QUALITY CHECK

THE NEW, BEST WAY TO MEASURE SCHOOL PERFORMANCE

BY LYNN OLSON

APRIL 2025

FutureEd
Independent Analysis, Innovative Ideas

KEYSTONE
POLICY CENTER

QUALITY **CHECK**

THE NEW, BEST WAY TO MEASURE SCHOOL PERFORMANCE

TABLE OF CONTENTS

Foreword

What's the best way to measure a school's quality? The answer depends on one's definition of a good school. And parents, educators, employers, and policymakers hold many different opinions on what schools should be about, from preparing young people to acquire employable skills to developing informed and engaged citizens.

Yet virtually every educational aim rests on the same foundation—giving students a strong academic grounding, helping them develop the knowledge and habits of mind that allow students to think critically, communicate effectively, and acquire new knowledge and skills as society changes.

At this challenging moment in American education, with student achievement in decline, FutureEd and the Keystone Policy Center decided to approach the question of how best to measure schools from scratch. We combed the research on the features of schools that make the greatest contribution to academic achievement in order to create a new measurement model to capture those features.

This report, researched and written by FutureEd Senior Fellow Lynn Olson, presents what we found. Encouragingly, there is a robust body of research spanning decades that points to the key features of successful schools and, importantly, to the metrics for measuring the presence of those features in schools. Yet we also learned that schools today are routinely rated in ways that fail to reflect the research on school quality, under measurement systems that fail to capture essential characteristics of quality schools and key student outcomes.

It is time, the research suggests, to rethink how we measure school performance. To help with that work, this report summarizes the research on school quality; outlines a new measurement model suggested by the research; and explores challenges to implementing the model at scale, as well as ways that policymakers can overcome those challenges.

We're grateful to Elaine Allensworth, director of the UChicago Consortium on School Research, and Chris Hulleman of the Motivate Lab at the University of Virginia for reading a draft of the report.

FutureEd Senior Policy Analyst Bella DiMarco managed the production of the report and Molly Breen and Merry Alderman on our editorial team lent their skills to editing and design.

Crown Family Philanthropies made the project possible.

Thomas Toch
Director, FutureEd

Van Schoales
Senior Policy Director, Keystone Policy Center

Measures of school quality vary widely, from those enshrined in federal law to state report cards, comprehensive school site visits, and commercial rankings by *U.S. News*, the *Wall Street Journal*, and others. There is no consensus on which measures are best.

This is partly because different stakeholders in education—educators, parents, employers, policymakers—hold different views about the purpose of schooling. Some see it chiefly as a means to train young people for employment, others as a means to develop informed and engaged citizens, still others as a way for children to pursue their dreams and lead fulfilling lives. Yet virtually every educational aim rests on the same academic foundation: providing students with the knowledge and habits of mind that allow them to think critically, communicate effectively, and continue to learn and adapt as society changes.

It's worth asking, then, at this challenging moment in American education, when student achievement is in decline: What constellation of school quality measures would emerge if education leaders built a new measurement model from scratch, drawing on research to identify the features of schools that provide students with the strongest possible academic foundation to create a new measurement model that captures those features?

In fact, there's a growing body of research, some of it going back decades, to guide such an exercise—research pointing to the need for a wider range of school performance metrics than are commonly used today. Standardized test scores, the predominant measure of school quality for several decades and required by federal law, are important but insufficient measures of school performance, unable to capture essential characteristics of quality schools and key student outcomes.

Rethinking school measurement is an especially important task in an era of expanding school choice, when families need the best available information on school performance so they can make the best possible choices for their children. This report summarizes the new research on school quality; outlines a new measurement model suggested by the research; and explores challenges to implementing the model at scale, as well as ways that policymakers can overcome those challenges.

The Attributes of Strong Schools

Research spanning nearly five decades reveals which school features provide students with a solid academic foundation and the habits of mind necessary for success.

In 1979, Ronald Edmonds, a high school teacher and education administrator who eventually joined the faculties at Harvard, the University of Michigan, and Michigan State, published a landmark study, "Effective Schools for the Urban Poor." It identified five characteristics of urban elementary schools that were producing strong academic results for students from low-income families: high expectations; a focus on basic skills acquisition and the ability to direct energy and resources toward that purpose and away from other activities; an orderly atmosphere; strong school leadership; and frequent monitoring of pupil progress.[1]

Edmonds undertook his research in response to the controversial 1966 report *Equality of Educational Opportunity* led by University of Chicago sociologist James S. Coleman, which found that family background and socioeconomic status were the major determinants of student achievement.[2] Edmonds set out to prove that schools do affect student learning, for better or for worse, kicking off decades of research into the characteristics of effective schools.

Three decades later, Anthony S. Bryk and colleagues at the UChicago Consortium on School Research published another landmark study on the dimensions of schools that produce strong student outcomes.

In 1988, the Illinois legislature had radically decentralized the Chicago public schools, granting parents and communities significant authority to reform their schools through the creation of Local School Councils that could hire principals and approve a school's discretionary budget and improvement plan. Bryk and his fellow researchers analyzed district wide elementary school data from 1991 to 2005 to identify the differences between schools that substantially improved student achievement and those that did not.

The resulting book, *Organizing Schools for Improvement: Lessons from Chicago*, identified five characteristics of schools that improved achievement: ambitious instruction, characterized by a coherent curriculum and significant academic demands; a school environment where students feel safe, supported, and respected by their teachers; committed teachers working to improve their schools and themselves; effective school leaders; and involved families.[3] The study found that schools strong on at least three of the five essential qualities were 10 times more likely to show gains in student learning than schools weak on three or more of the essentials. A low score in even one of the essentials reduced the likelihood of improvement to less than 10 percent.[4] These dimensions of quality schools became the basis for the 5Essentials surveys of teachers and students that are currently used in Chicago and throughout Illinois to inform school-improvement efforts.

Research conducted by the consortium in high schools in 2020 found that the five essentials also were positively and significantly related to a range of high school outcomes, including test scores, attendance, grade-point-averages, whether freshmen were on track to graduate, and college enrollment.[5]

"Those five essential supports, they're so important," says Elaine Allensworth, executive director of the consortium and one of the lead researchers of the study. "We keep finding over and over again that they tell you how likely a school is to improve that year and in the future."

In 2023, European researchers Spela Javornik and Eva Klemencic Mirazchiyski published a comprehensive literature review of the factors

contributing to school effectiveness. The review similarly identified strong leadership, expert teaching practices, positive school culture, and parental involvement as key.[6] School resources, such as funds for supplies and facilities, were also important, especially in under-resourced communities.

Much of the research on school quality points to positive school culture as a key feature of good schools, one that supports students' academic achievement as well as their social and emotional development. Research has found that the academic, social, and emotional growth of children are deeply connected. If students don't feel safe, supported, and challenged in school, they are not going to learn.

