

3. Report Levels and Audiences

A-Plus Communications. (1999). *Reporting results: What the public wants to know*. A companion piece to 1999 issue of Education Week's "Quality Counts." Arlington, VA: Author.

Beaton, A. E. (1992). *Methodological issues in reporting NAEP results at district and school levels*. Paper commissioned by the National Assessment Governing Board.

Berends, M., & Koretz, D. M. (1995). Reporting minority students' test scores: How well can the National Assessment of Educational Progress account for differences in social context? *Educational Assessment*, 3(3), 249-285.

This article investigates the adequacy of the National Assessment of Educational Progress (NAEP) for taking into account dissimilarities in students' family, school, and community contexts when reporting test score differences among population groups (i.e., racial and ethnic minorities). This question was addressed by comparing the NAEP to other representative data for Grades 8 and 12--the National Education Longitudinal Study (NELS) and High School and Beyond (HSB)--that contain richer social context measures. Our analyses show that NAEP lacks a number of important social context measures and that the quality of some (but by no means all) of NAEP's measures is low because of reliance on student self-reports and other unreliable data sources. These weaknesses of NAEP have important practical implications: Compared to HSB and NELLS, NAEP usually overestimates the achievement differences between students who come from different population groups but similar social contexts. However, at the secondary school level at which these analyses were conducted, these overestimates primarily reflect NAEP's lack of important measures rather than its reliance on student self-reports. [Authors' abstract]

Breithaupt, K., & Chuah, D. (2009, April). *Performance reporting for a licensing exam: What can, and should, we tell test takers?* Paper presented at the meeting of the National Council on Measurement in Education, San Diego, CA.

Bunch, M. B. (1986, April). *Building a user-oriented statewide score reporting system*. Paper presented at the meeting of the National Council on Measurement in Education, San Francisco, CA.

Burstein, L. (1990). Looking behind the "average": How are states reporting test results? *Educational Measurement: Issues and Practice*, 9(3), 23-26.

Means of interpreting norm-referenced tests to lead to more accurate reporting results are discussed, with particular emphasis on state-level and district-level data. Suggestions fall into the categories of documentation, frequency norm, and multiple form use. [Author's abstract]

In 1983 the Maryland State Department of Public Education (MSDE) issued a request for proposals for "The Development of the Score Reporting System for the Maryland Functional Testing Program." The MSDE called for a literature review, a national survey, a statewide survey of user needs and capabilities, an assessment of the state's report producing capability, and a final design for reports and a user's manual. Following a literature search, national and statewide surveys of reporting practices and information needs were conducted by Measurement Incorporated. Common and unique needs of district and building administrators, teachers and counselors, and parents and students were found. Using the nationwide search results, the information needs of students, parents, teachers, guidance counselors, principals, and district administrators in Maryland were surveyed. Score report design was based upon these studies emphasizing the accountability function of the tests. Four levels of reporting and seven content areas necessitated 28 separate score reports. Examples of four levels of reports (student, class, school, and local education agency) are presented. Each report is oriented to a specific audience, visual clutter is reduced, and diagnostic information is briefly presented. A user's guide provides thorough background on score interpretation at multiple levels. This score reporting system appears to meet the responsibilities and information needs of all its audiences. [Author's abstract]

Cieslak, P. (2000, February). *Milwaukee's experience with district-level NAEP results*. Paper presented at the workshop of the Committee on NAEP Reporting Practices: Investigating District-Level and Market-Based Reporting, National Research Council, Washington, DC.

DeVito, P. J., & Koenig, J. A. (Eds.). (1999). *Reporting district-level NAEP data: Summary of a workshop*. Washington, DC: National Academy Press. Retrieved March 31, 2009, from http://www.nap.edu/catalog.php?record_id=9768

DeVito, P. J., & Koenig, J. A. (Eds.). (2001). *NAEP reporting practices: Investigating district-level and market-basket reporting*. Washington, DC: National Academy Press. Retrieved March 31, 2009, from http://www.nap.edu/catalog.php?record_id=10049

Study questions focused on the: characteristics and features of the reporting methods, information needs likely to be served, level of interest in the reporting practices, types of inferences that could be based on the reported data, implications of the reporting methods for NAEP, and implications of the reporting methods for state and local education programs. [Authors' abstract]

Hambleton, R. K. (2002). How can we make NAEP and state test score reporting scale and reports more understandable? In R. W. Lissitz & W. D. Schafer (Eds.), *Assessment in educational reform* (pp. 192-205). Boston, MA: Allyn & Bacon.

Hambleton, R. K. (2002, February). *A new challenge: Making results from large scale assessments understandable and useful*. An invited presentation at the Provincial

Testing in Canadian Schools: Research, Policy, and Practice Conference, Victoria, British Columbia.

Hambleton, R. K., & Meara, K. (2000). Newspaper coverage of NAEP results, 1990 to 1999. In National Assessment Governing Board (Ed.), *Student performance standards of the National Assessment of Educational Progress: Affirmation and improvements*. Washington, DC: Editor.

