



May 12, 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 amendment to *603 CMR 2.00: Accountability and Assistance for School Districts and Schools*, as posted here:
<https://www.doe.mass.edu/news/news.aspx?id=26728>.

The disruptions caused by the pandemic, the emergence of new research, and the experiences with your accountability system to date give you the perfect opportunity to pause and reflect on the Massachusetts accountability system – its deep harms and ongoing failures.

The proposed amendment, as published on the Department of Elementary and Secondary Education (DESE) website, reads as follows:

(11) Commissioner's authority to suspend the reporting of measures and annual performance determinations for districts and schools The Commissioner may, for good cause arising from the COVID-19 pandemic, direct the Department to suspend the reporting of any or all of the reported measures and annual performance determinations, as described in 603 CMR 2.03 (2)-(3), for school year 2021-2022.

This proposed language is quite open-ended and could be interpreted as giving DESE discretion to pursue any of three options: 1) report all measures and annual performance determinations as it did before the pandemic; 2) report no measures or annual performance determinations; or 3) report some, but not all, of the information.

AFT Massachusetts strongly supports option 2 – the suspension of all reported measures and annual performance determinations. This should happen not just this year, but indefinitely.

Even before the pandemic, it was clear that the state’s accountability system is broken beyond repair. The pandemic has only shed more light on the inequities and biases built into the system. I touched on many of these themes in my letter to you dated Feb. 23, 2022. That letter highlighted recent research showing the bias and inaccuracy inherent in the school percentile metric – the crux of DESE’s accountability system. I am enclosing the letter as part of these comments and urge you to read it again.

A biased and inaccurate rating system causes real harm to low-income communities and communities of color. Among other harms, it:

- Falsely labels schools and districts as “low-performing,” stigmatizing students, families, and educators who are working hard to overcome the learning barriers tied to poverty and others forms of socioeconomic disadvantage.
- Sends the misleading message that “low-performing” schools are only to be found in low-income communities and communities of color – the corollary being that white suburban schools are doing just fine and require no improvement. This false perception reinforces systemic racism.
- Exacerbates segregation by race and class, steering families with means to higher-ranked schools, even if these schools aren’t truly of higher quality.
- Subjects low-income communities and communities of color to DESE interventions that are top-down, punitive, and ineffective. In the communities where DESE interventions occur, the costs are high: democracy is eroded; student learning experiences are narrowed; and school communities and students suffer from disruption and instability.

Ironically, the steep costs associated with DESE ratings and interventions are often justified on the grounds – the promise – that the interventions will “accelerate student achievement” and “close achievement gaps.” In short, that they will raise MCAS scores and improve school rankings.

But, according to DESE’s own data, they don’t even do that. Here is the track record of DESE’s interventions to date:

- Southbridge was put into receivership in 2016 – it is now the worst-performing district in the state, according to DESE’s rankings.¹
- Holyoke was put into receivership in 2015 – it is now the second-worst-performing district.²
- Springfield has undergone multiple state interventions since 2010 – it is now the third-worst-performing district.³

¹ See <https://www.doe.mass.edu/charter/enrollment/CapIncrease/>; click on “2020 district ranking.”

² See <https://www.doe.mass.edu/charter/enrollment/CapIncrease/>; click on “2020 district ranking.”

³ See <https://www.doe.mass.edu/charter/enrollment/CapIncrease/>; click on “2020 district ranking.”

- Looking specifically at schools in the DESE-run Springfield Empowerment Zone (established in 2014), these schools are among the lowest-ranked schools in the state. Of the 3,904 students who attend eight non-selective Zone schools with school percentile data from 2019, the most recent year available, 81 percent attend schools that DESE ranks in the bottom 5 percent – that is, with a school percentile of 5 or below.⁴
- Even highly touted Lawrence, which saw an initial uptick after being put in receivership in 2011, is back in the lowest 6 percent of districts,⁵ and residents are clamoring for local control.⁶
- Schools taken over by DESE have not fared much better. DESE declared four schools “chronically underperforming” in 2013 and has had full operational control since that time. Nine years later, no school designated as chronically underperforming has exited that status, presumably because there has not been enough progress to warrant release.

Given the significant harms caused by the rating system and interventions, and given the failure to deliver on promised results, it seems that some degree of self-reflection or policy course correction should occur before simply resuming business as usual. But Commissioner Riley’s March 14, 2022, memo on this topic to the Board indicates that getting back to business as usual is exactly what DESE intends to do, particularly regarding school percentiles. In that memo, Commissioner Riley writes:

*In the fall of 2022, the Department proposes to produce some, but not all, of the information associated with annual district and school accountability determinations. Under this “accountability lite” model, we intend to publish district, school, and student group-level performance data for each of the approved accountability indicators, as well as certain normative measures (e.g., **school percentiles**) that will help districts, schools, the Department, and the public assess the performance of all districts and schools using common measures. [bolding added for emphasis]*

While the memo goes on to say that DESE intends to suspend some measures this fall, those measures are generally ones that rely on comparisons to 2021 MCAS data, which DESE has already acknowledged to be unreliable given the disruptions to MCAS testing caused by the pandemic. The implication is that DESE will resume reporting and using these other measures as soon as doing so is practical and technically feasible. This displays an unwillingness to engage in any kind of serious self-reflection around the impact of harmful and failed policies.

⁴ See AFT Massachusetts research brief at https://btu.org/wp-content/uploads/2022/04/Springfield-Empowerment-Zone-research-brief_FINAL_April-20229.pdf.

⁵ See <https://www.doe.mass.edu/charter/enrollment/CapIncrease/>; click on “2020 district ranking.”

⁶ This WBUR radio broadcast features local Lawrence voices asking for restoration of local control: <https://www.wbur.org/radioboston/2021/10/21/lawrence-schools-receivership>

In my Feb. 23 letter, I applauded DESE for its recent efforts to re-examine its policies through a racial equity lens. But the question I asked bears repeating: Why has the accountability system—with its demonstrated racial and socioeconomic bias and clear mismeasurement of school quality—escaped this reckoning? Why does DESE continue to label and punish schools serving students of color based on biased, inaccurate, and discredited measures?

In the medical profession, practitioners are taught to “first, do no harm.” The same principle should apply here. The current “test, rank, label, and punish” approach to accountability is causing serious and demonstrable harm. Until that harm can be stopped, all accountability measures and DESE interventions should be suspended – not just for 2021-22 but indefinitely. It’s time to pause, reflect and reimagine school-accountability policy. We urge you to make that a top priority, and we look forward to being part of that work.

Sincerely,

A handwritten signature in cursive script that reads "Beth Kontos". The signature is written in black ink and is positioned above the typed name and title.

Beth Kontos
President, AFT Massachusetts

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

Enclosure: Letter dated Feb. 23, 2022, from AFT Massachusetts President Beth Kontos to
Commissioner Jeff Riley and Board members



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 metric⁷—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.⁸ 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.

⁷ 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.

⁸ See *Summary of the District and School Accountability System*, June 2019, at <https://www.doe.mass.edu/accountability/lists-tools/default.html>

In fact, when Commissioner Riley was the Lawrence Receiver, he was a consistent and forceful 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 study⁹ 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.”

⁹ 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 study¹⁰ 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

¹⁰ Angrist, J., et al. (2022). *Race and the Mismeasure of School Quality*. <https://blueprintlabs.mit.edu/research/race-and-the-mismeasure-of-school-quality/>

school and district accountability system—with its demonstrated racial and socioeconomic bias 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*¹¹, 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

¹¹ 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/>