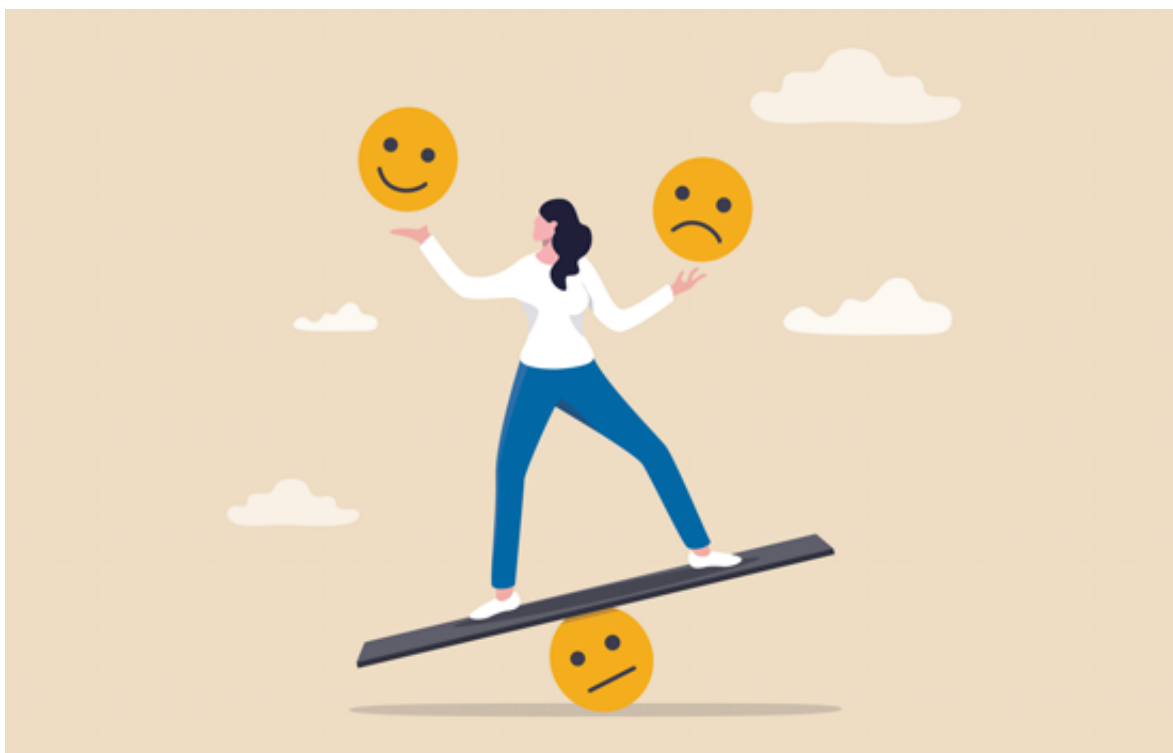# Beyond 'good' and 'bad': Disrupting narratives about school quality

James Noonan and Jack Schneider

October 24, 2022

State accountability systems don't give a complete picture of schools' strengths and weaknesses. Instead, they reflect community demographics and the inequitable distribution of resources.

Getty Images

Four-year-old Marcie's white, upper-middle-class family lived in a neighborhood largely populated by families like theirs. She was eligible for prekindergarten in her racially diverse urban school district. When choosing schools, Marcie's parents were allowed to rank up to 12 — some in their neighborhood and some in a nearby historically Black neighborhood. PreK

spots were in high demand throughout the city, but schools in their neighborhood were especially popular and had substantially longer waitlists than schools in the Black neighborhood. Marcie's family would have had a better chance of getting a spot at a school with a shorter waitlist. However, they decided to rank only those schools in their predominantly white neighborhood. Ultimately, Marcie remained on waitlists all year.

There are many reasons a family might make the choices that Marcie's did. The logistics of pickup and drop-off, for instance, can be a key determinant in enrollment decisions. Marcie's family, though — like many others from their socioeconomic background — used state accountability data when choosing a school. As they put it, they wanted to exercise due diligence. Of course, they already may have been inclined to conclude that schools predominantly serving students of color and low-income students were "bad" schools. But whether the data served as the basis for their decision or merely confirmed their beliefs, the result was the same: Marcie's family avoided the schools serving higher percentages of low-income students and students of color.

