



DeAngelis: 5 Reasons Why 5-Point Takedown of AEI Report on School Choice, Tests, And Long-Term Outcomes Misses the Mark

Corey DeAngelis

April 26, 2018

They say you shouldn't throw stones if you live in a glass house. Fordham Institute's President Michael J. Petrilli recently threw five of them. He released a five-part series that supposedly showed why an American Enterprise Institute report incorrectly concluded that "a school choice program's impact on test scores is a weak predictor of its impacts on longer-term outcomes." But Petrilli's flawed dissection of the AEI analysis failed to invalidate its conclusions or to establish that "impacts on test scores matter." Here's what he got wrong.

The most important weaknesses of the five-part series had to do with methodology. First, all of Petrilli's re-analyses were based on a highly truncated sample. In fact, he dropped over a third of the original review's studies linking test scores to high school graduation. Petrilli believed this was justified because they were not what he considered "bona fide" school choice programs. For example, he argued that career and technical education school evaluations should be dropped because "high-quality CTE could still boost high school graduation, postsecondary enrollment, and postsecondary completion" even though those schools spend less time shaping skills that are captured by standardized tests.

But that is precisely the point. A disconnect between programs' effects on standardized test scores and on long-term outcomes suggests that test scores are not good proxies for long-term success. And if we regulate teachers and schools based on them, educators may have a perverse incentive to focus less on the character skills that are necessary for true lifelong success.

And, in every single case, the dropped studies included schools that students chose to attend. But it shouldn't even matter if the studies were of schools of choice. Finding a divergence between short- and long-term outcomes — from any type of educational evaluation — should cause us to question the validity of test scores. Put simply, Petrilli should not have dropped over a third of the original report's observations.

But assume that dropping observations was a good call. The much more astonishing methodological error was in counting null results as positive or negative. Petrilli argued that it would be “reasonable” to look for matches by “seeing whether a given study’s findings point in the same direction for both achievement and attainment, regardless of statistical significance. In other words, treat findings as positive regardless of whether they are statistically significant, and treat findings as negative regardless of whether they are statistically significant.” No serious social scientist would call that approach “reasonable.” This is because null results are statistically indistinguishable from zero.

But even when treating zeros as positive or negative, Petrilli still found disconnects between test scores and high school graduation 35 percent of the time for math and 27 percent of the time for reading. However, the original report found that 61 percent of the effects on math test scores — and 50 percent of the effects on reading test scores — did not predict effects on high school graduation. In either case, effects on test scores are unreliable predictors of effects on attainment.

But that’s not all. The literature finding divergences isn’t limited to high school graduation and college enrollment. I have started to compile more evidence of these divergences that exist in the most rigorous private school choice literature. I’ve already found 11 disconnects between private schools’ effects on test scores and their effects on other arguably more important educational outcomes, such as tolerance of others, political participation, effort, happiness in school, and adult crime. For example, an experimental evaluation of a private school voucher program in Ohio found that winning the lottery to attend a private school had no effects on test scores but a 23 percent increase in students’ charitable donations in a lab setting.

And methods aren’t the only problem. There are some important logical errors to note as well.

Petrilli correctly points out that higher graduation rates could simply mean that individual schools have lowered their standards. In other words, high school graduation rates can be gamed. We have recent evidence of this in D.C. public schools. But Petrilli fails to point out that the same problem of gaming also applies to standardized tests. We also have lots evidence of this from places like Atlanta. In fact, the corruption involved with using top-down metrics — of any kind — for accountability is so widespread that social scientists have given the principle its own name: Campbell’s Law. This is just another reason why we should not regulate schools based on top-down metrics like test scores or even graduation rates.

But assume that no disconnects existed in the literature. Let’s also assume that test scores were indeed valuable predictors of all long-run outcomes we actually cared about. And let’s further assume that it was impossible to game the metric.

Regulators would still have a severe knowledge problem. How would they know which schools were the best at shaping test scores? Average test score levels would tell us nothing about how well the schools improved them. However, we could look at test score growth instead. And if the regulators were highly informed, they could use one of the most rigorous econometric methods social scientists currently have to determine schools’ effects on test scores: value-added methodology. The problem is that value-added methodology relies on the assumption that

children are randomly assigned to schools. By definition, schools of choice fail that assumption. Because, you know, kids don't choose their schools at random.

In other words, even if we all believed test scores were valuable, and even if regulators used the best methodology available, they could still close down schools that were doing good things for their students.

But what happens when regulators close schools that are actually low-quality? Obviously, this causes disadvantaged children to switch schools, which itself has been found to reduce student achievement by over two months of learning. But that isn't the only problem. Closing an objectively low-quality school could mean that children are displaced into an even worse institution. And there is absolutely no guarantee that a better institution will magically pop up.

The fact is, several studies show that test scores are weak proxies for the outcomes we actually care about. The weak predictive power of test scores suggests that policies incentivizing teachers and schools to improve these crude metrics could actually harm students in the long run.

But families already know this. When given the chance to choose their children's schools, families consistently prioritize things like school culture and safety over standardized test scores. Maybe families know a little something about their own kids that the experts don't know. And maybe the experts should learn to leave them alone.

Corey A. DeAngelis is an education policy analyst at the Cato Institute's Center for Educational Freedom.