2023's "Investing in Adolescents," also by the University of Chicago consortium, found that high schools that support students' development across multiple dimensions—helping them raise academic achievement but also promoting their social-emotional development and positive behavior — had up to double the positive impact on long-term
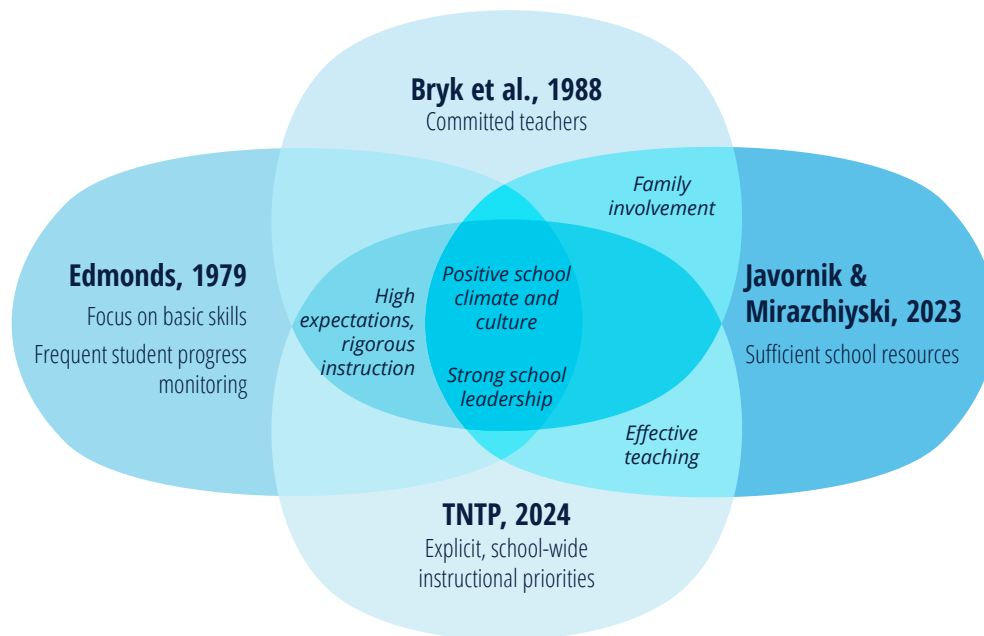
academic outcomes such as high school graduation and postsecondary enrollment and attendance, compared with schools that focused solely on test scores.[7] These "multi-dimensional" schools featured many of the five essentials, including a supportive environment, ambitious instruction, and collaborative teachers.

C. Kirabo Jackson, one of the co-authors of the study and a professor of human development and social policy at Northwestern University, says the findings "strongly suggest the need to take a more holistic view of both adolescents and schools."

A recent study by NewSchools, a nonprofit venture philanthropy that supports high-performing public schools, primarily in the charter sector, reinforced Jackson's point.

The organization analyzed more than 32,000 student surveys and data from 20,000 academic assessments administered during the 2022-23 school year. Its goal was to determine which specific student mindsets, habits, and skills, and which elements of school culture, drive academic results.[8]

Research-Based Ingredients of Successful Schools

Bryk et al., 1988
Committed teachers

Edmonds, 1979
Focus on basic skills
Frequent student progress monitoring

High expectations, rigorous instruction

Positive school climate and culture

Strong school leadership

Family involvement

Javornik & Mirazchiyski, 2023
Sufficient school resources

Effective teaching

TNTP, 2024
Explicit, school-wide instructional priorities

Its finding: students perform the best when they feel valued, secure, and connected to their school.

Elementary schools where students gave high ratings to school culture saw a 4.4-month boost in reading progress compared to schools with low school culture ratings. Middle school students who felt safe and secure at school, free from worries about violence or bullying, achieved 2.5 additional months of learning in math.

Most recently, a September 2024 study released by TNTP, "The Opportunity Makers: How a Diverse Group of Public Schools Help Students Catch Up—and How Far More Can," analyzed data from 28,000 elementary and middle schools where students start out below grade level and identified the 5 percent of those schools that help students gain more than 1.3 years of learning per academic year—a pace that catches the average student up to grade level by the time they leave their school.

After closely examining seven of these schools, TNTP, a non-profit that works with school districts to increase teacher supply, found that while they were all very different, they shared three common approaches: belonging (creating an emotional climate for learning that activates the ability of students to excel); consistency (delivering consistently good teaching and grade-level content for all students); and coherence (building a unified instructional program focused on individual students and setting priorities that are clear to all).[9] "Especially at the elementary level, but also at the middle school level, it was also about knowing a young person's family or community, whoever was caring for them," says Bailey Cato Czupryk, a TNTP vice president.

These findings about school culture align with what parents tell surveyors they want from schools. Nationally representative surveys by Learning Heroes and the National Parents Union have found that parents want to know how a school's students perform on state tests.[10] Researchers have found that parents value having their students in schools with high-performing peers.[11] But parents also want schools that are safe, support their children's social-emotional development as well as their academic growth, and foster a sense of connectedness.[12]

Chicago Public Schools (CPS) commissioned a survey of public-school parents, students, teachers, and other stakeholders about the school quality measures they would prioritize as part of efforts to redesign its school accountability system. Parents rated traditional measures—like dropout rates and students' performance on standardized tests—among the bottom 10 indicators.[13] They rated how well schools develop students' social-emotional skills and prepare them for life beyond high school near the top.[14]

A more holistic view of school quality would include both survey results and test-score growth. A recent study by researchers from Blueprint Labs, a nonpartisan research group at the Massachusetts Institute of Technology, studied how well test scores, surveys, and other performance measures predict students' long-term success, using data from the New York City Public Schools. It found that while surveys better predict high school graduation, test-score growth better predicts college enrollment and persistence.[15]

A Disconnect Between Research and Practice

While the research from the University of Chicago consortium, NewSchools, TNTP, and others suggests the need for comprehensive measurement systems to evaluate schools and drive improvement, for more than two decades states have instead judged schools mainly on state standardized test results and high school graduation rates, metrics driven by school "accountability" requirements in federal law.

The No Child Left Behind Act of 2001 required states to rate schools based largely on state test

scores—as a whole and by student subgroups based on race, ethnicity, income, disability status, and language fluency. But how students perform on standardized tests alone correlates tightly with student demographics and family income, making it difficult to gauge the real contributions of schools to improved student outcomes.[16] The federal government's emphasis on test scores and exclusion

of other key indicators of student success was, as a strategy, akin to determining a person's health based solely on blood pressure readings.

Mounting evidence that test scores alone weren't doing justice to students or schools led to revisions in federal requirements under the 2015 Every Student Succeeds Act (ESSA). They opened the door for states

SCHOOL RATINGS

Studies have found standardized test scores to be highly correlated with student race and family income. As a result, school ratings based on test scores may say more about student demographics than about the quality of the school. When schools rate higher simply because their student demographics are correlated with higher test scores (as can be the case with schools serving mostly white, affluent students) the outcomes are attributable to what is known as selection bias.

