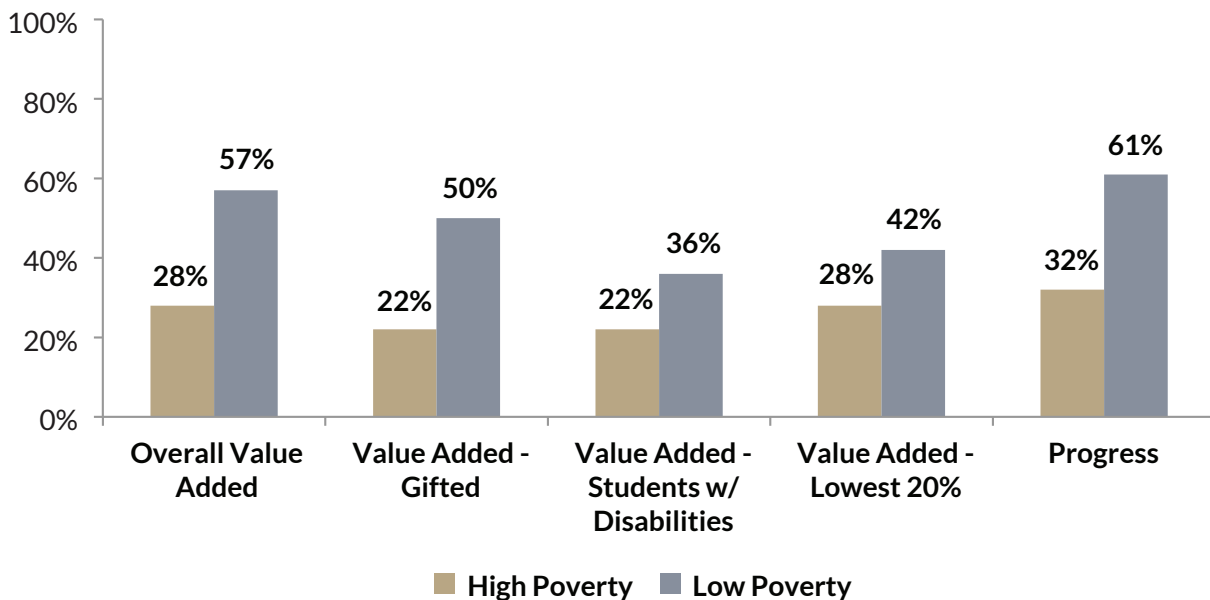
Figure 2 shows the extent to which high-poverty schools are disadvantaged in the school rating system. For instance, 82 percent of high-poverty schools received a D or F on the Performance Index—a weighted measure of student proficiency—while just 8 percent of low-poverty schools received those grades. Another way of viewing these ratings is that high-poverty schools have almost impossible odds of generating A’s and B’s (or often C’s) on many of Ohio’s report card measures.

To be certain, status measures have a place on school report cards. They tell us whether students measure up to state achievement or graduation standards, or exit high school ready for college or career. Many Ohioans likely view these data points as key indicators of the well-being of schools and the students

they educate. They *are* important markers as students prepare to exit into the “real world” and no one thinks that disparities in pupil performance should be swept under the rug. Indeed, we should acknowledge that most students in high-poverty schools have a long way to go if they are to leave school with the knowledge and skills necessary to pursue the American dream.

Yet Ohio has gone too far with status-based measures in the school grading system, meaning that it also de-emphasizes indicators of student academic growth over time (a.k.a., “value added”). Growth measures control for prior student achievement, giving schools credit for the gains students make, regardless of where on the performance spectrum they start out. This approach places schools with

Figure 3: Percentage of Ohio schools receiving A’s or B’s on growth measures by poverty level, 2016-17



Note: For descriptions of these components, see the Glossary (page 3) and Overview of Report Cards (page 12).

varying levels of poverty on a more even playing field, as growth is within a school’s control whereas pupil demographics are not. As Figure 3 shows, when we look at growth, we see that high-poverty schools receive fair numbers of A’s and B’s—about a quarter to a third of them—a proportion that tracks more closely with the fraction of low-poverty schools receiving those grades. The data indicate that all schools—no matter the backgrounds of their students—have a fair shot at performing well when growth is the focus.

Both status and growth measures deserve balanced treatment in school report cards, as they offer two important—and distinct—views of school performance. But Ohio policymakers haven’t yet struck the right balance. This is doubly important as Ohio is on the verge (starting in 2017-18) of issuing “summative” grades based on a rating system that places most of the weight on status-based measures (65 versus 35 percent, as Table 1 below shows). This formula will handicap high-poverty schools and lead to vast numbers of overall D’s and F’s in less advantaged communities.

These overall ratings may lead to unintended consequences. They’ll leave families in these areas with little ability to differentiate the performance of schools. Although such differentiation is crucial when choosing schools, it’s almost impossible to do when so many schools are lumped together as “failing.” It may reinforce the perception that all high-poverty schools are dysfunctional educational institutions when in fact there are some (though surely not enough) excellent, high-poverty schools in Ohio. This also has implications for charter school accountability: Because state policy restricts charters mainly to high-poverty communities, charter sponsors—held accountable for the performance of their schools—have gotten hammered on the

academic portion of their evaluations (this summative grading formula is already in use for sponsors). This could discourage these entities from approving promising startup schools in Ohio’s highest-need areas.

Table 1: Ohio’s overall school grading formula to be implemented in 2017-18

Component	Weight
Status components	
Achievement	20%
Gap Closing	15%
Graduation	15%
Prepared for Success	15%
Growth components	
Progress	20%
K-3 Literacy	15%

* * *

Ohio policymakers should modify the state’s school report cards to make them more streamlined, user friendly, fair, and balanced. These are the four areas where they should focus.

Recommendation 1: Reduce the number of A-F grades. Ohio now includes fourteen letter grades on report cards, with an overall rating scheduled for 2017-18. We suggest dialing back the number of letter grades to just six: an overall rating, plus these five components:

1. **Achievement**—based on Performance Index scores;
2. **Progress**—based on student growth (i.e., overall, or “schoolwide,” value added);

3. **Graduation**—based on the four-year graduation rate;
4. **Prepared for Success**—based on students meeting remediation-free targets on the ACT or SAT, receiving high-quality industry credentials, or earning passing scores on AP or IB exams or dual enrollment credits; and
5. **Equity**—based on whether various student subgroups meet Performance Index and growth goals. Subgroups include race/ethnicity, students with disabilities, and so forth.

This proposal eliminates the K-3 Literacy component, as well as the graded subcomponents of the report card. As we explain below, K-3 Literacy has limited ability to properly gauge school performance, as it relies on data from a range of fall diagnostic exams that may not generate comparable results. (Districts can select a diagnostic exam from a number of options.) The report card framework we propose here also meets the report card requirements under ESSA; Appendix A shows how.

Recommendation 2: Restructure the Gap Closing component and rename it “Equity.”

Instead of the byzantine calculations currently used to generate the Gap Closing rating—a fourteen-page guide is required to explain it³—lawmakers should base the component on a straightforward set of Performance Index and growth (or “value added”) indicators that schools’ subgroups are expected to meet. Table 5 on page 24 of this report shows a template for how this could be designed. The addition of subgroups’ student growth results—not a part of the current measure—also helps to balance the emphasis on achievement and growth in the summative ratings. We suggest renaming the component as well, as it includes subgroups that are not usually low-achieving (such as Asian American or gifted students). Instead, the com-

ponent name “Equity” would convey that schools are expected, under our proposal, to improve student learning across the board.

Recommendation 3: Create a summative rating formula that balances status and growth measures.

Summative or “overall” school grades roll up the various component ratings into a final evaluation of school performance. This rating offers the public the simplest and cleanest signal of the general health of a school and may be the only grade that some users look at. It’s important to get the summative rating formula right, striking a sound balance between status and growth measures. If state lawmakers adopt the two recommendations above, they could revise the summative weights in the following way: For elementary and middle schools, they could equally weight the three relevant components (Achievement, Progress, and Equity). A similar equal-weighting approach across all five applicable components could be employed for high schools and school districts. Other states, including Arizona, Colorado, Illinois, and Tennessee, are shifting to a more balanced approach to summative ratings, and Ohio should move in that direction as well.⁴

Recommendation 4: Report a wide range of informative data about schools, but don’t cram it all into the formal report card system.

Hard copies of Ohio’s school report cards are now twenty-five to thirty pages. Should schools or districts bother—and they absolutely should—to send home physical report cards, creating hundreds, if not thousands, of documents this bulky would be painfully burdensome and costly. Many of these pages, however, include supplemental data that most Ohioans are apt to skip over—e.g., statistics on extended graduation rates, mobility rates, instructional staff, and longer-term trends and comparisons to various averages

(Figure 5 on page 14 shows an example). While a wide range of school-specific information should be available to the public via online “data dashboards” or other publicly available data systems, policymakers should ensure that the report cards themselves focus on essentials.