The problems with accountability data

State accountability systems currently rely on indicators that correlate strongly with demographics. Consequently, no conversation about school quality can be separate from issues of race and class. Schools perceived as "good" tend to be in better-resourced districts and enroll higher percentages of wealthy and white students. Schools perceived as "bad" tend to be in more economically oppressed districts and enroll higher percentages of low-income students and students of color. Such determinations, endorsed by the state, reinforce the racial and socioeconomic status ideologies that permeate our assessments of schools. In general, Americans have come to accept that more privileged populations maintain better access to most of life's necessities. Why should schools be any different?

State accountability determinations are portrayed as objective evaluations unrelated to race and social class, even though the indicators they use are virtually guaranteed to rank schools by variables like family income. Thanks to neutral language about "school performance," state accountability systems imply that families choosing to live in white, wealthy neighborhoods are simply making smart decisions based on data.

Ironically, then, systems intended to strengthen schools and advance equity by holding them accountable for their performance often have the opposite effect. By steering privileged families toward particular communities, and away from others, current measurement and accountability systems exacerbate residential segregation by race and class.

We are not saying that all schools are the same. Glaring inequalities exist in the resources available to different school communities. Nor are we saying that measurement as an enterprise should be tossed aside. Information about student learning and school performance is essential for allocating resources equitably and empowering families and community members to be engaged and informed parties.

What we are saying is that many of the most egregious inequalities in education are the result of systemic racism and self-segregation, which, in turn, are exacerbated by the current measurement and accountability regime. Acting on what they believe is objective information, privileged families shape not only the schools they choose, but also the ones they don't. Schools with concentrations of families with economic, social, and political capital tend to have more resources to support students. Schools with concentrations of families from poverty struggle to provide these same resources. It's a self-fulfilling prophecy, in which the labels assigned by the state drive a sorting process that privileges some schools and disadvantages others.

Measurement and accountability are not objective sciences. On the contrary, they are subjective and value-laden. Moreover, the values embedded in state accountability systems presently run counter to the broad aims of racial justice and economic equality. These systems would be bad enough if they helped privileged families find the best schools to self-select into. But as it turns out, actual school strengths and weaknesses are almost entirely irrelevant to the process.

Failing to measure school quality

Existing measurement systems are narrow in design. By focusing on student standardized test scores, they fail to capture the full range of factors that Americans care about in schools (Rothstein & Jacobsen, 2006; Schneider, 2017). Consequently, the summative ratings of schools or districts produced by these systems end up communicating less than the public may assume. Characteristics like student engagement, authentic and culturally responsive curricula, and meaningful relationships between teachers and students are excluded. Meanwhile, aggregate ratings — often in the form of A-F grades — conceal elements that are combined into a total score. A summative rating, thus, is presumed to mean something far more than it does.

> State accountability systems currently rely on indicators that correlate strongly with demographics. Consequently, no conversation about school quality can be separate from issues of race and class.

Most perniciously, in addition to capturing *too little* of what matters most, existing measurement systems also capture *too much* of what they shouldn't. Specifically, because of the correlation between student test scores and race and family income, measurement systems often indicate more about family background than school quality (Koretz, 2017; Sirin, 2005). One recent study we conducted in Massachusetts, for instance,

found that a school's achievement percentile — that is, students' raw test scores — was strongly and negatively correlated with the percentage of economically disadvantaged students in the school (-0.56), as well as with the percentage of students identifying as Black or Latinx (-0.57). These correlations were weaker for measures like "student growth percentile," which are designed to account for different student starting points. But given the weight of achievement scores in the state accountability formula, the relationship between demography and performance ratings remained durable and troubling (Schneider et al., 2021).

These limitations are not trivial. As long as we rely on narrowly tailored measurement tools correlated with student demographics, our measurement systems — and the people who use them to make or justify enrollment decisions, including families like Marcie's — will do little to solve the problem of segregation. In addition, many state accountability systems are designed so that the "lowest achieving" schools and districts are persistently designated as "underperforming." Such a designation is accompanied by punitive sanctions that may include firing teachers and closing schools. Equally significant, such labels can drive families with resources to send their children elsewhere, further concentrating less-privileged families at these schools.

This pattern is accelerated by school ratings websites like Niche.com and GreatSchools.org, which offer easy-to-peruse summative ratings of schools and districts. These sites fill a perceived market need among families making high-stakes decisions about where to live and where to send their children to school. In fact, GreatSchools.org is embedded in several popular real estate websites, allowing users to filter by school rating. Their core audience is a subset of families with the resources to buy and sell homes and the ability to transport children to school. Because such families tend to be high in social capital and material resources, steering them toward particular schools —

and away from others — can have a powerful impact on the educational environment.