Concerns that such ratings can unfairly penalize schools serving high proportions of low-income students and students of color led to a major revision in federal law in 2015. The federal Every Student Succeeds Act permits states to rate schools based on students' year-to-year achievement or progress on state tests, not just their current achievement levels. Such growth models are viewed as a fairer and more accurate measure of how much schools contribute to student learning.

But a 2022 analysis by economists at Blueprint Labs at the Massachusetts Institute of Technology found that policymakers could go even further in ensuring that ratings do not simply reflect a school's racial make-up.[1]

Joshua Angrist, winner of a 2021 Nobel Prize in Economics, along with fellow economists Peter Hull, Parag Pathak, and Christopher Walters, analyzed the correlation between middle school

ratings in Denver and New York City and the schools' racial composition. Because both districts have school choice programs that assign students by lottery, the researchers were able to compare test scores outcomes for entering sixth graders who were assigned to their schools at random, thereby controlling for selection bias.

The study found that ratings based on achievement levels, or the percent of students "proficient" on state tests, were highly correlated with the share of enrolled students who were white. Progress ratings, based on the improvement in student test scores from fifth to sixth grade, were much less correlated with race, but some selection bias remained. By using a simple adjustment called "race-balanced progress" that controlled for the share of white students in a school, the economists were able to remove the correlation with race. This new measure was at least if not more predictive of school quality than conventional progress ratings, the researchers found.

Making such adjustments is important, the economists argue, not only to provide a more accurate measure of school quality but also to discourage families from conflating school quality with student composition when they are choosing schools, which could lead to the increased racial and economic segregation of schools.

1   John Angrist, Peter Hull, Parag A. Pathak, and Christopher R. Wren. (2022). "Race and the Mismeasure of School Quality." *MIT Blueprint Labs Discussion Paper #2022.01*.

to measure a more expansive set of features of school quality and student success.

The largest number of states—36 and the District of Columbia—added chronic absenteeism to their measurement systems to capture student engagement. A handful of others introduced school climate and culture surveys, according to a 2024 survey of state accountability systems by The Education Trust, a national nonprofit.[17] Four states—Connecticut, Georgia, Iowa, and West Virginia—capture student enrollment in postsecondary education. And about half the states now produce summative school ratings, such as A-F grades or 0-100 scores. *(See sidebar on Page 5.)*

But standardized test scores in reading and math and, to a lesser extent, science, continue to carry the heaviest weight in school ratings. Few states enable comparisons of schools with similar characteristics but different outcomes, or by the performance of different student demographic groups.

Importantly, the EdTrust report found that states are "largely failing to connect their school identification process to robust systems of support" to improve schools.[18] That conclusion mirrors a study by the U.S. Government Accountability Office, which found that states and districts are not meeting some of ESSA's basic requirements for supporting their lowest-performing schools, including conducting a needs assessment and selecting evidence-based interventions.[19]

In essence, ESSA changed the definition of school quality somewhat, but left it heavily focused on reading and math performance on standardized tests, while reducing states' obligation to act on failing schools. As C. Kirabo Jackson of Northwestern says, there continues to be a "disconnect between current metrics for assessing school quality and what high-quality schools are actually doing to promote student development."

A New Measurement Model

What would a school measurement system look like if it were better aligned to the research on effective schools? It would continue to include standardized test scores, of course, but they would become one of several features of a balanced system that more fully captures student academic outcomes and the contributors to student success. The research points to five key metrics:

Standardized Test Scores as an Important, but Not Overriding, Element of School Quality

Standardized test scores would continue to play a key role in measuring school quality. They help drive high standards and more rigorous instruction in education by signaling what students should know and be able to do. They provide parents, policymakers, and taxpayers with a reliable, comprehensible way to measure and compare achievement and growth across students, schools, and districts. And they help stakeholders evaluate which policies and programs are working and which are not. But research suggests they're currently playing an outsized role in school measurement. "There's so much about schools and what students are learning and experiencing in school that cannot be captured on a standardized test," says Elaine Allensworth of the University of Chicago consortium.

Addressing that reality requires reducing the weight that test scores currently bear in federal and state accountability systems, so that they do not become the primary means by which policymakers determine school quality. Instead, the focus should be on the transparency of test score information, including the performance of traditionally underserved students, and making that information both more accessible to parents, educators, and policymakers and easier to understand.

Measures of Students' Access to Ambitious Instruction and Advanced Coursework

To achieve at high levels, students need access to challenging instruction. A report on monitoring educational equity by The National Academy of Sciences recommends these measures include availability of and enrollment in advanced courses and programs such as Advanced Placement (A.P.), International Baccalaureate (I.B.), and dual-enrollment programs. Ensuring the rigor of advanced coursework is important. The A.P. and I.B. programs establish instructional standards for their courses and final exams are scored by external evaluators trained against the standards. Many dual-enrollment programs, however, rely exclusively on local teachers or college professors to rate student performance, and their expectations of students vary. States' level of commitment to rigor also varies. Colorado, for example, merely requires only that teachers of advanced courses have a master's degree.[20]

The Academies also recommend including measures of curricular breadth, such as availability of and enrollment in courses in the arts, social sciences, sciences, and technology. *The National Study of Learning Mindsets*, a longitudinal investigation of a nationally representative sample of 9th graders in the U.S., has found that completing one or more advanced math and science classes in high school not only predicts college readiness but also later health, job satisfaction, and wellbeing. David Yeager, a professor of psychology at the University of Texas at Austin and one of the principal investigators on the study, notes that because A.P. and I.B. course-taking correlates closely with school wealth, he would focus on the percentage of students who take college preparatory math and science courses rather than just the number enrolled in A.P. and I.B. programs to get a complete picture of schools' commitment to academic rigor.

There are also other means to assess instructional expectations. TNTP's report, "The Opportunity Myth,"

observed classrooms using an instructional rubric and examined student assignments and work samples to assess the rigor of instruction.[21] School Quality Reviews, such as those conducted in the United Kingdom, typically include classroom observations, as well as reviews of student work and discussions with students and teachers to assess instructional rigor and provide feedback for schools to improve. *(See sidebar on page 8.)* The 5Essentials Survey asks students questions related to academic expectations and support. And states such as Tennessee now track students' access to high-quality, standards-aligned instructional materials by collecting data on whether districts are using materials vetted by the state.

An important reason to capture the rigor of course content as well as student participation in advanced courses and programs is that the long-standing concentration of white and Asian students in advanced programs has spawned a movement to dismantle gifted-and-talented programs, exam schools, and other advanced opportunities because they promote racial and economic segregation in public education. That stance hurts the very students that equity advocates hope to help by ensuring they wouldn't be able to take advanced courses that would help them get ahead. They're already facing a vast opportunity gap: Vanderbilt University researchers have found that high-performing students from the wealthiest 20 percent of U.S. families are six times more likely to study advanced coursework than equally high-performing students from the poorest 20 percent.[22]

A better strategy is to abandon the scarcity mentality in advanced education that forces too many talented students to compete for too few seats. North Carolina now requires that all high-scoring third graders receive advanced math coursework in fourth grade. At the same time, a former president of the National Association for Gifted Children and other experts have recommended abandoning the term "gifted," signaling as it does that innate ability rather than hard work is the key ingredient of academic success,

while perpetuating long-debunked stereotypes that differences in educational performance are rooted in race and gender.