* * *

We recognize that no report card can capture all that happens inside schools—some intangibles are simply impossible to gauge and compare on an objective basis. Thoughtful people can also disagree about the details of report cards, such as the relative importance of various metrics or whether certain groups of students should receive greater attention than others. At the end of the day, Ohio should keep its report cards anchored in objective data—yes, for better and worse, that often means standardized exams—though the ratings based on them should be used more as a flashlight than a hammer against schools. At the same time, Ohio policymakers need to devise more balanced report cards, particularly by creating a better blend of status and growth measures. By adhering to commonsense notions such as “keep it simple,” state leaders can design and deploy report cards that work better for all Ohioans.

Analysis of Ohio’s Report Card

Overall structure

Starting with the 2012–13 school year, Ohio has gradually implemented its current version of school report cards. These retained several earlier elements, such as the use of proficiency rates, the Performance Index, and value added measures.⁵ But it also shifted Ohio to an A–F grading system for schools and added new components such as Gap Closing, Prepared for Success, and K–3 Literacy, while also eliminating several out-of-date elements, including the adequate yearly progress structure and overall designations such as “effective” and “continuous improvement.”

The present report cards have a hierarchical structure. Sitting at the top is the overall school grade that is slated to be used for the first time at the end of the 2017–18 school year. Below this are six main components: Achievement, Progress, Graduation, Gap Closing, K–3 Literacy, and Prepared for Success. These components form the direct basis of the overall grade. Below these are eight graded subcomponents that feed into the component grades. For example, the Performance Index and Indicators Met subcomponents are used to generate the Achievement grade. Table 2 provides an overview of the six main components and their subcomponents (more detailed explanations of the components and grading system follow in the analysis below). Figure 4 then shows the front page of a report card, while Figure 5 displays an example of a more detailed page on student achievement that is included behind the front page for users who want to view more data.

Table 2: Summary of Ohio's current school report card

Performance Indicator	2012-13	2013-14	2014-15	2015-16	2016-17	Brief Description
Overall Grade	Not Graded	Not Graded	Not Graded	Not Graded	Not Graded	Composite of the report card components; overall grades are expected in 2017-18. The overall grade is determined through a weighting and points system (see Table 1 above for the weights; points are earned by schools based on their component grades).
Achievement	Not graded	Not graded	Not graded	Graded	Graded	Composite of PI and Indicators Met grades (weighted 75% on PI and 25% on Indicators).
Performance Index	Graded	Graded	Graded	Graded	Graded	Weighted measures of student achievement, with more weight given to pupils who achieve at higher levels.
Indicators Met	Graded	Graded	Graded	Graded	Graded	Proficiency rate on each grade-subject exam (26 possible indicators). In addition, schools are evaluated on a gifted indicator yielding 27 total indicators.
Progress	Not graded	Not graded	Not graded	Graded	Graded	Composite of the overall and subgroup value added measures (weighted 55% on overall and 45% on the subgroup results listed below).
Value Added-Overall	Graded	Graded	Graded	Graded	Graded	Growth estimate based on gains of all tested students.
Value Added-Gifted	Graded	Graded	Graded	Graded	Graded	Growth estimate based on gains of students identified as gifted in math (on math exams), reading (on ELA exams), or superior cognitive (on both math and ELA exams).
Value Added-Students with Disabilities	Graded	Graded	Graded	Graded	Graded	Growth estimate based on gains of students with disabilities who do not take alternative assessments.
Value Added-Lowest Achieving	Graded	Graded	Graded	Graded	Graded	Growth estimate based on the gains of students within the lowest 20 percent in achievement statewide.
High School Graduation	Graded	Graded	Graded	Graded	Graded	Composite of the four- and five-year graduation rates (weighted 60% on four-year and 40% on five-year).
Four-Year Rate	Graded	Graded	Graded	Graded	Graded	Percentage of students who earn a diploma within four years of entering ninth grade.
Five-Year Rate	Graded	Graded	Graded	Graded	Graded	Percentage of students who earn a diploma within five years of entering ninth grade.
Gap Closing	Graded	Graded	Graded	Graded	Graded	Proficiency rates of student subgroups (i.e., students by their race/ethnicity, students with disabilities, and several other subgroups). Also known as Annual Measurable Objectives (AMOs).
K-3 Literacy	Not Graded	Not Graded	Graded	Graded	Graded	Percentage of K-3 students who go from being not on track to on track in reading proficiency on fall diagnostic tests.
Prepared for Success	Not Graded	Not Graded	Not Graded	Graded	Graded	Rating is based on various high school measures that include remediation-free rates on ACT/SAT exams, honors diplomas, industry credentials, AP/IB results, and dual enrollment credits.

Figure 4: Example of the front page of Ohio's school report cards

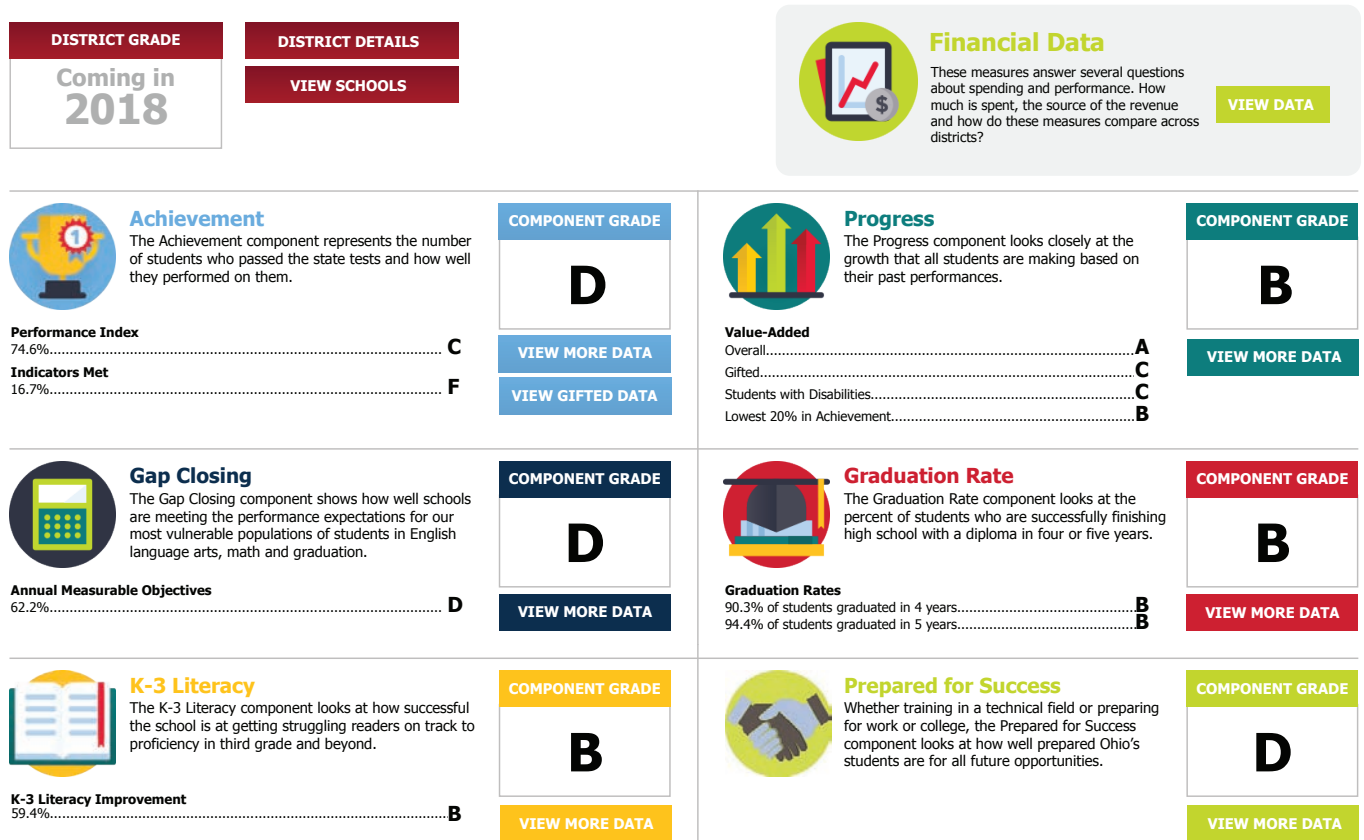
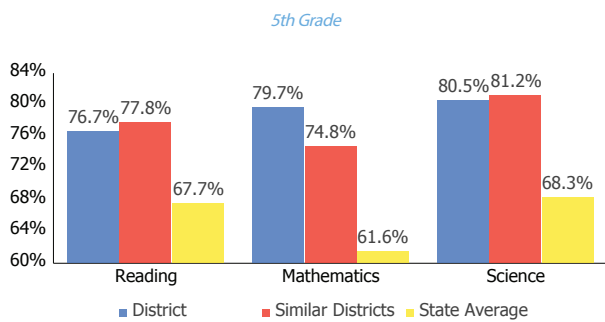
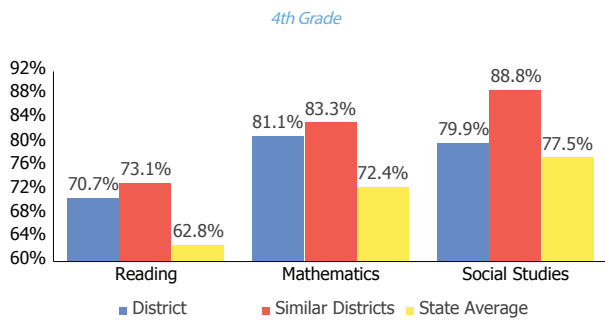
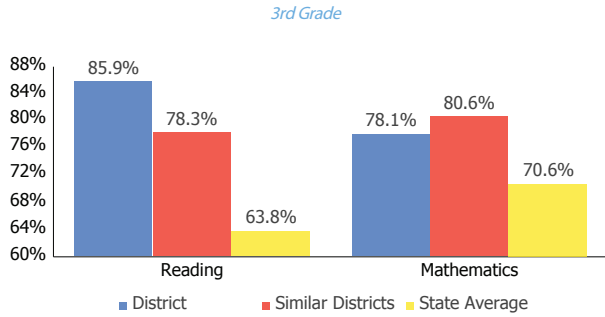
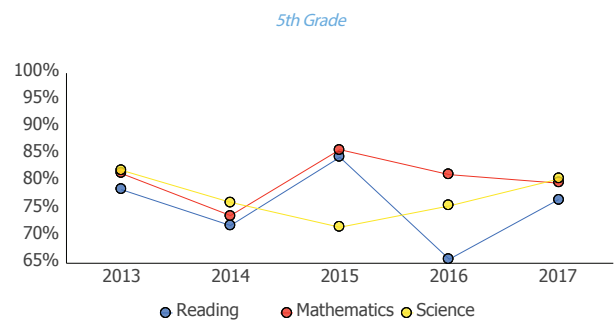
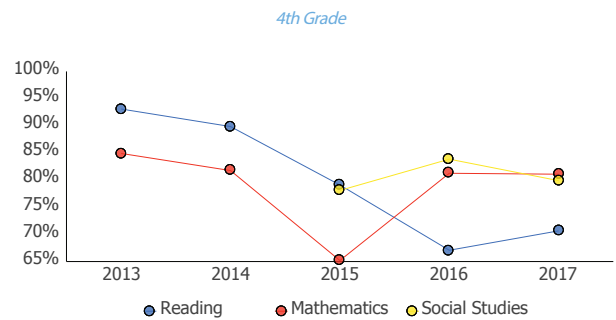
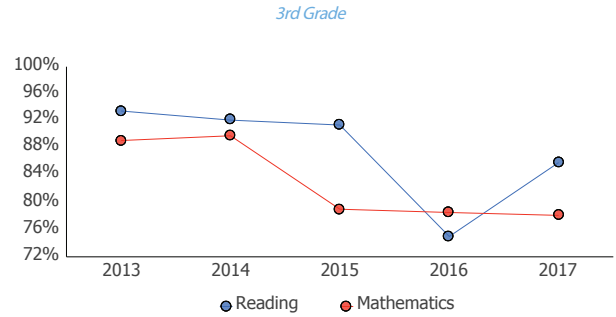


