



**K-12 Research Idaho**  
A Personal Public Service Project

May 27, 2016

TO: Alison.Henken@osbe.idaho.gov & Accountability Oversight Committee  
FROM: Bert Stoneberg Jr.  
SUBJECT: Comment on Idaho K-12 Accountability System Framework (April 2016)

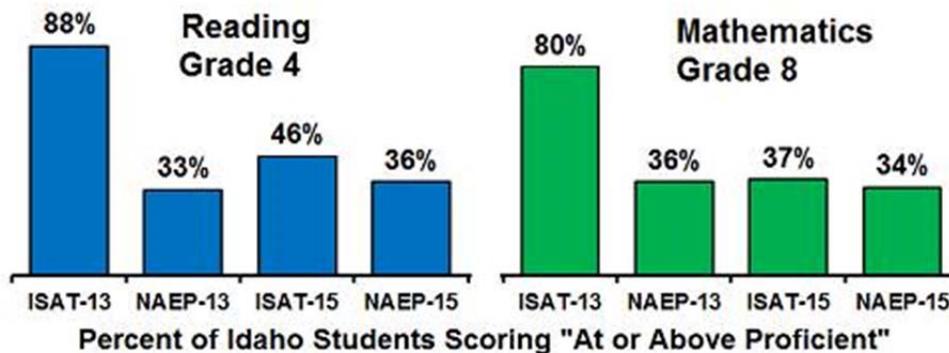
**THE PROBLEM.** Use of Percent Proficient Scores (i.e., Percent of Students Scoring Proficient or Advanced, as defined and used in ISAT and NAEP Achievement Level Results)

The U.S. Department of Education's proposed regulations for ESSA require each state's system of annual meaningful differentiation of schools to --

**Proposed §200.13.** Include *at least three distinct levels of performance* for schools on each indicator that are *clear and understandable to the public*, and set those performance levels in a way that is consistent with the school's attainment of the State's long-term goals and measurements of interim progress.

*Idaho K-12 Accountability System Framework (April 2016)* defines "ISAT Proficiency" as students who have scored proficient or higher on the ISAT, demonstrating appropriate grade-level knowledge. I suggest that achievement level indicators are neither clear nor understandable to the public.

The ISAT/SBAC developers made a concerted effort to define and produce achievement levels in line with NAEP's proficiency levels. These results from ISAT-13, ISAT-15 vs. NAEP demonstrate that the effort was rather successful.



NAEP has been using achievement levels since 1990. The National Academy of Sciences, in a Congressional mandated evaluation of NAEP, found that NAEP's achievement-level-setting procedures are fundamentally flawed. The judgment tasks are difficult and confusing; raters' judgments of different item types are internally inconsistent; appropriate validity evidence for the cutscores is lacking; and the process has produced unreasonable

results. *NAEP's achievement levels should be used on a developmental basis only.* Reports should strongly and clearly emphasize that the achievement levels are still under development, and should be interpreted and used with caution. (Pellegrino, 1999).

Why would the State Board of Education plan to make important decisions about Idaho schools based on "trial data"? No test publisher has ever been more diligent than the National Assessment Governing Board and the National Center for Education Statistics in developing an assessment of the highest quality. NAEP was developed via the technology of Item Response Theory, which estimates individual student scale scores, but NOT cut-scores for achievement levels. Cut-score setting has been and continues to be a political process.

Even post NCLB, NAEP achievement level percentages are supposed to be reported only on a trial basis, and to be interpreted and used with caution. **Federal law still requires that the NAEP achievement levels be used on a trial basis until the Commissioner of Education Statistics determines that the achievement levels are "reasonable, valid, and informative to the public."** So far, no Commissioner has made such a determination, and the achievement levels remain in a trial status. See <https://nces.ed.gov/nationsreportcard/achlevdev.aspx>

**THE SOLUTION.** The Idaho Accountability plan should call for collecting data and applying tried-and-true statistical methods commonly in use for program evaluation and educational research before politics -- rather than science -- demanded the use of problematic achievement level percentages for "accountability." **ISAT provides strong individual student scale scores.** These interval level data enable setting understandable and achievable goals and enable statistical methods for analysis of scale score averages, percentiles and effect sizes to determine progress on those goals.

A short paper from Learning Point Associates is available that recognizes averages, percentiles, and effect sizes as the past and future core of K-12 accountability, program evaluation, and educational research, and briefly describes what each of the three statistics can and cannot tell us. Highly recommended.

# # #

Reference Materials:

Learning Point Associates. (2009). *Connecting RESEARCH to PRACTICE: Knowing Who is Proficient Isn't Always Sufficient.* Available online:  
<http://files.eric.ed.gov/fulltext/ED509965.pdf>

Pellegrino, J.W., Ed. (1999). *Grading the Nation's Report Card: Evaluating NAEP and Transforming the Assessment of Educational Progress.* Available online:  
<http://files.eric.ed.gov/fulltext/ED446096.pdf>

Stoneberg, B.D. (2015). *The relationship between the rigor of a state's proficiency standard and student achievement in the state.* Available online:  
<http://files.eric.ed.gov/fulltext/ED558198.pdf>

NOTE: ERIC refers to a collection of studies and presentations in the Education Research Information Center, which is maintained by the U.S. Department of Education.