To ensure that accountability metrics around coursework do not have unanticipated negative consequences, districts should adopt strategies to ensure courses are staffed with teachers equipped to promote skill-building and a sense of belonging among a more diverse student body.[23]

Access to High-Capacity Teachers and School Leaders

Research has consistently identified the quality of teachers and school leaders as two of the most important school-based contributors to student learning.[24] High-quality teachers influence short-term outcomes, like test scores, as well as longer-term outcomes, such as graduation, college attendance, and earnings.[25] Mounting evidence over the past decade also has underscored the benefits of a diverse teaching force, particularly for students of color.[26]

The number of highly effective or effective teachers in a school (assessed through well-vetted teacher-evaluation systems such as those used in Tennessee and the District of Columbia) could serve as a measure of school quality. Research by the University of Chicago consortium suggests that overall professional capacity within a building also matters, as measured through educator surveys of principal-teacher and teacher-to-teacher trust. The consortium has found that schools in which teachers rate teacher-principal trust, principals' instructional leadership, and teachers' commitment to their school more highly also have higher teacher ratings.[27]

Similarly, a study by researchers Matthew A. Kraft and John P. Papay of Brown University found that teachers improved their effectiveness more over time in schools with a strong professional environment.[28]

INSPECTION SYSTEMS

In the United States, measuring school quality relies heavily on quantitative data such as achievement test scores and graduation rates. But many high-performing countries use "inspection systems" that combine those and other quantitative measures with information gathered by teams of trained experts who visit schools to gather information on other features of school success outlined in this report. Their objective is to provide a more wholistic picture of school performance and help schools improve.

These school quality reviews typically include a school self-assessment followed by team site visits that delve into the quality of teaching, learning, and leadership in the building. Depending on the jurisdiction, these teams may include former or current educators who are trained in evaluation as well as staff from the inspection agency. In addition to observing classrooms and talking with students, teachers, and school leaders, the visitors review administrative documents, lesson plans, samples of student work, test scores, and other data.

The inspection teams then produce a comprehensive report describing the school's strengths and weaknesses and recommending steps to improve instruction and student outcomes. The results of the school evaluations are reported publicly. In the United Kingdom, every school in England, Wales, Scotland, and Northern Ireland is visited on a two-and-a-half to four-year cycle, depending on whether they are found "good" or "outstanding" overall or "require improvement."[1] Some jurisdictions, such as the United Kingdom and Hong Kong, also conduct separate reviews across schools to identify trends that can inform public policy.[2]

Each year, for example, the United Kingdom's chief inspector of the Office for Standards in Education, Children's Services and Skills (Ofsted) draws on its school quality reviews to produce a report on the state of education. Ofsted's 2020 report included a comparison of "stuck schools," i.e., those that had not had a rating of "good" or "outstanding" since 2007, to schools that had improved. The report found that every school could get stronger

results "if leaders concentrated on improving a small number of key things, starting with behavior and high standards of teaching and learning."

The report continued: "It is no surprise that the schools that improved did so through planning an ambitious curriculum for all, focusing on phonics in primary schools and supporting staff to be experts in their subjects." The report also showed that "unstuck" schools typically benefited from strong support from a network of like-minded schools, known in the United Kingdom as a multi-academy trust.[3]

Still, it's hard to determine whether school quality reviews improve student achievement. Economist Iftikhar Hussain recently testified before the United Kingdom's Parliament about revisions to the Ofsted inspections. While his research has found inspections are associated with short-term gains in math and English achievement for schools that received a "failing" rating, he says there is a lack of evidence about whether Ofsted judgments lead to school improvements in the long run. Meanwhile, evidence does show that school ratings influence parental choice of schools, suggesting there is a demand for information about school quality.[4]

Such inspection systems have spread rapidly around the world. In 1994, the Organization for Economic Cooperation and Development founded the Standing International Conference of Inspectorates, which now includes 43 member jurisdictions.[5] Yet

only a handful of U.S. states, including Colorado, Maryland, and Massachusetts, use school quality visits, and only a similar number of districts, among them New York City, use school quality reviews to augment their accountability systems. The practice is more common in the charter school sector, both in the authorizing process and among charter school networks, including the IDEA Public Schools, Uncommon Schools, and Success Academy networks.[6]

Cost is a barrier to school quality reviews in the U.S., including training of on-site evaluators to ensure that their judgments are valid and reliable. In the 2018-19 school year, school inspections in the United Kingdom cost an average of $9,326 U.S. dollars per school.[7] Schools also need the capacity to act on an inspection team's recommendations.

In 2005, the New York City Department of Education launched its Quality Reviews program as part of multiple measures of school quality. Experienced educators use a rubric comprised of 10 quality indicators within three categories—the instructional core, school culture, and systems for improvement—to examine how well a school supports students and their achievement. For the 2024-25 school year, school leaders can opt into the review process, which has been reduced from two days to one day and focuses on the instructional core: curriculum, pedagogy, and assessment.[8]

1   Office for Standards in Education, Children's Services and Skills. (October 2024). "School inspections: a guide for parents." https://www.gov.uk/government/publications/school-inspections-a-guide-for-parents/school-inspections-a-guide-for-parents.

2   Bob Rothman. (May 30, 2018). "Inspection Systems: How Top-Performing Nations Hold Schools Accountable." Washington D.C.: The National Center on Education and the Economy.

3   The Annual Report of Her Majesty's Chief Inspector of Education, Children's Services and Skills 2019/20. (December 2020). London: Ofsted, Gov. UK.

4   Helena Mullineaux. (August 21, 2024). Inspecting Ofsted: An Economic Perspective. University of Sussex. https://www.sussex.ac.uk/broadcast/read/64922.

5   The Standing International Conference of Inspectorates. https://www.sici-inspectorates.eu/About-us/Who-we-are-and-what-we-do.

6   Erik W. Robelen. (March 2025). "The Full Measure of a School." EducationNext Weekly. https://www.educationnext.org/full-measure-of-a-school-student-test-scores-tell-only-part-of-the-story/.

7   National Audit Office. (2018). "Ofsted's Inspection of Schools: Report by the Comptroller and Auditor General." https://www.nao.org.uk/wp-content/uploads/2018/05/Ofsteds-inspection-of-schools.pdf.

8   NYC Public Schools. "Quality Review." https://www.schools.nyc.gov/about-us/reports/school-quality/quality-review.

Teacher surveys, as well as analyses of how much principals contribute to their schools' achievement growth, can also provide meaningful measures of school leadership quality. Several studies have found higher test scores in schools where teachers rate principal instructional leadership highly.[29]

Measures of School Climate and Culture Most Related to Student Success

Many states now measure chronic student absenteeism as a proxy for student engagement and the ability of schools to provide a culture and climate conducive to learning. It is a reasonable strategy. But absenteeism is a complex problem with many causes. So it is worth using other metrics to gauge schools' capacity to foster student engagement, measures that can capture schools' influence on students' social and emotional development, as well as their sense of themselves as learners and as members of learning communities, regardless of schools' absenteeism rates.