Figure 5: Example of a report card page showing more detailed data on student achievement

Achievement Levels by Grade



Proficient Percent Trend by Grade



Achievement

This component is based on schools' proficiency rates and Performance Index scores, measures that capture what students know and are able to do at a certain point in time. The Performance Index (PI) is a composite measure that awards additional credit—or weight—when students reach higher achievement levels on state exams. Table 3 displays these levels and their respective weight in the PI computations.⁶ Students' scaled scores on state exams are translated into these categories to facilitate a clearer understanding of their performance.⁷

Table 3: Ohio's achievement levels and weights used in the Performance Index calculation

Achievement Level	Weight
Advanced	1.2
Accelerated	1.1
Proficient	1.0
Basic	0.6
Limited	0.3

The Indicators Met subcomponent, in contrast, looks at whether students meet or do not meet Ohio's proficiency standard in tested grades and subjects (there is an additional indicator that uses data on gifted students). For example, a school receives credit for meeting the fourth-grade math indicator, if its proficiency rate is greater than 80 percent—the minimum rate to meet this indicator.⁸ PI is more akin to a weighted grade point average, while Indicators Met is more like a pass-fail rating. That said, these indicators are more similar than different—they both gauge student achievement, just from slightly different angles.

Both subcomponents feed into the larger Achievement component grade through a points and weighting system that places 75 percent weight on PI and 25 percent on Indicators Met. The table below displays an example of the computation, which involves three steps. First, ODE takes the raw data and converts it into points, which in turn determine the subcomponent letter grades.⁹ Second, a weighted average across the two subcomponents is calculated—3.06 points in the example below. Finally, that average is converted into a letter grade for the Achievement component.¹⁰

Table 4: Calculation of Achievement component letter grade based on Ohio's grading system and hypothetical performance data

	Raw Data	Points	Weight	Subcomponent Grade
Performance Index (% of PI points)	81%	3.25	0.75	B
Indicators Met (% of Indicators)	74%	2.50	0.25	C
Weighted Average Points		3.06		
Achievement Grade		C		

Comments

Student achievement remains an important part of school report cards; families and citizens deserve clear markers of how students currently perform against state academic standards. After all, ensuring that all students are on track to exit high school fully ready for their next step is a key goal of school accountability. Nevertheless, achievement measures should be used with care, as they do not account for students' prior achievement. As a result, the ratings and data from this component should be viewed as reflecting current pupil performance—not necessarily the instructional effectiveness of the school itself.

Ohio should maintain a rating based on student achievement. But the dual PI and Indicators Met ratings—along with the composite Achievement grade—communicate redundant information about how students fare against state standards. As a result, schools tend to receive similar grades on these metrics, and low (or high) grades can “pile up” on schools. For instance, the Cleveland school district—serving primarily low-income students—received F-F-F ratings on PI, Indicators Met, and Achievement on its 2016-17 report cards. This also happens at a school level: Sixty-nine out of 102 district schools in Cleveland received across the board F's on these metrics. The multiple ratings that signal similar things about student achievement could be seen as unnecessarily severe for high-poverty districts and schools.

There are tradeoffs to consider when deciding whether to use PI scores versus proficiency rates (i.e., Indicators Met). In a high-stakes accountability setting, PI scores tend to encourage schools to widen their attention to all students, as it offers them additional credit when students reach higher levels. Meanwhile, relying on proficiency rates creates incentives for schools to focus narrowly on stu-

dents just below the proficiency bar.¹¹ Proficiency rates have a more straightforward interpretation—how many students “passed” the test—while PI scores are somewhat less intuitive. Still, although proficiency rates are important to report, many scholars and educators now believe that PI scores represent a superior approach to accountability, as they encourage schools to serve students along all parts of the achievement spectrum.¹²

Conclusion

State lawmakers should eliminate the Indicators Met subcomponent based on proficiency rates, a move that would also eliminate the need for the points and weighting system used to generate an Achievement component grade. Instead, a single Achievement component rating based on Performance Index scores would best convey how all students in a school perform against state achievement standards. Proficiency rates, however, should continue to be reported as supplemental data.

Progress

This component gauges a school's impact on—or its “value added” to—student growth over time. It is based on individual student data and a statistical method known as value added modelling, which is commonly viewed as a “student growth measure.” Value added sets an expectation that all students—no matter their starting point—can at least maintain their position in the statewide achievement distribution from one year to the next. In a simplified example, a student at the 25th percentile at the end of one year would be expected to maintain that rank at the end of the next year—neither a loss nor a gain.¹³

The sophisticated modelling behind value added scores requires significant IT infrastructure and

statistical expertise, and Ohio has contracted with the SAS Institute, Inc., an analytics firm based in North Carolina, to generate results. SAS also works with Tennessee, Pennsylvania, and North Carolina to provide value added analysis.¹⁴ While the exact specifications of its statistical model are not public, detailed technical documentation and lay-friendly explanations of the modelling are available.¹⁵ In addition to Ohio-specific resources, there is a vast amount of literature on value added methods and other academic growth measures.¹⁶

