

FYI limited interpretations for “achievement level percentages” used by NAEP (and ISAT/SBAC).

Regarding the most recent congressionally mandated evaluation of the NAEP program: Edley, C., Koenig, J.A., Editors. (2017). *Evaluation of the achievement levels for mathematics and reading on the National Assessment of Educational Progress*. Washington, DC: The National Academies Press. Available online at <https://www.nap.edu/23409> or <https://eric.ed.gov/?id=ED571144>

Since 1992, NAEP has reported results in relation to three achievement levels: Basic, Proficient, and Advanced. However, the use of achievement levels has provoked controversy and disagreement, and evaluators have identified numerous concerns. This study examined the NAEP student achievement levels for reading and mathematics in grades 4, 8, and 12, to determine whether the achievement levels are ***reasonable, reliable, valid, and informative to the public***, and recommended ways that the setting and use of achievement levels can be improved.

A paragraph on page 208 of this 2017 evaluation report points out the misleading nature of using NAEP achievement level percentages, and the serious limitations of the “percent above cut-score” metric in general. The findings of this study directly relate to the ISAT/SBAC achievement level percentages.

1. Problem with achievement level percentages (trends over time):

“One of the most common and unwarranted inferences [using achievement level percentages] involves assessing the amount of progress students have made over time, particularly by population groups. For instance, news reports often focus not only on how students are doing at a particular time, but the extent to which the percentage of students scoring ***Proficient or above*** has (or has not) improved over successive NAEP years. When these comparisons are based on the scale scores, they provide useful information. When they are based on the “***percent Proficient or above***” metric and used to compare progress across groups, they can be misleading.

2. Problem with ALL percent above cut-score metrics (trends over time):

A report by Holland (2002) on this issue focused on misinterpretations associated with using the “percent above a cut score” metric. Although this metric is widely used (for NAEP and many other achievement tests), there are serious limitations to the inferences that [the percent above cut score metric] can support, particularly when evaluating ***trends over time, gaps among groups, or trends in gaps....***”

May 10, 2019