For example, well-designed and well-implemented student, parent, and educator surveys that touch on school culture—like the 5Essentials and those developed by the CORE districts in California—can identify schools' success in increasing student commitment to learning and help educators understand student needs and respond to them effectively. They can also provide key insights into where schools need to improve.[30] Recall the finding by the University of Chicago consortium's 2023 study "Investing in Adolescents": high schools that supported student development across multiple dimensions (academic and social-emotional) had up to double the positive impact on long-term educational outcomes, including high school graduation and postsecondary enrollment and attendance, compared with schools that focused solely on boosting test scores.[31]

The report by the National Academy of Sciences' on monitoring educational equity recommends that measures of school climate include academic support, an academically focused culture, students' perceptions of safety, and teacher-student trust.[32]

David Yaeger suggests the answers to two questions in particular reveal a lot about the culture of a school: Do students feel that teachers and other adults in the school treat them with respect? And do teachers trust the principal and believe that he or she has their best interests at heart?

Measures of Student Success After High School

High test scores have long served as a proxy for a larger student goal: amassing the skills to succeed in life after high school, including college attainment, career readiness, lifelong health and satisfaction, and civic engagement.

But readiness for adulthood cannot be measured solely with a standardized test. That task requires additional measures of school quality. Such measures include whether graduates enroll in a college, community college, apprenticeship or career training program, or a branch of the armed services, and whether they complete those programs. "We often evaluate a school based on what happens in that school," says Yaeger, "but the real measure is what happens in your life after you leave that school. That's a better way to evaluate schools."

Yet, as The Education Trust report found, few states currently measure outcomes such as college enrollment when evaluating schools or have data systems that can track students' long-term education and employment success. That level of tracking requires connecting PreK-12 data systems to postsecondary and labor market data and monitoring a range of post-high-school outcomes, including new postsecondary pathways that are emerging as alternatives to four-year colleges. State data systems are improving rapidly, however. A 2024 study by the Education Commission of the States found that 35 states incorporate postsecondary data into their state longitudinal data systems and 27 incorporate workforce data.[33]

Together, these five strands of school performance—standardized test scores, access to rigorous and advanced coursework, access to high-capacity teachers and school leaders, respectful and supportive school culture, and student success after high school—provide a richer, more complete picture of a school's quality.

Responsibility for weighting each strand and the specific measurement metrics within them should rest with state and local education officials. But each should play a significant role in evaluating school success.

EQUITY INDICATORS

Broader definitions of school quality look at the many factors that contribute to student achievement, such as students' access to advanced coursework and same-race teachers, and whether those opportunities are equitably distributed across student groups.[1]

In 2019, a committee of the National Academies of Sciences, Engineering and Medicine recommended that states and school districts monitor no fewer than 16 categories of educational equity and opportunity, ranging from the depth and breadth of classroom curricula to students' perceptions of school safety.[2] Part of the impetus was to ensure that states and school districts provide schools—especially those serving historically marginalized students—with the resources they need to help students thrive. This is the same rationale behind the Every Student Succeeds Act requirement that states conduct resource allocation reviews to support high-need districts. Districts and schools identified for improvement are then in turn required to address the resource inequities their students face.

Education Resource Strategies (ERS), a nonprofit consultancy that works with states, school districts, and schools on budgeting, has designed a template to help districts conduct resource-equity analyses, starting with a diagnostic blueprint to help districts understand how they are distributing resources across 10 dimensions:

- school funding,
- teacher quality and diversity,
- school leadership quality and diversity,
- empowering and rigorous content,
- instructional time and attention,
- a positive and inviting school climate,
- student supports and intervention,
- high-quality early learning,
- learning-ready facilities, and
- diverse classrooms and schools.[3]

"We know that states are holding schools accountable to outcomes, and that's important and good," says Hassaan Ebrahim, a manager at ERS. "But you can't just improve by looking at outcomes and saying, 'I want the outcomes to get better.'"

The diagnostics, rooted in data, help school systems understand which students have access to quality experiences and what actions systems might take to improve results across the 10 dimensions of school life. In public education, Ebrahim adds, districts rather than individual schools make many decisions about resource allocation that ultimately influence school quality, which is why ERS's diagnostics are primarily designed for district use.

---

1   See FutureEd's previous report, Changing the Narrative: The Push for New Equity Measures in Education.

2   National Academies of Sciences, Engineering, and Medicine. (2019). Monitoring Educational Equity. Washington, DC: The National Academies Press. https://doi.org/10.17226/25389.

3   Alliance for Resource Equity. Dimension Guidebooks. https://educationresourceequity.org/toolkit/guidebooks/.

The Challenges to a New Measurement Model

Despite widespread interest in using more comprehensive measures of school quality, challenges remain—particularly if the results inform high-stakes decisions around staffing or school closures. A school climate survey administered for the purpose of school accountability might feel very different to students or staff than one administered with the goal of self-improvement. For example, staff members concerned their jobs are on the line might feel pressure to burnish their responses about school culture, leading to inaccurate results.[34]

To relieve this pressure while still encouraging teachers, students, parents, and other stakeholders to complete surveys, Illinois opted to rate schools on survey participation rather than survey results.

"I am a proponent of using student surveys. I'm also a proponent of using qualitative measures to understand the student experience," says Chris S. Hulleman, a professor of education and public policy at the University of Virginia who has studied student motivation extensively. "I am definitely not in favor of using them as accountability metrics."

"When you tie accountability to these metrics, they're very easy to influence, to bias, to game," he explains. "So, I would say yes to embedding these as leading indicators of school quality and tying them to improvement but not tying them to accountability."[35]

As a result, not all measures of school quality can be used responsibly in high-stakes decisions. Some, including surveys, are best used to provide a fuller picture of school quality and to help schools improve.

A second challenge is the logistics of collecting and reporting additional data such as teacher and student surveys. "You have to get participation rates up, administer the survey, make sure it's psychometrically sound," says C. Kirabo Jackson of Northwestern University. "All of these are logistical reasons why some places may not do it." Many schools administer school culture surveys, but the response rates are so low that the results cannot necessarily be trusted, adds David Yaeger.

States could help by providing free access to research-based surveys such as the 5Essentials or PERTS to ensure that schools are using well-validated measures. Other measures—such as access to and rates of advanced course-taking—could be collected as part of the federal Office of Civil Rights data collection.

Another challenge is integrating data from multiple sources to make it both readily accessible and understandable to the range of education stakeholders, especially parents. This includes survey data, academic data, and longitudinal information, including such long-term outcomes as college enrollment, employment, and measures of health and wellbeing.