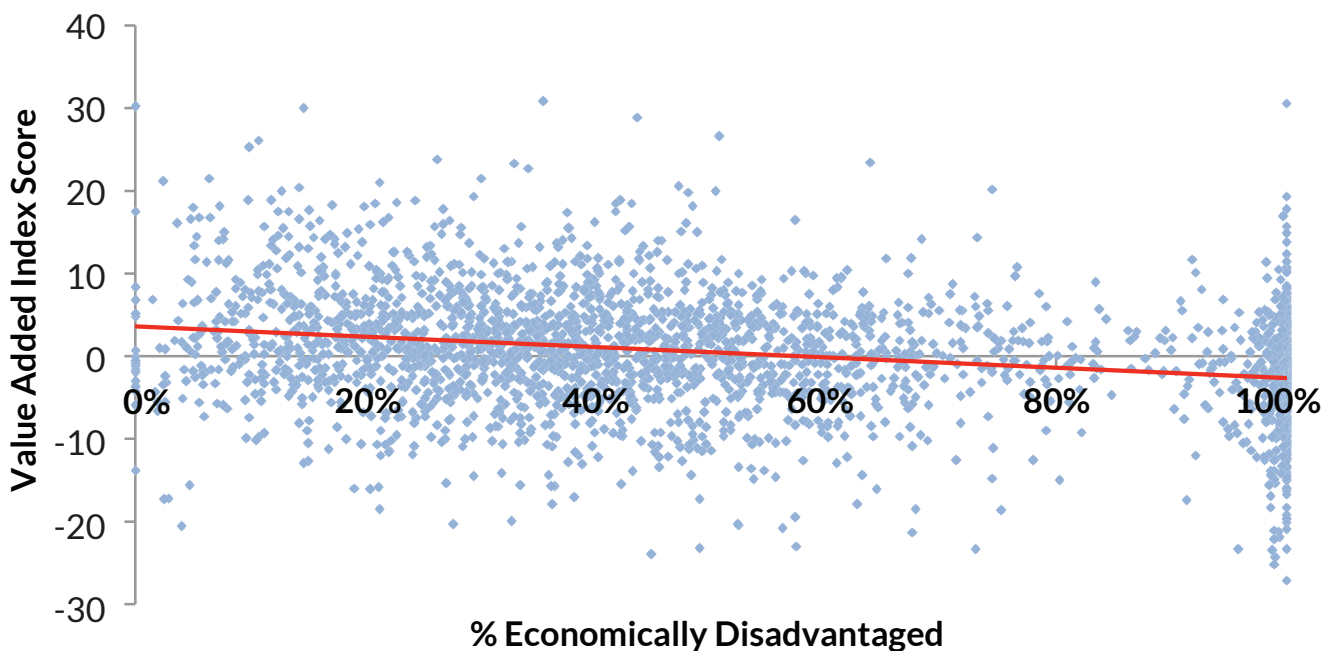
Value added results are reported in a way that incorporates both the magnitude of the estimated gains or losses and the precision of those estimates. The measure of the magnitude is known as the “value added gain,” while the measure of precision is the “standard error.” Based on the estimated gain or loss and standard error, value added “index scores” are computed; those scores form the direct basis of A-F grades. The index score should be seen as a measure of certainty that the average amount of

growth in a school or district is statistically different from zero.¹⁷ Beginning with the 2016–17 school year, Ohio has moved to multi-year averages of schools’ index scores to generate ratings.¹⁸ Value added index scores are translated into A-F grades on the following scale:¹⁹

Index Score	-2.0 or below	-1.99 to -1.0	-0.99 to +0.99	+1.0 to +1.99	+2.0 or above
Grade	F	D	C	B	A

As Figure 6 indicates, value added index scores are only weakly associated with schools’ poverty levels. The chart displays each school’s index score versus its percentage of economically disadvantaged students. The horizontal trend line indicates a loose correlation between the two variables, suggesting that the data are “telling us something about what is going on in schools, not the home.”²⁰ Previous years’ data also show a lack of correlation between value added scores and poverty levels.²¹

Figure 6: Relationship between percent economically disadvantaged (horizontal axis) and value added index scores (vertical axis), Ohio schools, 2016-17



Ohio assigns an overall (or schoolwide) value added grade and three subgroup value added grades when schools have sufficient numbers of students in each subgroup. Ohio calculates subgroup value added scores for gifted students, students whose achievement is in the lowest 20 percent in the state, and students with disabilities. A composite Progress component grade is then determined based on the three subgroup value added grades (assigned a combined 45 percent weight), while the overall value added grade is given 55 percent weight. These weights are used in a points system to calculate the school's Progress grade in a manner similar to the Achievement component (see Table 4 above). State law includes an additional rule that prohibits a school from earning an A on Progress unless all three subgroup grades for value added are at least B's.²²

Comments

Student growth measures are a critical dimension of school report cards, as they provide a picture of school quality apart from students' prior achievement levels or background characteristics. This information can help Ohio families, citizens, and policymakers better understand how schools perform—i.e., their educational “effectiveness.”²³ Because growth measures are less correlated with poverty measures, they can help lower-income communities identify quality schools—those in which students make gains even when they haven't yet reached rigorous proficiency standards. As is clear from Figure 6, high-poverty schools can and do receive both high value added scores (and low ones as well); thus, the measure can be said to effectively differentiate among schools. In contrast, virtually all high-poverty schools receive low grades for Achievement.

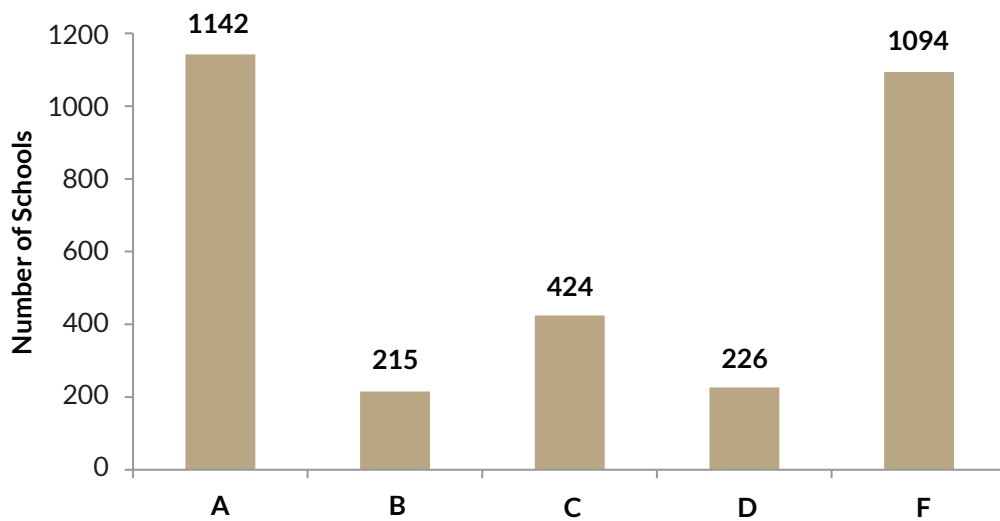
Growth measures also better align incentives with productive behaviors on the part of teachers and schools. Because schoolwide growth measures gauge

the progress of all pupils, they send the message that all students matter—not just those achieving near the proficiency cut-score. In addition, growth measures should weaken perverse incentives for schools of choice to avoid serving lower-achieving pupils—incentives that may be strengthened by proficiency and graduation metrics. They may also encourage schools to pursue innovative, personalized learning strategies that may be difficult to implement when the emphasis is solely on meeting grade-level standards.

Yet growth measures—value added especially—have challenges, too. The calculations that underlie them are notoriously complex, and the results are not intuitive. In Ohio's case, the statistical algorithms used to calculate them are not accessible for independent inspection. This may lead to uncertainty on how to interpret or apply the ratings and data. Recent fluctuations in Ohio's value added grades—e.g., a school or district going from an A in one year to an F in the next—led to questions about what value added measures actually convey about school quality.²⁴ (Ohio's recent shift from ratings based on single-year scores to multi-year averages may help to stabilize the ratings.)²⁵ Some also worry that the value added measure does not expect students and schools to make enough growth over time. If a student is performing at the 10th percentile, is it enough for her school to keep her at the 10th percentile a year later?

Meanwhile, the grading scale that translates value added index scores into A-F grades yields an odd distribution of grades. As Figure 7 indicates, 72 percent of Ohio schools received either an A or an F rating on overall value added with relatively few schools in the B to D range in 2016-17—a pattern that is also noticeable in prior years' data. When almost 40 percent of Ohio schools receive top marks on this measure, it becomes more difficult to pinpoint the very highest-performing schools; the same could be said for very low performers.

Figure 7: Distribution of A-F grades on overall value added, Ohio schools, 2016-17



Conclusion

Student growth measures should play a key role in Ohio’s report card system, including carrying substantial weight in the overall grading formula. They offer the clearest indication of schools’ effectiveness, as they are less influenced by students’ prior achievement or their background characteristics. However, they can still be improved.

First, the Progress component should be streamlined to focus on the overall—or schoolwide—value added measure. This in turn would eliminate the three subgroup value added subcomponents from the Progress component, along with the points and weighting system used to generate the component rating. We propose that two of these subgroup value added ratings be incorporated into a restructured Gap Closing component (more on that beginning at page 22).

Second, given the distribution of A-F ratings, state policymakers should consider ways to alter the

method that translates value added data into A-F grades. Possible alternatives include moving away from using the value added index scores as the basis for grades or adjusting the benchmarks (or “grading scales”) that translate the index scores into ratings.

Third, over the longer term, given the complexity of value added, policymakers might consider alternative approaches to gauging student growth. While value added is regarded by many researchers as a rigorous method, less complex models might strengthen public understanding. Nevertheless, the tradeoffs among various growth methods should be carefully weighed (the Council of Chief State School Officers’ *A Practitioner’s Guide to Growth Models* report is a good place to start).

Graduation

Ohio calculates two high school graduation rates—both a four- and five-year graduation rate—which feed into a larger graduation component. The four-

year rate is based on the percentage of students who earn a high school diploma within four years of entering ninth grade; similarly, the five-year rate is based on the percentage of students earning a diploma within five years. On a school's report card, the four- and five-year rates do not represent the same cohort—the five-year rate is one cohort behind to allow the fifth year to pass. For instance, on its 2016–17 report card, a school's four-year graduation rate is based data from the class of 2016, while its five-year rate is based on the class of 2015.²⁶

Ohio uses an “adjusted cohort graduation rate” calculation that adds to a school's cohort those students who transfer in, while subtracting those who transfer out. Students in the adjusted cohort who do not receive a diploma within four or five years are, in effect, held against the school.²⁷

Each of the two graduation rates receives a letter grade based on grading scales that are partially defined in state law.²⁸ The targets for the five-year rates are slightly higher, as more students are expected to graduate within the longer timeframe. For example, to earn an A on the four-year component, a school must post at least a 93 percent graduation rate; meanwhile, to earn an A on the five-year component, it must post a 95 percent graduation rate. Like the Achievement and Progress components, the state uses a weighting and points system to generate a composite rating (see Table 4 above) with the four- and five-year rates weighted at 60 and 40 percent, respectively.

