



# Executive Summary

---

## Technical Report: The Factorial Validity and Reliability of the 2010 AdvancED Survey Series

Prepared for AdvancED by: Scott Weaver, Ph.D. and John Barile, Ph.D.

12/1/2011

# Executive Summary

---

In 2010, *AdvancED* re-developed its Opinion Inventory Series in accord with a redesign of its Learning Outcomes. The Opinion Inventory Series elicits teacher/staff, parent, and student level of agreement with statements that correspond to 5 new topic areas (subscales) linked to the learning outcomes: (1) Purpose and Direction, (2) Governance and Leadership, (3) Teaching and Assessing for Learning, (4) Resources and Support Systems, and (5) Using Results for Continuous Improvement. The Elementary Student Opinion Survey consists of 20 questions written to elicit the perceptions that 3<sup>rd</sup> through 5<sup>th</sup> grade students had of their school's performance along the aforementioned topic areas. The Middle and High School Student Opinion Survey measures the perceptions of students in grades 5 through 13+ with 31 items covering the same topic areas. The opinions of teachers and staff (Staff Opinion Survey) and of parents (Parent Opinion Survey) were similarly gathered with 53 and 35 items surveys, respectively. As the surveys are designed to measure school-wide perceptions and opinions, survey responses are aggregated over items and persons within the same school to provide measurements that reflect the collective perceptions a respondent group (e.g., elementary students) holds for any particular school.

Following the re-development of the Opinion Inventory Series, *AdvancED* conducted a study to evaluate the reliability and validity (factorial) of the redesigned Opinion Surveys. Sampling and data collection for this pilot study was conducted by *AdvancED* during March-May 2011, and the statistical analyses were conducted by independent consultants, Drs. Scott Weaver and John Barile, with expertise in survey methodology and psychometrics. Participating were 19,982 elementary students, 25,427 middle- and high-school students, 12,516 parents of students in selected schools, and 8,323 staff and teachers from 285 schools that are accredited by or seeking accreditation from *AdvancED*. These data were subjected to multilevel factor analysis aimed at determined the number and nature of factors that are measured by the surveys, to evaluate the validity of each item, and to estimate the reliability for the scale scores. Correlations between school characteristics variables (e.g., accreditation status, enrollment size) and opinion survey results were also examined.

Based on the results of these analyses, each opinion survey reliably captures a single dimension or factor reflecting the collective student, parent, and staff perceptions of the activities and best-practices schools engage in to promote student learning and achievement. All items were supported as valid and reliable, and reliability estimates of a composite score of a school performance on the survey computed by averaging over items and respondents within the same school indicated exceptionally high reliability (i.e., very minimal measurement error). In fact, the attained degree of reliability affords the scale

developers an opportunity to shorten the scales (e.g., to ease respondent burden) while still maintaining acceptable or excellent reliability. Whereas the results indicate a single factor rather than five factors corresponding to the five topic areas, this is not to suggest that schools should not consider computing and interpreting subscale scores. Examination of subscale scores might be useful, for instance, if a school is evaluating a concerted effort to improve its performance in a particular domain, such as providing high quality resources and support systems to its students. However, in general, these subscale scores would be expected to be highly correlated with one another and thus convey little unique information, and at the cost of lower reliability (as scales with fewer items tend to be less reliable than scales with more items) relative to the composite scale score. Further analyses would be necessary to ascertain whether subscale scores were *uniquely* predictive of student outcomes and other important variables.

With the exception of the student opinion surveys, the surveys exhibited criterion validity in that they were able to statistically discriminate accredited schools from those schools with "accreditation on advisement" status: the latter were rated 0.57 to 0.63 standard deviations lower than those with "accreditation" status. Private schools and schools with larger enrollments were rated significantly lower by teacher/staff and middle- and high-school students. However, additional analyses found that respondents from private schools tended to disagree with items more often than respondents from public schools, even when adjusting for overall differences. This may indicate that the metric of the composite score may differ between private and public scores, and thus comparisons between private and public schools should be made with caution.

In summary, the results of this study support the validity and reliability of the Opinion Inventory Series. Whereas the focus of these analyses was on the internal structure (factorial validity) of the scales, that two of the four surveys were associated with accreditation status also supports the predictive validity and meaningfulness of the surveys. The detailed technical report that follows elaborates on the sample, methodology, and analyses of the opinion surveys. This technical report is organized into four sections, each with parallel subsections, pertaining to the four respondent surveys.