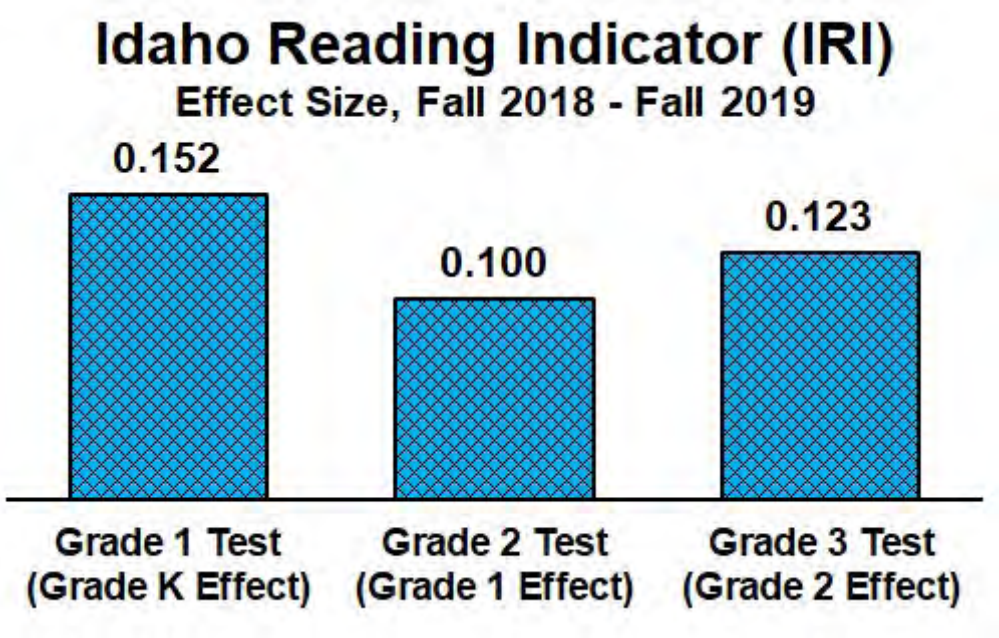
Bert Stoneberg, Ph.D.
K-12 Research Idaho
Idaho Educ Statistics

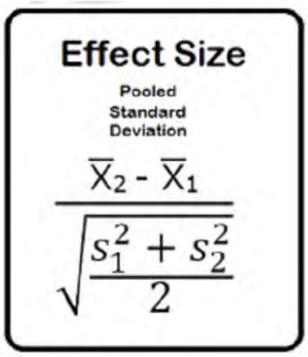
Idaho Reading Indicator (IRI) Effect Sizes

Fall 2018 to Fall 2019, Grades K-2, All Students

The IRI "simple effect size" enables us to compare how two or more grade-level groups performed by putting their Fall to Fall IRI average test scores onto the same scale. The proposed use of the IRI test is to evaluate the state reading initiative. Since kindergarten students have not had any experience in the state reading initiative when they enter school in the fall, it does not make sense to look at kindergarten progress from fall of last year to fall of this year. However, Grade 1 test data lets us look at student learning from the end of kindergarten last fall and the end of kindergarten this fall; i.e., it enables the calculation of a kindergarten effect size. Likewise, Grade 2 test data reveal a Grade 1 effect size, and Grade 3 test data reveal a Grade 2 effect size.

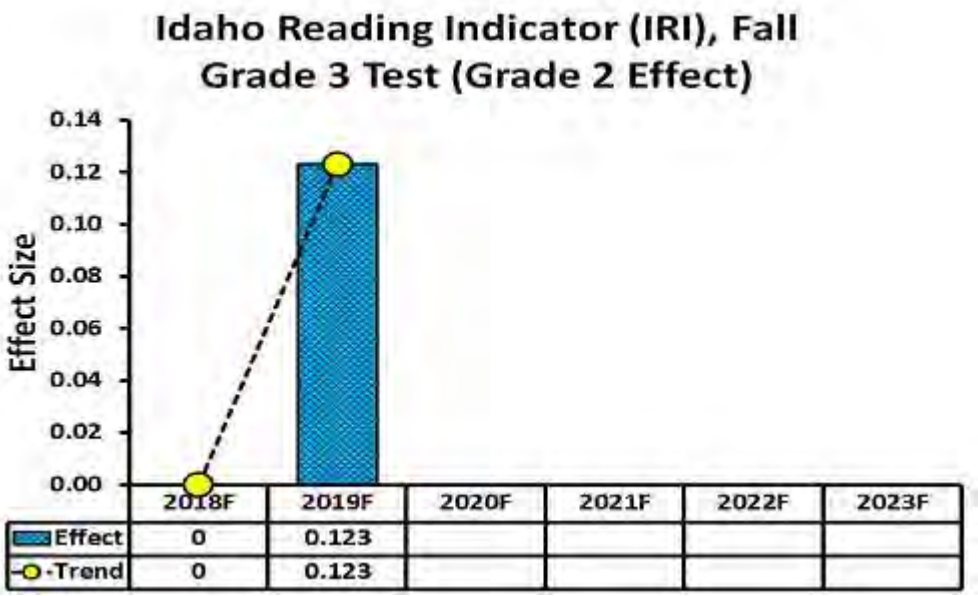


- ✓ The largest effect size for reading on the IRI fall 2018 and fall 2019 tests was on the Grade 1 test (+0.152), the magnitude of the growth for Idaho kindergarten students between the two fall tests.
- ✓ The second largest effect size for reading on the IRI fall 2018 and fall 2019 tests was on the Grade 3 test (+0.123), the magnitude of the growth for Idaho Grade 2 students between the two fall tests.
- ✓ The effect size for reading on the IRI fall 2018 and fall 2019 Grade 2 tests was (+0.100), which magnitude of the growth for Idaho Grade 1 students was between that of kindergarten and second grade students.



The calculation formula for the simple effect size used here is shown in the figure to the left, where the "x bar" is an average score and the "s" is a standard deviation. The effect size has the characteristics of a z-score.

As fall to fall IRI reading effect size results for grade K-2 are announced each year, they will be posted on a graph like the one below.



It would be nice if all effect sizes were positive every year, but that's not likely. It should be hoped, maybe even expected, that all of their trend lines live in positive territory even if the effect size on a given years is negative.

# # #

NCLB required states to report PAC statistics, but that requirement was not continued into ESSA. The ESSA accountability plans for five states, unlike Idaho's plan, use their state scale scores (with better properties) to report achievement and growth trends for mathematics and reading.

Ho (2007) concluded as useful as PAC (percent above a cut score) statistics have been in communicating test results to the public, their properties as trend statistics render them ill-suited for trend comparison. It seems perfectly reasonable to maintain PAC statistics as a primary means of NCLB reporting while conducting more serious trend analyses using statistics with better properties. Averages and average-based statistics should be the default consideration.

Ho, A.D. (2007). Discrepancies between score trends from NAEP and state tests: A scale-invariant perspective. *Educational Measurement: Issues and Practice*, 26(4), pp. 11-20.