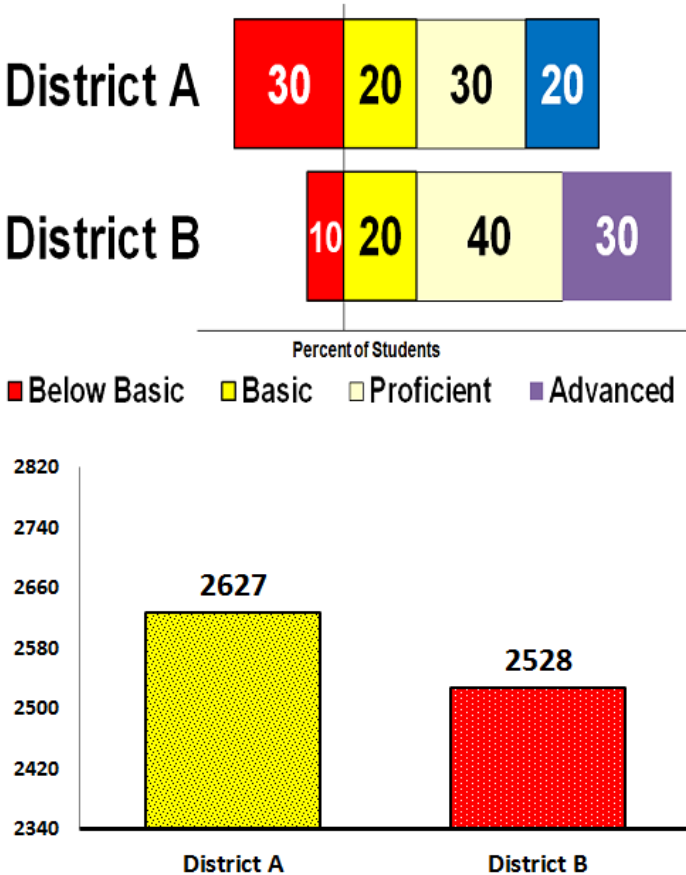




## Scale Scores Are Needed to Understand Achievement Test Results.

Which district (A or B) had the better results for this ISAT/SBAC 8th grade mathematics test? Simulated graphs of achievement level percentages and average scale scores are below.



Using SBAC achievement levels, District B had the better performance with 70% Proficient or Advanced with only 10% scoring Below Basic. District A had only 50% Proficient or Advanced with 30% scoring Below Basic. Average SBAC scale scores, however, tell a different story. District A's average score of 2627 was better than District B's 2528. Describing the average scale scores by achievement level, District A can be labeled as Proficient and District B as Basic.

Collapsing multiple scale scores into only four achievement levels *always* creates error in the original meaning of the data. When the error is great enough, achievement level results will wrongly contradict scale score results.

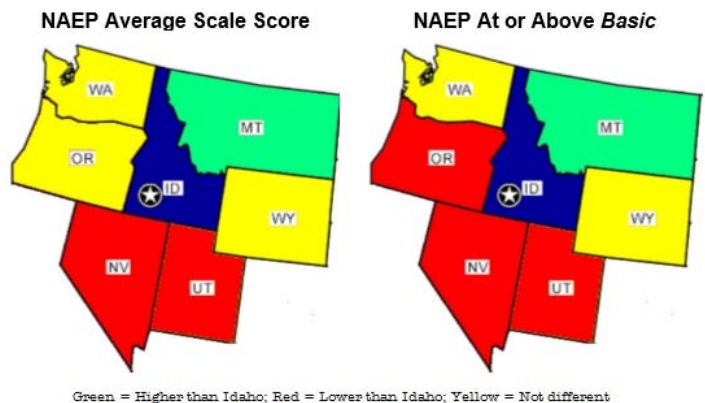
The State Board of Education and the State Department of Education have released only achievement level results from the 2015 ISAT/SBAC. When achievement level results are reported, scale score results should also be reported. *Scale score results can stand alone, but achievement level results cannot.*

NAEP implemented achievement level scores in 1990, The National Assessment Governing Board has done more than any other test publisher to make them as meaningful and useful as possible, *but major problems still persist.*

The NAEP Reauthorization Act (P.L. 107-279, Title III) states, "The [NAEP] achievement levels shall be used on a trial basis until the Commissioner for Education Statistics determines [snip] that such levels are *reasonable, valid, and informative to the public.*" No Commissioner has ever made such a determination. NAEP documents reporting achievement level results must state they are still used only on a trial basis.

NAEP believes achievement levels are useful for looking at trends, but they do not have the qualities necessary for use in statistical trend analysis. Ho (2007) noted, "As useful as [achievement levels] statistics have been in communicating test results to the public, their properties as trend statistics render them ill-suited for trend comparison."

Side-by-side displays of NAEP scale scores and achievement levels from one test may provide contradictory interpretations of the results. From the NAEP 2009 8th grade mathematics test (go with average scale scores):



Green = Higher than Idaho; Red = Lower than Idaho; Yellow = Not different