The most obvious harm is done to historically marginalized and low-income students. However, summative ratings of schools derived from a narrow range of indicators undermine the aim of school improvement everywhere. If perceptions of schools are driven by something other than quality, then measurement and accountability systems are sending false signals. Even highly rated schools are poorly served by such systems, which paper over their weaknesses and praise them for qualities peripheral to their mission.

An alternative approach

In Massachusetts, eight school districts joined an effort to find a new way of talking about and measuring school quality. Their efforts have the potential, over time, to shake loose the tightly coupled relationship between segregation and school quality. Organized in 2016 and jointly governed by superintendents and teachers union presidents, the Massachusetts Consortium for Innovative Education Assessment (MCIEA) seeks to build a fairer and more effective accountability system that might be adopted statewide. As one consortium superintendent explained, "The goal is to demonstrate that school quality is more nuanced than just test scores."

Beginning in 2014, our team tested how public perceptions of schools in one urban district varied according to the kind of data they were given. Participants generally had more favorable opinions of schools when given a broader range of data (Schneider et al., 2017). In fall 2016, the data framework was revised for use with the broader consortium (Schneider, 2017).

In adopting a framework for school quality that could be used across

multiple districts with varying levels of resources and demographics, we conducted 31 focus groups with 261 participants. Those groups included students, teachers, family members, school leaders, and district administrators. They were asked to reflect on the question, "What makes a good school?" (Famularo et al., 2018). Though the focus of these conversations was measurement, discussions regularly addressed issues of equity and justice. One middle school math teacher, for instance, reflected on her own education in a "good school" — specifically, how the outward perceptions of her school felt tightly bound to district demographics. As she explained:

I grew up in a fairly rich suburban area with "good schools," and I had *terrible* math teachers throughout my high school career. But all the kids in my school did great because parents could afford to get tutors and they went to SAT prep programs, and so our school had a very high college acceptance rate. Did that speak specifically to the quality of the school or did that speak to these other factors?

The revised MCIEA School Quality Framework sought to measure factors that are important to stakeholders but seldom included in state accountability data. These factors include student engagement, the promotion of social and civic competencies, a broad and culturally sustaining curriculum, and access to the arts. Just as important, these factors are not as strongly correlated to student demographics as test scores (Schneider, 2017). Having thus established a complete set of aims, we worked to identify or develop measures aligned with this framework that could be taken from district administrative data sets, as well as from teacher and student perception surveys.

MCIEA first rolled out its framework during the 2016-17 school year, collecting surveys from more than 25,000 students in grades 4-12 and more

than 5,000 teachers. A suite of school quality indicators — from surveys and administrative data — has been collected each subsequent year, including the pandemic years in several consortium districts. The indicators are publicly available on MCIEA's School Quality Measures Dashboard. These data were used first to inform district decisions about school improvement. But our subsequent analysis revealed critical insights into the ways more holistic school-quality measurement — an approach capturing more about schools and less about student demography — could create a *more* nuanced conversation about education. By disrupting the narrative about "good" and "bad" schools, we might begin to dismantle the infrastructure that reinforces segregation. Most encouragingly, the Massachusetts Legislature recently provided funding for a new undertaking, the Education Commonwealth Project (ECP), which will make the tools we developed inside MCIEA freely available. Working with public schools and districts across Massachusetts, ECP seeks to demonstrate what a more valid, democratic, and equitable approach to assessing school quality might look like.

'Good' and 'bad' in the new system

The MCIEA/ECP approach to measuring school quality challenges popular notions of "good" and "bad" at both the district and the school level.

Contrary to the narrative about good and bad *districts*, we found that — on MCIEA/ECP indicators — school districts with very different resources seemed remarkably similar in other ways. Across school quality indicators — from physical or emotional safety to sense of belonging to civic participation — no one district appeared substantially better or worse than any other. This finding directly challenges the use of state measurement and accountability systems to rate and rank districts.

> In a truly just system, schools would be held accountable for results, but

> the state would be accountable for ensuring that each school has the capacity necessary to succeed.

Although districts appeared similar, we found their schools varied quite a bit — though not in the way one might think. Rather than revealing themselves as either good or bad, the schools in our sample were neither. Instead, the data painted portraits of highly complex institutions, all of which had strengths and areas for growth. Contrary to the notion of school quality as one-dimensional, with good schools at one end and bad schools at the other, our data showed school quality to be multidimensional and dynamic. All schools, it seems, have both strengths and weaknesses — something that current measurement and accountability systems fail to recognize, particularly considering their summative ratings.

