



February 23, 2022

Jeffrey C. Riley, Commissioner
Department of Elementary and Secondary Education
75 Pleasant Street
Malden, MA 02148

Massachusetts Board of Elementary and Secondary Education
75 Pleasant Street
Malden, MA 02148

Dear Commissioner Riley and Members of the Board of Elementary and Secondary Education:

Thank you for the opportunity to comment on your proposed modifications to the school and district accountability system for school year 2021-22, as posted for public comment here: <https://www.doe.mass.edu/accountability/lists-tools/default.html>.

Regarding your proposal to suspend school-specific accountability targets for the year, it's hard to see how anyone could object to that, given the lack of reliable data from the 2020-21 school year to use as a baseline.

What troubles us greatly, however, is your stated intent to move forward with the school percentile metric¹—a metric that research shows to be biased and deeply flawed. The disruption to MCAS data caused by the pandemic only exacerbates the problems with this metric.

As you know, the school percentile metric is used to rank order schools, giving primary emphasis to MCAS achievement. Specifically, the current system weights achievement and growth using a 3:1 ratio.² Because MCAS achievement is so strongly linked to students' socioeconomic status (SES), it has long been regarded as a biased and inaccurate measure of school quality. The school percentile metric is similarly biased and inaccurate since its primary component is MCAS achievement.

There has always been a strong research base to support these concerns, and past debates in Massachusetts about how much weight to give achievement vs. growth reflected those concerns. In fact, when Commissioner Riley was the Lawrence Receiver, he was a consistent and forceful

¹ More recently, DESE has referred to this metric as an “accountability percentile.” In our comments, we treat the two terms—school percentile and accountability percentile—interchangeably.

² See *Summary of the District and School Accountability System*, June 2019, at <https://www.doe.mass.edu/accountability/lists-tools/default.html>

proponent for giving more weight to growth, as the Lawrence Public Schools routinely showed significantly higher ratings on growth than achievement, and it was understandable for LPS leaders to want credit for that progress. Other urban superintendents have made similar arguments.

Recent research elevates these concerns, underscoring the bias and error inherent in achievement scores when used as measures of school quality. Two studies in particular warrant a close look.

Center for Education and Civil Rights Study

A 2020 study³ of the Massachusetts accountability system led by UMass-Lowell Professor Jack Schneider examined the relationship between the accountability/school percentiles calculated by DESE and the racial composition of schools. Among the main findings:

- “Students of color are disproportionately represented in schools rated as low performing according to Department of Elementary and Secondary Education’s accountability system,” whereas “schools rated most highly on DESE’s accountability system, on average, are overwhelmingly white.”
- Similar disparities are evident for other historically disadvantaged groups: “Schools serving high populations of English language learners and Economically Disadvantaged students are disproportionately represented in the lowest accountability quintile.”
- “In sum, one can generally predict whether a school will be designated as lower performing...by looking at the demography of its student population...the schools most likely to be sanctioned are the schools serving the largest shares of the least advantaged.”
- These correlations, however, do not mean that schools serving high-need students are of lower quality. Because current accountability measures—namely, achievement levels—correlate strongly with student background variables, “heavy reliance on standardized achievement scores systematically disadvantages schools with higher percentages of low-income students and students from historically marginalized racial groups. As a result, the existing system appears to reflect demography more than it does school quality.”
- Despite its obvious flaws, the existing system has profoundly negative effects, including increased racial segregation: “The present accountability system not only ignores racial diversity, but appears to actively *discourage* it. Existing accountability mechanisms currently produce sanctions mostly for schools with majority populations of color; at the same time, they single out predominantly white schools for praise, drawing the attention of families with the privilege to choose where they send their children to school.”

