



APPLES TO AARDVARKS: WILL REPLICATES ERRONEOUS PRIVATE SCHOOL FINDINGS



Last year, NEPC published a [review](#) of a [report](#) from the Wisconsin Institute for Law and Liberty (WILL) titled, *Apples to Apples: The Definitive Look at School Test Scores in Milwaukee and Wisconsin*. Our reviewer found the report more opaque and misleading than definitive. A year later, WILL has released a new “Apples to Apples” report, which fundamentally repeats the errors of the first.

Each report compares public and private school sectors for a single year. Yet measuring effectiveness requires at least two years of data, since different schools start from different places. This shortcoming is compounded by selection bias: Different populations with different test scores choose to attend different schools. When schools have different scores, is it the type of school that makes the difference or is it the students who attend? In Wisconsin, as in many other states, private or charter schools sometimes even have admission requirements, which means some schools have test score advantages before they get out of the gate.

The WILL researchers attempt to resolve this selection bias problem by considering (controlling for) outside factors in their analyses. This is the basis for the “apples to apples” claim. Unfortunately, this consideration is limited to five apples out of a decent-sized barrel: (1) school-based enrollment counts, (2) race, (3) ELL status, (4) economically disadvantaged and (5) grade levels served by the school (p. 5). The analyses do not include prior test scores of individuals or even schools. That is insufficient for claiming one sector is performing better or worse than another.

Our reviewer noted substantial missing data for private schools—which the WILL report acknowledges, yet still insists it is producing “something approximating an ‘apples-to-apples’ comparison.” Then, the school is rated on the aggregated percent proficiency for schools, rather than a full continuum of test scores. This is arbitrary and different grades and subject matters have different cut scores. An improvement from last year is the use of the now com-

monly required ACT test, but how this is used is unclear.

Perhaps the most unusual research method they applied involves the calculation of disability rates. Not trusting the state's reported rates, they used estimates from an earlier University of Arkansas study on Milwaukee. This assumes the whole state has the same disability rates as the city. But that's not all. The earlier study was based on an estimated range rather than a count. Referring to the high end of this estimated distribution, "I *assume* (emphasis added) the disability rate is a factor of 8.125." (p. 14). That is, they eyeballed the data, plucked a number and used it as the basis of their statistical analysis. Strange things happen when you pick a number – such as one school having "a disability rate exceed(ing) 100%" of their enrollment. The report acknowledges that these assumptions are "very rough."

Most of these shortcomings were explained in last year's expert review by Professor Ben Shear, and they were thus available to the WILL researcher, yet they were repeated in the latest report. The result is not apples-to-apples, it is applesauce.

NEPC Resources on Accountability and Testing

The National Education Policy Center (NEPC), housed at the University of Colorado Boulder School of Education, produces and disseminates high-quality, peer-reviewed research to inform education policy discussions. Visit us at: <http://nepc.colorado.edu>