

Standard-setting for Use of the NCSEAM Measures to Address the SPP/APR Parent/Family Indicators

Rationale

Rigorous measurement instruments yield consistent measures reportable in a uniform metric. This fact allows the meaning of the measures to be interpreted similarly by all users. However, the question of whether a particular measure (score) obtained through application of the measurement tool is adequate for a particular purpose should be determined by those who hold a stake in the consequences of using the measurement system.

There are many examples of standard-setting using well-known measurement tools. For example, colleges often set a particular SAT score as a minimum requirement for admission. States establish scores on their statewide public school tests that represent different levels of proficiency.

Use of the NCSEAM instruments to address the parent/family indicators requires the determination of a standard. For Part B, the standard is defined as the measure at which there is adequate evidence of schools' facilitation of parent involvement. For Part C, the standard is defined as the measure at which there is adequate evidence of families' achievement of specified outcomes.

In July 2005, NCSEAM convened a national group of stakeholders including parents, state Part B and Part C directors, advocates, service providers, and researchers, to recommend standards for the Part B and Part C indicators. Their recommendations are reported in the NCSEAM Summer Institute Plenary Session presentation [provide link].

Procedure

The standard setting process implemented by NCSEAM was a modification of the process described in Stone, G. E. (2001). Objective Standard Setting (or Truth in Advertising), *Journal of Applied Measurement*, 2(2), 187-201.

- Convene a workgroup with broad representation of families, state and local agencies, advocates, and other key stakeholders.
- Distribute a list of all the items constituting the scale for which a standard is to be set. The items should be in their calibration order from lowest (greatest amount of agreement) to highest (lowest amount of agreement). The items will have been scaled such that the item calibrations represent a combined .95 likelihood of a response across the three agree categories (agree, strongly agree, very strongly agree).
- Reach consensus as to the highest item with which participants would require an “agree” response in order to have confidence that the meaning of the indicator (e.g., schools are facilitating parent involvement) is being achieved. Descriptively, “If families don’t agree with this item” – and, by implication, with all those below it – “then we could not say that we had acceptable quality in this area.”

- The measure that corresponds to the selected item - or items, when several items are in the same statistical range - represents the standard.
- Performance on the indicator is calculated as the percent of parents or families with measures at or above the established standard.
- To take measurement error into consideration, construct a confidence interval around the percent based on the estimate of measurement error. We will then have 95% confidence that the true percent of parents at or above the measure is within this % interval.