Comments

Graduation rates are one of the most familiar statistics in education and offer the public a broad sense of how many young people have met basic educational requirements at the end of their K–12 experience. This is information that should be widely accessible and deserves inclusion on report cards. At

Sidebar: Ohio's graduation requirements

Starting with the class of 2018, Ohio policymakers instituted new graduation standards that require students to either: 1) earn passing scores on the state's high school end-of-course exams, 2) achieve a college-ready score on the SAT or ACT, or 3) meet career and technical education requirements.²⁹ However, after much debate, in June 2017 state lawmakers created alternative graduation pathways that apply to the class of 2018—the first class that would have been subject to these updated requirements. These alternatives include meeting senior-year attendance goals, completing a capstone project, receiving a satisfactory senior-year grade point average, or accruing a certain number of volunteer or work hours. Legislators added these alternatives based on concerns that too many students would not meet Ohio's new graduation standards. While it's uncertain whether Ohio lawmakers will extend the alternative pathways to future classes, these debates illustrate how policy decisions on standards impact key statistics on student outcomes.³⁰

the same time, graduation rates can be overemphasized in the school report card system.

First, graduation rates alone send weak signals about whether students are well-prepared for the rigors of college or the workforce. In recent years, the percentage of students earning diplomas has increased nationally, and in many parts of Ohio, percentages exceed 90 percent. Yet college remediation rates and concerns shared by employers suggest that far fewer students leave secondary education well prepared for life after high school. The disconnect can

be traced to modest graduation standards, as well as indications that some schools rely on questionable “credit recovery” programs that push students toward graduation—as is their incentive when held accountable for graduation rates—but may not impart the knowledge and skills needed to succeed after high school.³¹

Second, the adjusted cohort graduation rate calculation does not properly assign responsibility for graduation or non-graduation when students transfer high schools. Receiving schools become fully accountable for transfer students’ on-time graduation, though they may have fallen significantly behind while in their former school.³² High schools that enroll large numbers of credit-deficient students are especially at risk of low ratings when this calculation is utilized.

Third, graduation rates tend to correlate with schools’ demographics. National and state data show disparities between more advantaged students and their peers,³³ as pupils from high-poverty areas are more likely to struggle to meet state requirements—even relatively soft ones—compared to their peers. Nor is it surprising to see lower graduation rates for higher-poverty schools (see Figure 1 above).

Fourth, the combined use of both four- and five-year graduation rates can be confusing, as they supply information on two different cohorts of students side by side. On 2016-17 report cards, for example, almost 30 percent of high schools reported *higher* four- than five-year rates. At first blush, this is counterintuitive—one expects higher five-year rates—but this is of course explained by the use of data from different cohorts.

Conclusion

State lawmakers should assign a single A-F letter grade for the graduation component based on

the four-year graduation rate, the most widely understood measure of attainment in elementary and secondary education. Moving to just a four-year rate also eliminates the points and weighting system needed to generate a component grade; it would also remove the confusion associated with reporting graduation rates from different classes on the same report card. State lawmakers should also explore an alternative graduation rate calculation that holds schools accountable in proportion to the amount of time students are enrolled. While Ohio may not be able to replace the adjusted cohort method due to federal regulations, it could be included as supplemental data on the report card.

Prepared for Success

In 2015-16, Ohio fully implemented Prepared for Success as a graded report card component that applies to districts and high schools. It contains various indicators of readiness for college or career, including remediation-free scores on college entrance exams (ACT or SAT), honors diplomas, industry-recognized credentials, scores on Advanced Placement (AP) or International Baccalaureate (IB) tests, and credits earned through the state’s dual high school/college enrollment program, known as College Credit Plus (CCP). Though not used in the grading system, the college enrollment and completion data of a school’s graduates are also reported.

To calculate component grades, ODE employs a points system. Students generate one point by earning a remediation-free score on ACT or SAT,³⁴ an honors diploma,³⁵ or industry credentials.³⁶ Furthermore, students who earn one point on the primary measures can generate an additional 0.3 “bonus” points by scoring a 3 or above on at least one AP test, scoring a 4 or above on at least one IB test,

or earning three CCP credits.³⁷ Each student generates zero, one, or 1.3 points.

For each school or district, its total amount of points is divided by the number of students in its graduating classes (data from the two most recent cohorts are used).³⁸ Those percentages are then converted to A-F ratings based on a scale is set by the State Board of Education; that scale was recently adjusted to reflect higher benchmarks starting in 2016-17.

Comments

The Prepared for Success component contains valuable information about the readiness of students to enter college or the workforce after graduating high school. The ACT/SAT remediation-free scores help to communicate how many students meet college-ready benchmarks. The inclusion of an industry credential is also critical, as many young people join the workforce directly after high school. The AP, IB, and CCP data also deserve inclusion, as they encourage schools to serve pupils likely to benefit from advanced coursework and exams.

However, incorporating the advanced coursework (and honors diploma) data into this component obscures the view of students exiting high school as college or career ready. The bonus structure, for example, based on AP, IB, and CCP outcomes requires a points system; hence, users no longer receive a straightforward rating based on the percentage of *students* who are college or career ready, but a percentage of points earned by a school. Meanwhile, AP and IB exams are not universally administered, nor do all students participate in CCP. In turn, there is wide variation in student participation across Ohio schools and districts. This could lead to questions about whether schools in which few students earn passing AP or IB scores simply reflect low participation rates, or indicate poor student performance or ineffectual instruction.

Conclusion

State lawmakers should retain the current structure of the Prepared for Success component, including the bonuses that provide additional points when students earn passing scores on advanced exams or receive dual enrollment credit. Policymakers should consider one improvement: Report cards should clearly and prominently display the percentage of students who meet college- or career-ready standards before they graduate (i.e., meet either remediation-free thresholds on college exams or earn an industry credential). This data point would offer the public a clean view of how many students exit high school ready for success, apart from the bonuses based on advanced coursework.

Gap Closing

The Gap Closing component looks at the achievement, graduation rates, and test participation rates of certain student subgroups as defined in federal law: all students; American Indian/Alaskan Native; Asian/Pacific Islander; Black; Hispanic; multiracial; White; economically disadvantaged; students with disabilities; and English language learners. The rationale for reporting subgroup performance separately is that schoolwide averages could mask the performance of particular subgroups. Under current state policy, a district's or school's subgroup must have at least thirty students to be included in this rating.

ODE uses a series of computations to produce this component's letter grade based on subgroup proficiency and graduation rates. The key steps are: 1) Compare each subgroup's proficiency rate in math and reading to statewide proficiency goals to gauge whether an achievement gap exists.³⁹ If the subgroup meets that goal, then the school receives "full credit." 2) If the subgroup

misses the goal, then a school may receive “partial credit” if there are sufficient year-to-year increases in subgroup proficiency rates.⁴⁰ The points across all subgroups are then totaled and divided by the total possible points, yielding a preliminary A-F component grade. Ohio then applies certain demotions based on low proficiency, graduation, or test participation rates, resulting in the final letter grade that is displayed on report cards.

Ohio’s ESSA plan notes that minor changes will be made to the structure of the Gap Closing component starting in 2017–18. These include a shift from subgroup proficiency rates to Performance Index scores—a step in the right direction (see page 16 for discussion on the use of PI versus proficiency rates). It also incorporates the ESSA-required indicator of English language learner progress into this component.

Comments

While reporting data by subgroup remains an important element of report cards, there are several problems with the Gap Closing component.

- The computations are difficult to follow. Rather than enhancing transparency about subgroup performance, this can create a murkier view of subgroup outcomes.
- The use of the year-to-year changes in subgroup proficiency rates (or Performance Index scores) could reflect changes in student composition rather than actual improvements.⁴¹
- The year-to-year calculation sends an unrealistic signal that a school can meaningfully close an achievement gap in a single year. Research has shown that it takes years of sustained work from high-growth schools to help disadvantaged students make up substantial academic ground.⁴²

- Survey data suggest that Gap Closing is viewed by Ohio parents as the “least useful” report card component. In focus groups, some parents express concerns that the component might reinforce stereotypes about students from certain backgrounds.⁴³
- The interpretation of the letter grades for the Gap Closing component is not intuitive. Do D’s and F’s suggest that schools are causing achievement gaps to increase?
- Assigning F’s on Gap Closing to high-growth, high-poverty schools can lead to puzzling interpretations of their performance. Last year, 161 high-poverty schools received Fs on Gap Closing and A’s on value added. Taking these ratings at face value, it’s unclear whether these schools are helping low-income children make gains (A’s on value added) or lose ground (F’s on Gap Closing).