States also need to do a much better job communicating the research about good schools and what goes into their school quality ratings. School measurement systems can be well-designed, but unless they're easily understood by families and communities, they're not likely to be of great value. Analyses by EdTrust and Waypoint Education Partners have found that many states provide confusing and frequently outdated information on their websites about school performance, and that schools often don't communicate individual students' test score reports to parents clearly.[36]

Better communication about the most effective indicators of school quality is important, as many parents judge schools based on factors with intuitive appeal, such as marginally smaller

classes, that data have not shown to be primary contributors to students' academic success.
In addition to providing parents with valuable information as they choose among schools, new school quality measures could challenge their

preconceptions of which schools, and which kinds of schools, excel.

The biggest barrier to the use of a wider range of research-backed school quality measures

NEW METRICS FOR MEASURING SCHOOL QUALITY

Using a broader set of school performance metrics offers a more complete picture of school quality. While all these measures can support school improvement, not all are suitable for accountability systems. Metrics that lack sufficient reliability, validity and comparability should not be used to rank schools or put educators' jobs at risk. Here we identify which measures can be confidently and responsibly included in accountability systems, which measures may be used for accountability under certain conditions, and which should be reserved for school improvement efforts only.

Can it be used for accountability?

✅ YES   ➖ POTENTIALLY   ❌ NO

GROWTH IN STUDENT ACHIEVEMENT

✅ Growth in student test scores in classrooms/schools over a school year

ACCESS TO RIGOROUS INSTRUCTION

➖ Availability of and enrollment in A.P., I.B., and dual-enrollment programs[1]

✅ Access to broad course offerings, including the arts, sciences, and technology

❌ Classroom observations

➖ Student work reviews focused on rigor[2]

❌ Comprehensive school "inspections" by trained evaluators

➖ Student surveys on teacher/school academic expectations and support[3]

✅ Use of high-quality, standards-aligned instructional materials

EFFECTIVE SCHOOL STAFF

➖ Percentages of effective/highly effective teachers (based on sound teacher evaluation systems)[4]

❌ Teacher/staff surveys of trust and principal supportiveness

➖ Analyses of principal impact on student success[5]

SUPPORTIVE SCHOOL CLIMATE

➖ Surveys of students, parents, and teachers on school safety and culture[6]

❌ Indicators of mutual respect and trust among students and staff

POSTSECONDARY OUTCOMINGS

✅ Rates of college attendance and completion, career training, military enlistment

---

1. if accompanied by student supports
2. collecting and evaluating data would need to be systematized
3. with safeguards against adult influencing of responses

4. if evaluation systems rely on multiple measures of teacher quality and trained classroom observers
5. with the use of multiple measures and several years' worth of student-achievement data
6. with sufficient validity and reliability safeguards

is that educators and education policymakers don't make it a priority. "I would say that's the biggest challenge, because people are so focused on getting test scores up," says Allensworth. "They're very focused on the hard accountability measures." This is due in part to federal ESSA requirements and in part to inertia: test scores are what has traditionally mattered most when it comes to assessing schools.

"One of the biggest holes that we're trying to dig out of," says Jeffrey Broom, the director of school quality measurement and research for Chicago Public Schools, which is currently trying to develop a broader set of school quality measures, is the idea that "outcomes equal quality." Changing that mindset will require extensive community conversations, he says, about what different sources of information can and cannot tell you.

It also will require addressing the ongoing tension between using broader quality measures for high-stakes accountability decisions—such as replacing a school's staff—and for continuous improvement.

Moving from Measurement to Improvement

Just having a richer set of school quality metrics isn't enough. Schools also need help in using the data to get better, whether they're low-performing schools targeted by policymakers for improvement or good schools working to become great schools. That requires coaching educators to understand school data and the steps needed to improve specific indicators. And it requires leadership at the school building level.

Although ESSA requires a comprehensive needs assessment for the lowest performing schools and the development of a school improvement plan, data from the Education Trust and Government Accountability Office studies show these tasks are easier said than done. "There just isn't enough support right now to help schools and districts do that well on their own," says Nicholas Munyan-Penney, assistant director of P-12 policy at The Education Trust. "Sometimes it feels like the only voices they're listening to are other school district leaders, who are equally confused about what they should be doing, or vendors who have their own agenda."

Surprisingly, perhaps, Chicago Public Schools is in the vanguard of districts implementing a richer set of measurement metrics for school improvement. Despite the turmoil in Chicago over school funding and leadership, CPS has adopted a new accountability policy to do away with summative ratings of schools and provide parents and the public with a more comprehensive picture of the conditions and learning environments necessary for student success.

In 1988, the Illinois legislature passed the Chicago School Reform Act, which emphasized bottom-up accountability through the creation of Local School Councils citywide. But since 1995, state law also has required the district to create a top-down system of school ratings, with less autonomy and more centralized decision-making for schools that fail to meet standards. The most recent iteration of that rating system, the School Quality Rating Policy (SQRP), was adopted by the school board in 2013 and, in many ways, previewed some of ESSA's principles, including a focus on measuring annual growth in student achievement.

But many community members perceived SQRP as punitive and resented that it had been developed without their input. As a result, in 2019, a new school board appointed by then-mayor Lori Lightfoot approved modifications to SQRP on the condition that the district would replace it with something different developed in partnership with the Chicago community.

That decision led to a multiyear process that engaged more than 21,000 stakeholders and resulted in the adoption of a new district policy on Continuous Improvement and Data Transparency

in April 2023. The new policy committed CPS to using a broader set of metrics to guide school improvement and the use of district resources. According to the policy, "an effective and fair

COMMERCIAL RATINGS

Widely available commercial rankings of schools—by organizations such as GreatSchools. org, Niche.com, and *U.S. News*—are often the first sources parents turn to when choosing a neighborhood or school for their children. Until recently, these organizations relied heavily on end-of-year performance on standardized tests as the source for their rankings. Yet these scores, which are strongly related to student demographics and family income, can be misleading about the quality of schools.

As a fairer way to judge schools' contribution to student learning, GreatSchools in fall 2020 updated its summary ratings to prioritize students' progress on state tests from year to year. In addition to this new focus on student growth, GreatSchools added college readiness data, including graduation rates and how students perform on college admissions exams.

One-third of Niche's rankings are based on the percentage of students at or above the proficiency level on state tests, as are 20 percent of *U.S. News*' rankings of public high schools, with another 30 percent based on the proportion of 12th graders who took and earned a qualifying score on at least one Advanced Placement or International Baccalaureate exam. But those courses tend to be more available in schools serving more affluent students.

The commercial ranking organizations are constrained in their ability to make comparisons across schools by a dearth of publicly available data. Jon Deane, CEO of GreatSchools, says his company made the decision to give more weight to student growth data as more of it became available: "There was better student progress

data in more places than there had been earlier. So, as the data got more consistent and viable in more places, we gave more weight to it."

Yet GreatSchools and other organizations don't include data on school culture, climate, and safety practices in their school ratings, even though research indicates they are important contributors to student success and parents value them. The reason, Deane says, is that "a lot of states don't release [these] measures or don't release them broadly. They don't put them in any kind of a publicly accessible place."