³ Schneider, J., Piazza, P., Carey, A. J., and White, R. S. (2020). *School integration in Massachusetts: Racial diversity and state accountability*. Beyond Test Scores Project and Center for Education and Civil Rights. https://static1.squarespace.com/static/60aea369b2b1517a8d2afb0c/t/60b00f9a2dd26f3a72e1c78f/1622151076845/Demography_Report_FINAL_7.24.20.pdf

MIT Study

A 2022 study⁴ led by Nobel Prize-winning economist Joshua Angrist of MIT examined two school accountability systems—Denver’s and New York City’s—that generate ratings similar to those used in Massachusetts. The study concluded that the supposed inferiority of low-rated schools serving predominantly students of color is due to inaccuracy in the underlying measures (achievement levels) and is not due to lower-quality education. Among the main findings:

- “Selection bias drives the correlation between widely used ratings and student racial composition: many schools rate higher simply because they serve students who tend to have higher test scores regardless of school quality (e.g., higher-income students).”
- “True school quality appears unrelated to race.”
- The main driver of the discrepancy between ratings and true school quality is the reliance on achievement levels, which “are strongly correlated with race” but “only weakly related to [true] quality due to selection bias.” In contrast, progress—or growth—ratings “predict school quality much more accurately than levels ratings,” although “some selection bias remains.”
- Using inaccurate ratings based on achievement levels has troubling implications: “School ratings that fail to adjust for [selection bias] conflate differences in school quality with differences in student composition; recent research suggests such selection bias is pervasive. Rating schemes that reward family background rather than educational effectiveness are likely to direct households to low-minority rather than higher-quality schools, while penalizing schools that improve achievement for less-advantaged groups.”

Implications for Massachusetts

The Center for Education and Civil Rights study focused on Massachusetts, clearly showing the disparate impact of the DESE accountability system on low-income communities and communities of color.

While the MIT study focused on Denver and New York City, the implications extend to any state or district system that treats high achievement scores as proof of high school quality and low achievement scores as proof of low school quality. With its school percentile metric that weights achievement and growth using a 3:1 ratio, the Massachusetts rating system is certainly one of those systems, as explained well in this news article: <https://schoolyardnews.com/new-proof-that-the-state-accountability-system-is-biased-against-boston-b41e42896a57>.

This new research and the disruptions caused by the pandemic give us the perfect opportunity to pause and reflect on the Massachusetts system. We applaud DESE for its recent efforts to re-examine its policies through the lens of racial equity and justice. We wonder, however, why the school and district accountability system—with its demonstrated racial and socioeconomic bias

⁴ Angrist, J., et al. (2022). *Race and the Mismeasure of School Quality*. <https://blueprintlabs.mit.edu/research/race-and-the-mismeasure-of-school-quality/>

and clear mismeasurement of school quality—has escaped this reckoning. Why does DESE continue to label and punish schools serving students of color based on biased, inaccurate, and discredited measures?

This larger philosophical and moral question is the one we are urging you to grapple with. But even if one believes philosophically in using school percentiles in their current form, it will be nearly impossible methodologically to calculate 2022 school percentiles that have any integrity. That’s because the school percentile metric has components, such as student growth percentiles, that rely on having accurate baseline data from previous years. Also, historically, the metric has encompassed at least two years of achievement/growth data and sometimes up to four years. Yet, there were no MCAS scores in 2020 and the 2021 data are compromised, as DESE has acknowledged.

Again, the larger point is this: It’s time to pause and reflect. For 30 years, and especially since the No Child Left Behind law of 2001 and the Achievement Gap Act of 2010, this state has implemented “test, rank, label, and punish.” It hasn’t worked to close opportunity and achievement gaps, *even by DESE’s own measures*⁵, and it certainly won’t work after a pandemic that has hampered students’ social-emotional well-being and academic development.

The school percentile metric is at the heart of DESE’s biased, inaccurate, and punitive accountability system. It should be suspended—not just for 2021-22 but indefinitely. We owe it to our students to reckon with policies that have caused them harm and to abandon those policies. It’s time to develop new approaches to meet students’ social-emotional and academic needs and to foster their learning and growth. We look forward to having those conversations with you.

Sincerely,



Beth Kontos
President, AFT Massachusetts

cc: Rob Curtin, Department of Elementary and Secondary Education
Russell Johnston, Department of Elementary and Secondary Education

⁵ For example, the Board of Elementary and Secondary Education voted to place Holyoke in receivership in 2015 and Southbridge in 2016. Today, Southbridge and Holyoke are the worst-performing and second worst-performing school districts in the state, according to DESE’s most recent district rankings, found here: <https://www.doe.mass.edu/charter/enrollment/CapIncrease/>