Conclusion

The Gap Closing component is unnecessarily complex and yields perplexing results. We propose a more straightforward way of gauging subgroup performance that gives credit to schools when subgroups meet Performance Index or student growth (or “value added”) targets. The component could be designed in the way displayed in the table below; this hypothetical school would have met ten out of seventeen indicators (59 percent), which could then be converted into a letter grade. State legislators should rename the component as well, calling it “Equity,” since the component takes into account students who are not typically low achieving (e.g., Asian or gifted).

Table 5: Recommended structure of a revised Gap Closing component

	Performance Index Goal	Student Growth Goal
Race/Ethnicity		
African American	Met	Met
Asian	Met	Not Met
Hispanic	Not Met	Met
Multiracial	Met	Not Met
Native American	*	*
White	Met	Met
Economically disadvantaged	Not Met	Met
English language learner	Not Met	Met
Gifted students	Met	Not Met
Students with disabilities	*	*
English language learner (progress) ⁴⁴	Met	

Notes: Using the current value added system, Ohio could give schools credit for meeting the goal if a subgroup’s value added index score is greater than -1.0 (equivalent to a C or above). Annual performance index score goals in math and ELA by subgroup are documented in Ohio’s ESSA plan, as are the progress goals for ELL students.⁴⁵ Under Ohio’s ESSA plan, Performance Index goals vary by subgroup, depending on each subgroup’s baseline scores in 2015-16.

K-3 Literacy

Research has shown that children who struggle to read early in life have difficulty catching up in middle and high school and face a greater likelihood of not earning a high school diploma.⁴⁶ With the adoption of the third grade reading guarantee, Ohio is placing an increasing emphasis on early literacy. While the centerpiece of the guarantee is its grade retention provision, Ohio has also implemented other policies that aim to increase attention onto early reading, one of which is the addition of the K-3 literacy component of school report cards.

First used as a graded component in 2014-15, K-3 Literacy considers whether “off-track” (for reading proficiency in third grade) students improve to “on track” from one year to the next, as indicated

by state-required reading diagnostic tests administered in the fall. Schools may select these assessments from an approved vendor list published by ODE or they may use a state-developed diagnostic test. ODE has approved thirteen vendor exams, with each vendor establishing a cut score that determines on- or off-track status.⁴⁷ For Kindergarten, schools may use the statewide Kindergarten Readiness Assessment (KRA)—the only compulsory statewide assessment given to students prior to grade three—or another diagnostic test.

The K-3 Literacy calculations focus principally on off-track students. For example, schools are penalized if a student is off track in the fall of first grade and off track again in the fall of second grade. However, schools receive credit when students move

from off- to on-track status from one autumn to the next. The only case in which on-track students play a role in the calculations is when they fall short of proficiency on their third-grade ELA state assessment; in that case, points are deducted from the school. These calculations produce a K-3 Literacy improvement percentage that is then converted into an A-F component rating.

Comments

Early literacy is critical to long-term academic success, and done well, holding schools accountable for student learning in these pivotal grades makes sense. Yet the state’s present K-3 Literacy component is problematic in several respects.

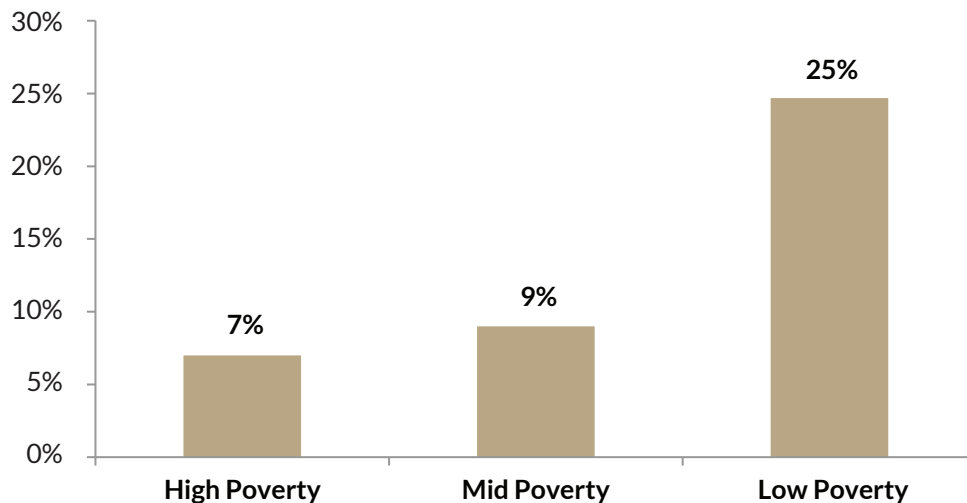
Most concerning is the use of an array of vendor diagnostic exams in an accountability setting. Schools may be tempted to search for the least challenging assessment or the lowest cut scores to gain an advantage on this high-stakes measure. This calls into

question whether schools’ ratings are being artificially influenced by their choice of tests. Moreover, schools self-report on- and off-track data to ODE—an “honor system” that could be subject to gaming unless a robust data verification process is in place.

Second, the use of fall testing data in school accountability settings is unusual. K-3 Literacy evaluates schools based on whether students meet the previous year’s literacy standards in the fall,⁴⁸ in effect gauging reading achievement on a two- to three-month lag. This is cause for concern, given the well-known problem of summer learning loss that could distort results, especially for high-poverty schools.⁴⁹ This also raises the question of whether fall diagnostic assessments—used to inform instruction for the year—should be used at all for accountability purposes.

Third, K-3 Literacy calculations represent a coarse approach to calculating growth. Since only two performance categories are considered, schools receive

Figure 8: Percentage of elementary schools receiving no K-3 Literacy rating by poverty level, 2016-17



Note: Elementary schools are identified as any school reporting second grade enrollment; high-poverty schools have 66 to 100 percent economically disadvantaged (ED) students; mid-poverty = 33 to 66 percent ED; low-poverty = 0 to 33 percent ED. The number of exempt schools was 104 low-poverty; 53 mid-poverty; and 48 high-poverty schools.

no credit when off-track students make improvements but still fall short of the on-track bar. Nor does it give schools credit when on-track students make additional progress. The current calculations may be the best that can be done with extant data—schools report only on or off track status—but they do not yield a clear picture of whether schools are helping *all* youngsters improve.

Fourth, though K-3 Literacy results from 2016-17 were not closely linked with schools' poverty rates, a fair number of low-poverty schools—those with insufficient numbers of “off track” learners—are excluded.⁵⁰ As Figure 8 on page 25 indicates, a larger proportion of low-poverty elementary schools received no rating on this component in 2016-17. In effect, this measure tends to hold higher poverty schools more accountable for their early readers—putting them at risk of receiving a low grade—than low-poverty schools that enjoy an exemption.

Conclusion

The K-3 Literacy component offers an uncertain picture on whether schools are truly boosting early learning. The ability of schools to select among various diagnostic assessment options creates comparability problems, and the performance of all students—or even all schools—is not taken into account. Ohio lawmakers should eliminate this component, for now, until better approaches to gauging early learning outcomes are developed.

This does not suggest that policymakers neglect early education. However, they would better focus their efforts on the proper implementation of the third grade reading guarantee, a policy that already provides a strong incentive for schools to pay close attention to early learners. Most notably, policymakers

should maintain a high bar that students must reach on their third grade reading exams in order to advance to fourth grade. They should also make sure (per state law) that families receive timely information from their district or school when their children are deemed not on track for reading proficiency according to diagnostic exams.

In Closing: A Better Balance

Creating well-functioning school report cards entails careful balancing of multiple tradeoffs. Among the most important are simplicity versus comprehensiveness; growth versus status measures; and schoolwide versus subgroup performance. As Ohio's report cards have evolved, policymakers have not always struck the right balance when reconciling these tensions. For example, Ohio has erred too much on the side of a more inclusive report card and sacrificed elegance and simplicity. Policymakers have also placed too heavy an emphasis on status measures and not enough on indicators of student growth. Ohio policymakers should make course corrections to the present report cards, focusing on the following actions:

- Reduce the number of A–F grades by eliminating the K–3 Literacy component and the graded subcomponents.
- Restructure the Gap Closing component as a set of indicators that student subgroups are expected to meet, based on their Performance Index and value added scores; legislators should also rename this component “Equity.”
- Devise a summative rating formula that balances status and growth measures to create a more even playing field for schools enrolling students of varying backgrounds.
- Reduce the amount of supplemental data found on formal report cards to help the public focus on the most critical gauges of school quality.

With these changes, state leaders will devise a simpler, clearer, and more evenhanded report card that works better for the people of Ohio.

Appendices

Appendix A: Federal report card requirements

Federal education law—the Every Student Succeeds Act (ESSA)—requires states to evaluate schools along at least five indicators. The table below shows how the report card that we recommend in this report would comply with ESSA, along with examples of comparable report card components from other states that have already been approved by the U.S. Department of Education (USDOE) under their ESSA plans.