One exception is the state of Illinois, which administers a 5Essentials Survey to teachers and students statewide. GreatSchools has used the survey data to create a school-improvement or "thrive" badge for Illinois schools that score above a certain percentage on at least three of the five essentials, which research has found to be related to improved student achievement. Parents are more likely to want to enroll their child in a school that scores the highest on traditional indicators of academic quality. But if two schools are tied on those indicators, Deane says, parents prefer a school that scored well on the 5Essentials by a large margin.

To provide a more holistic picture of school quality, GreatSchools incorporates user reviews of schools from students, parents, staff, and others. It is trying to build tools to make it easier to leave such reviews and to make them more meaningful to others. Niche incorporates self-reported parent and student surveys on academics into its rankings, though the content of those surveys is not available on the organization's website.

approach to improving school quality considers a broad range of indicators of success, including, but not limited to student academic progress."[37]

Revised school profiles, debuting this school year, focus on four key aspects of school life:

- evidence of student learning and wellbeing, including such traditional measures as test-score achievement and growth, chronic absenteeism rates, and college enrollment and persistence;

- students' daily learning experiences, including their access to high-quality curricula, rigorous instruction, and a supportive learning environment;

- adulty capacity and continuous improvement, including school vision and continuous improvement practices; and

- creation of an inclusive and collaborative school and community, including out-of-school-time learning and enrichment opportunities and school and community partnerships.

Jeffrey Broom, CPS's director of school quality measurement and research, says stakeholders essentially wanted answers to three questions: What is it that schools should be doing to support students academically, socially, and emotionally and why are those things important? Is my school doing those things? To the extent that my school isn't doing those things, how is the district supporting its efforts to get better?

"The answers to those three questions not only meet stakeholder needs, but also drive continuous-improvement efforts," Broom says. CPS expects to publish digital profile pages for each of its 634 schools by the end of the year. While some of the 24 indicators—such as college enrollment and persistence—are already being measured, others—such as access to high-quality curricula—are expected to be ready in fall 2025.

Chicago is also addressing another important aspect of school improvement: the role of school districts in supplying schools with the funding and technical assistance they need to raise student achievement, often called "reciprocal accountability." *(See sidebar on page 11.)*

Each school is expected to develop an improvement plan and track its progress based on the data in its school profile. But the school district also will use the schools' results to set its own priorities, budgets, and staffing to better support schools. "We need to make schools' problems our problems," Broom says.

Education advocates agree, although the details of how reciprocal accountability will operate is still a work in progress. Daniel Anello, the CEO of Kids First Chicago, a parent advocacy organization, says parents and other stakeholders wanted to base school ratings on student performance, and "they also wanted to measure and hold the district accountable for living up to the promise of providing the school what it needed to be successful, which is not something that's happened in the past."

The need for dependable ways to measure the contributions schools make to the academic success of their students has never been greater—to guide the school-choice decisions of parents, to hold schools accountable for their performance, to inform policymaking, to gauge the return on taxpayer investments in education and, ultimately, to ensure that students get the academic grounding they need to be successful regardless of the postsecondary path they follow.

While outcome-based measures of student performance such as test scores and high school graduation rates are necessary elements of the best measurement models, research makes clear that they can neither provide a complete profile of a quality school nor identify the best ways to improve schools. As Daniel Anello of Kids First Chicago says: "You need to create a measurement tool that looks

at inputs, outputs, and outcomes, that is much more well-rounded than just reading and math scores."

Technical challenges, the tensions around the uses of the results, and simple inertia currently form significant hurdles to building school quality measurement models that pay the biggest dividends in student achievement. Yet efforts in

Chicago and elsewhere show meaningful progress toward the creation of a more comprehensive set of metrics better able to promote learning, inform state and district policy decisions, put more information in the hands of parents, and guide school improvement.

GUIDELINES FOR BUILDING NEW SCHOOL MEASUREMENT MODELS

Given the Trump administration's intention to reduce the role of the federal government in education policy, it's likely that improvements in school quality measures will have to come from state and school district leaders in the months ahead.

Here are guidelines they can use in developing comprehensive, research-based school-measurement models:

Prioritize a manageable set of school quality metrics based on research. These should include outcome measures such as student academic progress and postsecondary success, as well as school conditions most related to student success. Develop the metrics in partnership with local communities so they understand and trust them.

Provide access to research-based surveys and other measures of student wellbeing. This would help ensure that schools have well-validated measures to identify some of the reasons behind low student outcomes. Illinois provides all its schools with the 5Essentials survey.

Lean into the school improvement process. Ensure that districts and schools have timely access to data to conduct a need assessment for low-performing schools, as well as data analysis and coaching support. Tennessee, for example, pairs schools that are struggling with those that have successfully tackled common challenges.[1] Help schools focus on improving just a few things at a time. "We heard consistently from rapidly improving schools that 'We didn't get here by doing a bunch of big stuff all at once'," says Cato Czupryk of TNTP. "I do think, sometimes, we expect low-performing schools to do everything at once. I don't think there's any evidence that's how change happens."

Hold districts accountable for providing schools with the people, time, and money they need to succeed. Ensure that districts conduct resource allocation reviews for low-performing schools and align their support for schools with schools' needs, as Chicago is trying to do.

---

1   Craig Jerald, Kati Haycock, and Allison Rose Socol. (March 2017). For Equity-Oriented State Leaders: 9 Ideas for Stimulating School Improvement Under ESSA. Washington D.C.: The Education Trust.

ENDNOTES

1   Edmonds, Ronald. (1979). "Effective Schools for the Urban Poor," Educational Leadership, 37, 15-24.

2   Coleman, James S. (1966). Equality of Educational Opportunity, Washington, D.C.: U.S. Department of Health, Education, and Welfare.

3   Anthony S. Bryk, Penny Bender Sebring, Elaine Allensworth, Stuart Luppescu, and John Q. Easton. (2010). Organizing Schools for Improvement: Lessons from Chicago, Chicago: The University of Chicago Press.

4   Uchicago Impact. "5Essentials," https://www.uchicagoimpact.org/our-offerings/5essentials.

5   Holly Hart, Christopher Young, Alicia Chen, Anrew Zou, and Elaine M. Allensworth. (August 2020). "Supporting School Improvement: Early Findings from a Reexamination of the 5Essentials Survey," Chicago: UChicago Consortium on School Research.

6   Spela Javornik and Eva Klemencic Mirazchiyski. (2023). "Factors Contributing to School Effectiveness: A Systematic Literature Review," European Journal of Investigation in Health, Psychology and Education, 13, 2095-211. https://doi.org/10.3390/ejihpe13100148.

7   Shanette C. Porter, C. Kirabo Jackson, Sebastian Kiguel, and John Q. Easton. (April 2023). "Investing in Adolescents: High School Climate and Organizational Context Shape Student Development and Educational Attainment," Chicago: UChicago Consortium on School Research.

8   Jason Atwood. (May 7, 2024). "How an Expanded Definition of Student Success Drives Learning: Three New Insights," Oakland, CA: NewSchools Venture Fund. https://www.newschools.org/blog/how-an-expanded-definition-of-student-success-drives-learning-3-new-insights-from-the-newschools-portfolio/.