When Marcie's family ranked schools, their preconceived notions of good and bad went unchallenged. The ratings they encountered reinforced dominant narratives. But if they had access to more and better information, Marcie's family might have acted differently. They might have seen that all the schools on their list had strengths and weaknesses. As a result, they might have thought in more nuanced ways about school fit. They might have been able to visualize Marcie thriving at schools that scored higher on the indicators that reflected her strengths, needs, and interests, regardless of location. And their conversations about race and social class, if they had them, would have been more open and direct, not cloaked in the guise of school quality.

Because we are concerned with matters of equity, we want to explicitly mention the importance of including indicators of school *capacity* alongside indicators of school *performance*. Nuanced views of school quality are useful for many reasons. However, they are insufficient without a clear understanding of whether a school has access to the resources it needs.

Current measurement and accountability systems suggest that schools improve via pressure — pressure from the state and pressure from market-style competition. This theory fails to recognize the importance of inputs, which at present are unequally distributed across schools. In a truly just system, schools would be held accountable for results, but the state would be accountable for ensuring that each school has the capacity necessary to succeed.

The late Richard Elmore (2004) wrote that accountability depends on the principle of reciprocity: "For each unit of performance I demand of you, I have an equal and reciprocal responsibility to provide you with a unit of capacity to produce that performance, if you do not already have that capacity" (pp. 244-245). This kind of "reciprocal accountability" would go a long way in transforming perceptions of good and bad. It would help the public understand how school performance in many ways reflects adequate resources.

Part of a larger movement

We are not going to solve the problem of segregated schools and neighborhoods simply by building better measurement and accountability systems. The movement to disrupt the generational legacy of segregation is going to require brave housing policy, fairer distribution of resources, and a stronger commitment to the public good. But better measurement and accountability systems might disrupt the way we think about and talk about our schools. It's possible that school quality and student demographics need not be inexorably linked.

Furthermore, new and more holistic data systems that measure what Americans care about might encourage some families to enroll their children in more diverse schools. Rather than reluctantly decamping for whiter and

more affluent districts, they might enroll their children in schools that perform well on measures that matter to them. White families may or may not send their children to schools with high numbers of students of color. However, more and better data will at least mean they'll no longer be able rationalize those decisions in terms of avoiding supposedly bad schools.

Taken together, such shifts may — over time, and with a mindful and humble approach — make possible the heavy lifting needed if we are to create and sustain a just and equitable public education system.

References

Elmore, R.F. (2004). *School reform from the inside out: Policy, practice, and performance*. Harvard Education Press.

Famularo, J., French, D., Noonan, J., Schneider, J., & Sienkiewicz, E. (2018). *Beyond standardized tests: A new vision for assessing student learning and school quality* [White paper]. Center for Collaborative Education.

Koretz, D.M. (2017). *The testing charade: Pretending to make schools better*. The University of Chicago Press.

Rothstein, R. & Jacobsen, R. (2006). The goals of education. *Phi Delta Kappan, 88* (4), 264-272.

Schneider, J. (2017). *Beyond test scores: A better way to measure school quality*. Harvard University Press.
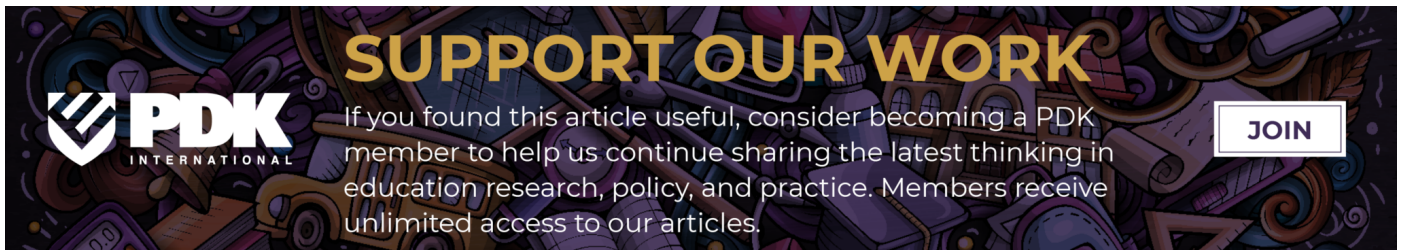
Schneider, J., Jacobsen, R., White, R., & Gehlbach, H. (2017). [Building a better measure of school quality.](#) *Phi Delta Kappan, 98* (7), 43-48.

Schneider, J., Noonan, J., White, R.S., Gagnon, D., & Carey, A. (2021). Adding "student voice" to the mix: Perception surveys and state accountability

systems. *AERA Open, 7* (1), 1-18.

Sirin, S.R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75* (3), 417-453.

This article appears in the November 2022 issue of *Kappan,* Vol. 104, No. 3, pp. 6-11.