Federal Indicator	Proposed Implementation	Notes on Compliance
Achievement	Achievement component (i.e., Performance Index score)	USDOE has approved the use of performance indexes to fulfill ESSA's achievement indicator. ⁵¹
Other academic indicator	Progress component (i.e., overall value added)	USDOE has approved the use of value added scores to fulfill ESSA's other academic indicator. ⁵²
Graduation	Graduation component (i.e., four-year graduation rate)	The use of the four-year graduation rate is required under ESSA.
English language learner (ELL) progress	Indicator of ELL progress within the Gap Closing component	The ELL progress indicator is required under ESSA.
Student success or school quality	Gap Closing component (renamed "Equity")	USDOE has approved the use of an indicator based in part on subgroup performance to fulfill ESSA's Student Success or School Quality indicator; ⁵³ other states that recently submitted their plans for approval have proposed using subgroup data to meet this ESSA indicator. ⁵⁴

Appendix B: Distribution of school ratings by schools' poverty levels, 2016-17

These tables show a breakdown of A-F ratings on each of Ohio's grade components or subcomponents by schools' poverty status. "High-poverty" schools are those reporting 66 to 100 percent economically disadvantaged (ED) students; mid-poverty schools as 33 to 66 percent ED; and "low-poverty" schools as those reporting 0 to 33 percent ED. Several schools may be misclassified due to participation in the federal Community Eligibility Program, which allows eligible high-poverty districts to offer meals to all students at no cost regardless of household income. In turn, participating districts report 100 percent ED pupils.⁵⁵

	Indicators Met			Performance Index			Achievement		
	High-Poverty	Mid-Poverty	Low-Poverty	High-Poverty	Mid-Poverty	Low-Poverty	High-Poverty	Mid-Poverty	Low-Poverty
A	<1%	1%	7%	<1%	<1%	7%	<1%	1%	7%
B	<1%	2%	16%	3%	19%	51%	1%	7%	31%
C	<1%	2%	9%	15%	50%	34%	7%	33%	40%
D	2%	10%	23%	51%	29%	7%	47%	57%	21%
F	97%	84%	45%	31%	2%	0%	45%	3%	<1%
N Schools	1,076	1,190	935	1,076	1,187	938	1,080	1,190	939

	Gap Closing			K-3 Literacy			Prepared for Success		
	High-Poverty	Mid-Poverty	Low-Poverty	High-Poverty	Mid-Poverty	Low-Poverty	High-Poverty	Mid-Poverty	Low-Poverty
A	2%	6%	20%	2%	6%	5%	0%	0%	5%
B	5%	14%	22%	12%	25%	32%	1%	<1%	11%
C	2%	11%	14%	45%	54%	53%	1%	9%	44%
D	4%	13%	16%	38%	13%	9%	17%	68%	37%
F	86%	57%	27%	4%	2%	2%	80%	22%	3%
N Schools	1,062	1,183	933	715	588	331	205	279	313

	Four-Year Graduation Rates			Five-Year Graduation Rates			Graduation		
	High-Poverty	Mid-Poverty	Low-Poverty	High-Poverty	Mid-Poverty	Low-Poverty	High-Poverty	Mid-Poverty	Low-Poverty
A	15%	41%	79%	15%	37%	77%	13%	35%	81%
B	8%	28%	17%	12%	39%	19%	12%	39%	16%
C	14%	17%	3%	17%	15%	3%	13%	15%	2%
D	8%	7%	0%	16%	4%	0%	19%	5%	0%
F	55%	7%	1%	40%	5%	1%	42%	6%	2%
N Schools	204	278	312	199	275	311	206	278	314

	Overall Value Added			Value Added - Gifted			Value Added - Lowest 20%		
	High-Poverty	Mid-Poverty	Low-Poverty	High-Poverty	Mid-Poverty	Low-Poverty	High-Poverty	Mid-Poverty	Low-Poverty
A	21%	41%	49%	12%	23%	37%	17%	21%	24%
B	7%	6%	8%	10%	15%	13%	11%	15%	18%
C	15%	14%	13%	38%	38%	27%	26%	34%	38%
D	10%	6%	6%	18%	13%	9%	12%	12%	11%
F	48%	33%	24%	22%	12%	13%	35%	18%	9%
N Schools	1,040	1,150	911	348	865	795	1,006	1,061	709

	Value Added - Students with Disabilities			Progress		
	High-Poverty	Mid-Poverty	Low-Poverty	High-Poverty	Mid-Poverty	Low-Poverty
A	13%	20%	24%	11%	16%	21%
B	9%	11%	12%	21%	36%	40%
C	28%	27%	31%	18%	15%	14%
D	16%	13%	14%	31%	26%	21%
F	34%	30%	19%	20%	7%	4%
N Schools	887	1,050	799	1,040	1,150	911

Endnotes

- ¹ Editorial board, “Ohio Report Cards Paint Incomplete Picture,” *Record-Courier* (September 24, 2017).
- ² Saperstein Associates, “State Report Cards: Focus Group Findings” (2016) and Edge Research and HCM Strategists, “Parent Perspectives on Ohio School Report Cards: Findings from an Online Survey” (2017).
- ³ Ohio Department of Education, “2016-17 AMO Gap Closing Measure”: <http://education.ohio.gov/getattachment/Topics/Data/Report-Card-Resources/Gap-Closing-Measure/Technical-Documentation-AMO-Calculation.pdf.aspx>
- ⁴ Brandon L. Wright and Michael J. Petrilli, *Rating the Ratings: Analyzing the First 17 ESSA Accountability Plans*, Fordham Institute (2017).
- ⁵ Ohio’s older report cards can be accessed via archives: <http://reportcard.education.ohio.gov/Pages/default.aspx>.
- ⁶ “Advanced plus” is also included with a weight of 1.3 for students scoring proficient or above on an above grade-level exam; typically, less than 1 percent of students reach that status. In addition, zero weight is assigned when students do not take an exam.
- ⁷ For the range of scaled scores that equate to each achievement level, see Ohio Department of Education, “Understanding Ohio’s State Test Score Report, 2016-17”: http://oh.portal.airast.org/ocba/wp-content/uploads/Understanding_State_Tests_Reports_2016-2017.pdf.
- ⁸ The proficiency rate necessary to meet an indicator is 80 percent across all grade-subject testing combinations; see Ohio Department of Education, “2016-17 Indicator Targets”: <https://education.ohio.gov/getattachment/Topics/Data/Report-Card-Resources/Achievement-Measure/Indicators-Table.pdf.aspx>.
- ⁹ The tables that translate raw data into subcomponent points, as well as weighted average points into component grades, are available at Ohio Department of Education, “ESSA Plan: Appendix B” (pp. 56-68): <http://education.ohio.gov/getattachment/Topics/Every-Student-Succeeds-Act-ESSA/ESSA-Template-Appendix-B.pdf.aspx>
- ¹⁰ The results of this calculation are not intuitive. One would expect the Achievement component grade to be a B, given the greater weight on PI. The explanation is that the PI score is in the lowest quartile of the B range, while the Indicators Met score is toward the middle of the C range; together, this leads to a score (3.06) that does not quite meet the Achievement component’s minimum threshold for a B rating (3.125 points).
- ¹¹ Derek Neal and Diane Whitmore Schanzenbach, *Left Behind by Design: Proficiency Counts and Test-Based Accountability*, National Bureau of Economic Research (2007): <http://www.nber.org/papers/w13293>.
- ¹² Morgan Polikoff, “A Letter to the U.S. Department of Education: Final Signatory List,” (2016): <https://morganpolikoff.com/2016/07/12/a-letter-to-the-u-s-department-of-education/>
- ¹³ For a more detailed explanation of Ohio’s value added model, see SAS, *Technical Documentation of EVAAS Analysis* (2016): <http://education.ohio.gov/getattachment/Topics/Data/Report-Card-Resources/Ohio-Report-Cards/Value-Added-Technical-Reports-1/Technical-Documentation-of-EVAAS-Analysis.pdf.aspx>.