9   TNTP. (September 25, 2024). The Opportunity Makers: How a Diverse Group of Public Schools Helps Students Catch Up—and How Far More Can, New York City: TNTP. https://tntp.org/publication/the-opportunity-makers/.

10 Echelon Insights and National Parents Union. (October 2020). "National Parents Union Survey of K-12 Public School Parents," Boston, MA: National Parents Union. Gallup, Inc. and Learning Heroes. (2023). "B-flation: How Good Grades Can Sideline Parents." (2023). Washington DC: Gallup, Inc. https://bealearninghero.org/wp-content/uploads/2023/11/B-flation_Gallup_Learning-Heroes_Report-FINAL.pdf.

11 Abdulkadiroglu A., Pathak P., Schellenberg J., Walters C. (2017). Do parents value school effectiveness? (No. w23912). National Bureau of Economic Research. https://doi.org/10.3386/w23912.

12 National Urban League and UnidosUS. (2023). "Education Assessment, Accountability & Equity: 2022 Phase 1: The Final Report," Washington DC: National Urban League and UnidosUS.

13 Chicago Public Schools. (2023). "CPS Accountability Redesign: Stakeholder Survey Results," Chicago: Chicago Public Schools. https://www.cps.edu/strategic-initiatives/accountability-redesign/.

14 ibid.

15 Joshua Angrist, Peter Hull, Russell Legate-Yang, Parag Pathak, and Christopher Walters, (2025): "Putting School Surveys to the Test." Blueprint Labs Working Paper #2025.02. Boston: Blueprint Labs, M.I.T.

16 Cailyn Torpie, Noman Khanani, Bradley Quarles and Mary Rauner. (2023). Equity in Measuring School Quality: Comparing the Robust and Equitable Measures to Inspire Quality Schools (REMIQS) Framework and State Accountability Systems in Arizona, Kentucky, Massachusetts, Texas and Virginia. WestEd; KnowledgWorks.

17 Nicholas Munyan-Penney, Abigail Jones, and Shayna Levitan. (2024). Reassessing ESSA Implementation: An Equity Analysis of School Accountability Systems. Washington, D.C.: The Education Trust. https://www.future-ed.org/whos-in-chronic-absenteeism-under-the-every-student-succeeds-act/.

18 ibid.

19 United States Government Accountability Office. (January 2024). "K-12 Education: Education Could Enhance Oversight of School Improvement Activities," Washington DC: US Government Accountability Office, https://www.gao.gov/assets/d24105648.pdf.

20 Keystone Policy Center. (2024). Sweating the Small Stuff: Rural and Urban Perspectives on Post-Secondary Preparation. https://www.keystone.org/works/sweating-the-small-stuff-rural-and-urban-perspectives-on-post-secondary-preparation/.

21 TNTP. (2018). The Opportunity Myth: What Students Can Show Us About How School Is Letting Them Down—and How To Fix It. New York: TNTP.

22 Peg Tyre. (April 2024). Excellence with Equity: The Case for Rethinking Gifted Education. Washington D.C.: FutureEd.

23 See, for example: Elaine Allensworth, Takako Nomi, Nicholas Montgomery, & Valerie E. Lee (2009). College Preparatory Curriculum for All: Academic Consequences of Requiring Algebra and English I for Ninth Graders in Chicago. Educational Evaluation and Policy Analysis, 31(4), 367-391. https://doi.org/10.3102/0162373709343471 (Original work published 2009). Also:

    Charles T. Clotfelter, Helen F. Ladd and Jacob L. Vigdor, The Aftermath of Accelerating Algebra: Evidence from District Policy Initiatives.

24 Kenneth Leithwood, Karen Seashore Louis, Stephen Anderson and Kyla Wahlstrom. (2004). How Leadership Influences Student Learning. New York: The Wallace Foundation.

25 Isaac. M. Opper. (Dec. 4, 2019). Teachers Matter: Understanding Teachers' Impact on Student Achievement. Santa Monica, CA: RAND Corporation.

26 Seth Gershenson, Michael Hansen, and Constance A. Lindsay. (2021). "Teacher Diversity and Student Success: Why Racial Representation Matters in the Classroom." Harvard University Press: Cambridge, MA.

27 Jennie Y. Jiang and Susan E. Sporte, Teacher Evaluation in Chicago: Differences in Observation and Value-Added Scores by Teacher, Student, and School Characteristics. Research Report: January 2016. Chicago: UChicago Consortium on School Research.

28 Matthew A. Kraft and John P. Papay, (2014). Can Professional Environments in Schools Promote Teacher Development? Explaining Heterogeneity in Returns to Teaching Experience. *Educational Evaluation and Policy Analysis, 36*(4), 476-500. https://doi.org/10.3102/0162373713519496 (Original work published 2014).

29 Steven Rivkin, Lauren Sartain, and Jeffrey Schiman, (September 2023). Training and Retaining Effective School Leaders: Results from Efforts to Improve Principal Effectiveness in Chicago Public Schools. Chicago: UChicago Consortium on School Research.

30 Thomas Toch and Raegan Miller. (February 2019). Core Lessons: Measuring the Social and Emotional Dimensions of Student Success. Washington D.C.: FutureEd.

31 Shanette C. Porter, C. Kirabo Jackson, Sebastian Kiguel, and John Q. Easton. (April 2023). "Investing in Adolescents: High School Climate and Organizational Context Shape Student Development and Educational Attainment," Chicago: UChicago Consortium on School Research.

32 National Academies of Sciences, Engineering, and Medicine. (2019.) Monitoring Educational Equity. Washington, DC: The National Academies Press. https://doi.org/10.17226/25389.

33 Education Commission of the States. (2021). "50-State Comparison: Statewide Longitudinal Data Systems." https://www.ecs.org/state-longitudinal-data-systems/.

34 Thomas Toch and Raegan Miller. (February 2019). Core Lessons: Measuring the Social and Emotional Dimensions of Student Success. Washington D.C.: FutureEd.

35 FutureEd laid out many of the challenges to using student surveys in a 2019 report by Thomas Toch and Raegan Miller, Core Lessons: Measuring the Social and Emotional Dimensions of Student Success. https://www.future-ed.org/wp-content/uploads/2019/02/FutureEd_Core_Report.pdf.

36 Shayna Levitan. (April 2024). Making Assessment Reports More Meaningful for Students and Families. Washington D.C.: The Education Trust.

37 Chicago Public Schools, District Policy for Continuous Improvement and Data Transparency, Section 302.15. Board Report 23-0426-P04, Adopted April 26, 2023, Chicago: Chicago Public Schools. https://www.cps.edu/strategic-initiatives/accountability-redesign/.

QUALITY CHECK

THE NEW, BEST WAY TO MEASURE SCHOOL PERFORMANCE

Future*Ed*

Independent Analysis, Innovative Ideas

KEYSTONE

POLICY CENTER