

"Federation Corner" column
The Montgomery Sentinel - August 7, 2008

How Weast blames Simpson's Paradox for his failures

by Wayne Goldstein
Chair, MCCF Education Committee

Simpson's Paradox does not concern which choice of disastrous paths of least resistance that Homer Simpson will next make. Instead, Simpson's Paradox is a mystery that has been observed for 105 years and repeatedly studied for its statistical, philosophical, and public policy implications. Although Udny Yule first described it in 1903, it is now called "Simpson's Paradox" (or the Yule-Simpson effect), for the 1951 work of Edward H. Simpson, instead of, say, "Yule's Rule." It "is a statistical paradox in which the successes of several groups seem to be reversed when the groups are combined."

Superintendent Weast explained it as such in February 2005: "The mathematical effect that occurs when student achievement increases for all subgroups of students but the overall average score declines because the percentage of low-scoring students increases." Weast should be quite glib with his explanation, because he has been using it as a way to avoid responsibility for declining test scores almost from the day he arrived here in August 1999. In fact, at a September 1999 staff meeting, he provided a graph that "shows the mathematical effect that occurs when student achievement increases for all groups of students, but the overall average score declines because the percentage of low-scoring students rises."

In February 2000, a question from a BOE member about declining MCPS MSPAP (Maryland School Performance Assessment Program) test scores brought this response: "Using Simpson's Paradox, Dr. Weast opined that the systemwide scores might continue to fall over the next two years; however, full implementation of the new programs in the 55 schools that were the most affected by a significant decline in student achievement would eventually raise scores and every child would be educated to his/her full potential."

A study of immigrants in suburbia in the 2000s, that refers to an August 2002 news account about SAT scores, states: "The influx of minority children--black, Asian and Latino, many of them immigrants, some of them poor--threatened to disrupt the unwritten compact described by Weast. For one, these new students have lower test scores than their white native born counterparts, and these lower scores could potentially lower the average test results for the county school system, and hence do damage to their national reputations and standing. This is what Weast referred to as Simpson's paradox: the idea that scores for the school system's traditional middle-class white students could be going up, but average scores could still be dragged down by the performance of lower performing students, many of them immigrants and minorities."

In August 2004, Weast wrote: "The overall average [SAT] scores for African American and Hispanic students largely did not change, remaining steady among African American (917) and declining by one point among Hispanic students (944). However, the average scores for African American and Hispanic students receiving federal meal assistance increased significantly--up 17 points among African American students (854) and six points among Hispanic students (837).

"This is an important point because the percentage of both groups of students participating in the Free and Reduced-price Meal System (FARMS) is increasing, lowering the overall average score for each group even as the specific subgroup scores increase. In the case of Hispanic students, the average scores for both FARMS and non-FARMS students increased but the overall average declined. This effect is known mathematically as 'Simpson's Paradox' and figures prominently in our understanding of the

effects of the changing demographics of Montgomery County and the impact of greater poverty among students. Overall, the average score among students receiving meal assistance increased by 16 points to 894... The average score among students in the English for Speakers of Other Languages (ESOL) program was 757, down 54 points from 2003..."

This memo was followed by Weast's February 2005 statement I quoted at the beginning of this column. In a September 2005 study: "An Examination of the SAT Results for the Class of 2005," done by the MCPS Department of Shared Accountability, the author wrote: "The disparity between the overall and subgroup changes in mean SAT scores of female students was the result of a mathematical effect known as Simpson's Paradox. Simpson's Paradox can be understood by examining two things--changes in SAT participation and relative SAT performance of females of different races/ethnicities. Compared with the Class of 2004, a greater proportion of female SAT test takers in the Class of 2005 were African American or Hispanic students, two groups whose mean SAT combined verbal and math scores are below the district average. The greater number of low SAT scores relative to high SAT scores resulted in a statistically significant decrease in the overall mean for females, even though SAT performance of females of different races/ethnicities held steady or improved significantly."

According to an August 2007 news account: "Montgomery students took the test in record numbers, Superintendent Jerry D. Weast said. The county's performance was 'a perfect Simpson's paradox,' a statistical scenario in which the success of individual groups reverses when they are combined, he said. Scores rose for Hispanic and white students in Montgomery and declined for black and Asian students. But the overall county average was down 10 points. The county average is pulled down, Weast said, by students who are new to college-entrance testing and score low because of poverty and inadequate preparation. Participation on the test reached 79 percent in the county, an all-time high, with black and Hispanic students accounting for three-quarters of the increase."

An August 2007 MCPS press release also stated: "The MCPS analysis of the exam results yielded important findings. First, the longer a student studies in MCPS, the better he or she performs on the test. On average, students with only one to three years experience in MCPS scored substantially lower (181 points) on the SAT than those with four or more years in MCPS. The effect is more apparent among African American students than any other group, with African American students here longer than four years scoring 84 points higher than African American students here only one to three years. The second finding demonstrates that the greater level of participation means more under-prepared students are taking the exam, creating a variation of the complex mathematical effect known as 'Simpson's Paradox.' The paradox occurs when high scores continue to improve or remain strong but no longer offset the comparatively low scores of a larger number of students."

Let's recap the number of different ways that Weast has blamed Simpson's Paradox for unsatisfactory test scores:

1999: It's an increase in "low-scoring students/"

2000: "Full implementation of the new programs" would overcome the effect of Simpson's Paradox and "would eventually raise scores and every child would be educated to his/her full potential."

Mid 2000s: "The influx of minority children--black, Asian and Latino, many of them immigrants, some of them poor--average scores could still be dragged down by the performance of lower performing students, many of them immigrants and minorities."

2004: "Students participating in the Free and Reduced-price Meal System (FARMS) is increasing, lowering the overall average score for each group even as the specific subgroup scores increase."

2005: "A greater proportion of female SAT test takers in the Class of 2005 were African American or Hispanic students, two groups whose mean SAT combined verbal and math scores are below the district average."

2007: "Students with only one to three years experience in MCPS scored substantially lower (181 points) on the SAT than those with four or more years in MCPS. The effect is more apparent among African American students than any other group, with African American students here longer than four years scoring 84 points higher than African American students here only one to three years. The second finding demonstrates that the greater level of participation means more under-prepared students are taking the exam."

As you can see, Weast's targets of blame for Simpson's Paradox change almost from year to year, from "low-scoring students" to all minority immigrants to FARMS participants to female minorities to African-American students attending MCPS for less than 4 years to under-prepared students. Explaining the meaning of Simpson's Paradox is about showing that ALL groups are making progress, even if at different rates and from different starting points. But Weast undermines his all-purpose excuse when he also writes: "The average score among students in the English for Speakers of Other Languages (ESOL) program was 757, down 54 points from 2003," and does not say whether minorities attending MCPS for less than 4 years scored better compared to previous years. If standard sub-groups, as well as ones of his own creation, score lower from year to year, that is not an example of Simpson's Paradox at work. That's evidence that such students have not been "educated to his/her full potential."

Jerry Weast appears to be the only superintendent in the country who has used Simpson's Paradox as the primary excuse to explain away why his school system's SAT test scores have not risen for certain groups of students. It shows the lengths to which he will go to make use of even the most arcane of concepts to create "Weast's Excuse" to free him from accountability.