- ¹⁴ States use various approaches to calculating individual student growth of which value added is one of them. Several states—e.g., Colorado, New Jersey, and Washington—use a complex statistical approach known as Student Growth Percentiles (SGP), while other states use simpler methods such as tracking individual students’ transitions from one achievement level to another (Florida) or changes in scaled scores (Connecticut). For more on the approaches to growth, see Kathleen E. Castellano and Andrew D. Ho, *A Practitioner’s Guide to Growth Models*, Council of Chief State School Officers (2013): http://www.ccsso.org/Resources/Publications/A_Practitioners_Guide_to_Growth_Models.html
- ¹⁵ SAS, *Technical Documentation of EVAAS Analyses*, Ohio Department of Education (2016): <http://education.ohio.gov/getattachment/Topics/Data/Report-Card-Resources/Ohio-Report-Cards/Value-Added-Technical-Reports-1/Technical-Documentation-of-EVAAS-Analysis.pdf.aspx> and SAS, *Misconceptions about Value-Added Reporting in Ohio* (2012): http://static.battelleforkids.org/documents/ohio/Misconceptions_About_OH_EVAAS_v1%200.pdf.
- ¹⁶ See for example, Douglas N. Harris, *Value-Added Measures in Education* (Harvard University Press: 2011); Mark Ehlert, et al., “Choosing the Right Growth Measure,” *Education Next* (2014): <http://educationnext.org/choosing-the-right-growth-measure/>; Carnegie Knowledge Network, “Knowledge Briefs”: <http://www.carnegieknowledgenetwork.org/briefs/a-focus-on-value-added/>; and RAND Corporation: “Value-Added Modeling 101”: <https://www.rand.org/education/projects/measuring-teacher-effectiveness/value-added-modeling.html>.
- ¹⁷ Note the evidentiary language in SAS’s interpretation of the value-added ratings: e.g., <https://ohiova.sas.com/valueAdded.html?as=a&aj=a&w4=9&w4=9&x9=-8&yb=-8&x7=-8&ww=98847>.
- ¹⁸ During the recent testing transitions, Ohio used single-year value added scores, though prior to the transition it used multi-year ratings. State law requires the use of three years of value added data as available; ORC 3302.03(C)(1)(e).
- ¹⁹ State statute prescribes this scale for 2012-13 [ORC 3302.03(A)(1)(e)(i-v)] but does not appear to require its use in any year thereafter.
- ²⁰ SAS, “Value Added Reporting in Ohio” (2016): http://search-prod.lis.state.oh.us/cm_pub_api/api/unwrap/chamber/131st_ga_ready_for_publication/committee_docs/cmte_h_education_1/testimony/cmte_h_education_1_2016-05-17-0930_1443/sashb524.pdf.
- ²¹ Aaron Churchill, “House Bill 524: An Opportunity for Ohio to Strengthen its Value Added System,” *Ohio Gadfly Daily* (2016).
- ²² ORC 3302.03(C)(1)(e).
- ²³ Tom Kane, a Harvard University researcher, writes: “There is now substantial evidence that value-added estimates capture important information about the causal effects of teachers and schools.” Tom Kane, “Do Value-Added Estimates Identify Causal Effects of Teachers and Schools?” Brookings Institution (2014).
- ²⁴ Bill Bush and Shannon Gilchrist, “Districts question poor marks on state report cards” *Columbus Dispatch* (February 26, 2016): <http://www.dispatch.com/article/20160225/news/302259712>.
- ²⁵ The tradeoff to using a multi-year average is that, by including past performance, the results may not be indicative of the current performance of a school. Experts, however, recommend using multi-year averages; e.g., Douglas N. Harris, *Value Added Measures in Education* (p. 208).

- ²⁶ The four-year rate is already lagged by one year to account for summer graduates.
- ²⁷ In 2008, the U.S. Department of Education issued regulations that required states to report graduation rates using the adjusted cohort method. See Sterling C. Lloyd, “44 States Now Using the Same Grad.-Rate Formula,” *Education Week* (June 1, 2012).
- ²⁸ State law sets only a benchmark for an A letter grade, which is 93 percent for an A on the four-year rate and 95 percent for an A on the five year rate; see ORC 3302.03(A)(1)(d).
- ²⁹ For more on Ohio’s graduation requirements: <http://education.ohio.gov/Topics/Ohio-Graduation-Requirements>.
- ³⁰ This is also an issue with proficiency rates: Policymakers can set a low bar for “proficiency” and in turn allow the large majority of students to be deemed “proficient.”
- ³¹ Sonali Kohli, “Schools Are Boosting Graduation Rates by Offering ‘Credit Recovery.’ But What Are Students Learning?” *Los Angeles Times* (July 2, 2017); Robert Pondiscio, “The Phoniest Statistic in Education” Fordham Institute (January 13, 2017): <https://edexcellence.net/articles/the-phoniest-statistic-in-education>.
- ³² The reverse is also true: A school could enroll an “on track” student late in her high school career, and receive full credit for that student’s graduation while her former school receives no credit.
- ³³ National Center for Education Statistics, “Public High School Graduation Rates” (2017): https://nces.ed.gov/programs/coe/indicator_coi.asp#.
- ³⁴ The Ohio Department of Higher Education sets the qualifying scores. For the ACT, they are 18 in English, 22 in math, and 21 in reading; for the SAT, they are 430 in writing, 520 in math, and 450 in reading. Provided the student takes the same entrance exam, the highest score on each section is used when students take the ACT or SAT multiple times.
- ³⁵ Honors diploma requirements are posted at: <http://education.ohio.gov/Topics/Ohio-Graduation-Requirements/Graduation-Requirements-2014-2017/Honors-Diplomas>
- ³⁶ The State Board of Education has approved dozens of industry recognized credentials, each with a point value attached to them. Most are worth three to six points, with some credentials generating the maximum twelve points. To meet the Prepared for Success condition, a student must earn twelve points within a single career field. For the list of credentials, see: <http://education.ohio.gov/Topics/Ohio-Graduation-Requirements/Graduation-Requirements/Industry-Recognized-Credentials-and-WorkKeys/Industry-Recognized-Credentials>.
- ³⁷ A score of 3 on AP exams and 4 on IB exams is usually considered by colleges to be minimum score needed to earn credit. The maximum AP and IB scores are 5 and 7 respectively.
- ³⁸ Due to the bonus structure, a school could theoretically receive more points than it has students. The state currently caps the grading scale at 100 percent, though no schools reached the 100 percent level in 2016-17; the highest percentage was 98.
- ³⁹ In 2016-17, Ohio set a single, annual proficiency goal that all subgroups must meet: <http://education.ohio.gov/getattachment/Topics/Data/Report-Card-Resources/Gap-Closing-Measure/Technical-Documentation-AMO-Calculation.pdf.aspx>.

- ⁴⁰ These steps are repeated for each subgroup's graduation rates.
- ⁴¹ For more on this problem, see Matthew DiCarlo, "If Your Evidence Is Changes In Proficiency Rates, You Probably Don't Have Much Evidence," Albert Shanker Institute (2012) and Steven M. Glazerman and Liz Potamites, *False Performance Gains: A Critique of Successive Cohort Indicators*, Mathematica Policy Research (2011).
- ⁴² Research on the exceptionally high-performing Boston charters indicates that they reduce the black-white achievement gap by about one-third to one-fifth in a year. See Tom Kane, "The Cost of the Charter School Cap," *CommonWealth* (October 5, 2016): <https://commonwealthmagazine.org/education/the-cost-of-the-charter-school-cap/>.
- ⁴³ Edge Research and HCM Strategists, "Parent Perspectives on Ohio School Report Cards" (January 2017) and Saperstein Associates, Inc., "State Report Cards: Focus Group Findings" (2016).
- ⁴⁴ The ELL progress indicator is included to meet federal ESSA requirements; the main difference in this indicator is that it is based on Ohio's alternative assessments given to ELLs.
- ⁴⁵ Ohio Department of Education, "Appendix A: Long Term Goals": <http://education.ohio.gov/getattachment/Topics/Every-Student-Succeeds-Act-ESSA/ESSA-Appendix-A.pdf.aspx>.
- ⁴⁶ Annie E. Casey Foundation, "Early Warning! Why Reading by the End of Third Grade Matters" (2010): <http://www.aecf.org/resources/early-warning-why-reading-by-the-end-of-third-grade-matters/> and Ohio Department of Education, "Third Grade Reading and College and Career Readiness": <http://education.ohio.gov/getattachment/Topics/Early-Learning/Third-Grade-Reading-Guarantee/TGRG-CCR.pdf.aspx>.
- ⁴⁷ For more see, Ohio Department of Education, "Diagnostic Tests Guidance Manual": <http://education.ohio.gov/getattachment/Topics/Testing/Diagnostic-Assessments/Diagnostic-Tests-Guidance-Manual/Diagnostic-Manual-2016-2017.pdf.aspx>.
- ⁴⁸ See the note from Ohio Department of Education, "Diagnostic Tests Guidance Manual" (emphasis added): "The results of the Sept. 30 and Nov. 1 reading diagnostic test determines whether a child is considered on track or not on track in reading. A child will be considered on track at the beginning of each grade if the child is reading at the level set by Ohio's Learning Standards for *the end of the previous grade*."
- ⁴⁹ Jennifer Sloan McCombs, et. al., *Making Summer Count: How Summer Programs Can Boost Children's Learning*, Rand Corporation (2011).
- ⁵⁰ State law exempts any school or district that has fewer than 5 percent of their Kindergartners reading below grade level in the current school year from a K-3 literacy rating. At a district level, 63 districts (just over 10 percent of districts statewide) received no K-3 literacy rating in 2016-17. The exempt districts were, on average, 27 percent economically disadvantaged.
- ⁵¹ See Connecticut's approved state plan: <https://www2.ed.gov/admins/lead/account/stateplan17/map/ct.html>.
- ⁵² See Tennessee's approved state plan: <https://www2.ed.gov/admins/lead/account/stateplan17/tnconsolidatestateplan817.pdf>.
- ⁵³ See Arizona's approved state plan: <https://www2.ed.gov/admins/lead/account/stateplan17/azconsolidatedstateplan.pdf>.

⁵⁴ See Mississippi and New Hampshire’s submitted state plans which propose a growth measure based on low-achieving students to fulfill this ESSA indicator: <https://www2.ed.gov/admins/lead/account/stateplan17/msconsolidatedstateplan.pdf> and <https://www2.ed.gov/admins/lead/account/stateplan17/nhconsolidatedstateplan.pdf>

⁵⁵ For more, see Ohio Department of Education, “Community Eligibility Provision”: <http://education.ohio.gov/Topics/Other-Resources/Food-and-Nutrition/Resources-and-Tools-for-Food-and-Nutrition/Community-Eligibility-Option>.



100 E. Broad Street, Suite 2430

Columbus, OH 43215

614-223-1580

www.edexcellence.net/ohio-policy