Hambleton, R. K., & Slater, S. (1997). *Are NAEP executive summary reports understandable to policy makers and educators?* (CSE Technical Report 430). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Teaching. Retrieved March 31, 2009, from <http://research.cse.ucla.edu/Reports/TECH430.pdf>

This research study is a follow-up to several recent studies conducted on NAEP reports that found policy makers and the media were misinterpreting test, figures, and tables. Our purposes were (a) to investigate the extent to which NAEP Executive Summary Reports are understandable to policy makers and educators, and (b) to the extent that problems are identified. Several recommendations are offered for improving the NAEP reports: First, all displays of data should be field tested prior to their use in NAEP Executive Summary Reports. A second recommendation is that NAEP reports for policy makers and educators should be considerably simplified. A third recommendation is that NAEP reports tailored to particular audiences may be needed to improve clarity, understandability, and usefulness. [Authors' abstract]

Hambleton, R. K., & Smith, T. (1999). *A focus group study of the general/public 1996 NAEP Science Reports* (Laboratory of Psychometric and Evaluative Research Report No. 361). Amherst, MA: University of Massachusetts, School of Education.

Haney, W., & Madaus, G. F. (1991). Caution on the future of NAEP: Arguments against using NAEP tests and data reporting below the state level. In R. Glaser, R. Linn, & G. Bohrnstedt (Eds.), *Assessing student achievement in the states: Background studies*. Stanford, CA: National Academy of Education.

Impara, J. C., Divine, K. P., Bruce, F. A., Liverman, M. R., & Gay, A. (1991). Teachers' ability to interpret standardized test scores. *Educational Measurement: Issues and Practice*, 10(4), 16-18.

To what extent do teachers possess the competence to interpret state testing program results properly? [Authors' abstract]

Jaeger, R. M. (1996). *Reporting large scale assessment results for public consumption: Some propositions and palliatives*. Paper presented at the meeting of the National Council on Measurement in Education, New York, NY.

Johnson, E. G. (1994). *Standard errors for below-state reporting of National Assessment of Educational Progress*. Paper prepared for the National Assessment Governing Board. Princeton, NJ: Educational Testing Service.

Koretz, D. M. (1991). State comparisons using NAEP: Large costs, disappointing benefits. *Educational Researcher*, 20(3), 19-21.

Suggests that the proposed state-by-state National Assessment of Educational Progress (NAEP) will be unable to provide information about which state programs are responsible for differences in test scores. Raises concerns about its cost effectiveness and potential loss of validity if used in state comparisons. [Author's abstract]

Koretz, D., & Diebert, E. (1993). *Interpretations of National Assessment of Educational Progress (NAEP) anchor points and achievement levels by the print media in 1991*. Santa Monica, CA: RAND.

Levine, R., Rathbun, A., Selden, R., & Davis, A. (1998). *NAEP's constituents: What do they want? Report of the National Assessment of Educational Progress Constituents Survey and Focus Groups* (NCES 98-521). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

McDonnell, L. M. (1994). *Policymakers' views of student assessment*. Report commissioned by the Office of Educational Research and Improvement, U.S. Department of Education. Santa Monica, CA: RAND Institute on Education and Training.

O'Reilly, J. (2000, February). *District level and market-basket reporting: A district perspective*. Paper presented at the workshop of the Committee on NAEP Reporting Practices: Investigating District-Level and Market-Based Reporting, National Research Council, Washington, DC.

Patelis, T., & Matos, H. (2009, April). *Efforts to produce relevant score reports to school, district, and state officials on national tests*. Paper presented at the meeting of the National Council on Measurement in Education, San Diego, CA.

A historical overview of score reporting at the College Board is documented within this paper. Efforts to make score reports more meaningful and valuable to score reports users are described through the developmental activities that were underway during the production of the College Board's SAT Skills Insight reports for both students and state officials. Reflections of lessons learned throughout the report development process are also provided, along with the College Board's vision for future score reports. [Authors' abstract]

Robert, E. D. (1994, February). *Guidelines for the use of NAEP at the district and school levels*. Paper commissioned by the National Assessment Governing Board.

Rust, K. (1999). *NAEP sample designs and district level reporting*. Paper prepared for the National Research Council Workshop on District-Level Reporting, Washington, DC.

Selden, R. (1991). The case for district- and school-level results from NAEP. In R. Glaser, R. Linn, & G. Bohrnstedt (Eds.), *Assessing student achievement in the states: Background studies*. Stanford, CA: National Academy of Education.

Sicoli, F. (2002). What do school-level scores from large-scale assessments really measure? *Educational Measurement: Issues and Practice*, 21(4), 17-26.

Although assessments of mathematics, reading, and writing are assumed to measure distinct academic skills, this may be difficult owing to the pervasive influence of general ability on performance. Factor analyses of school-level data from 14 large-scale assessment programs revealed that 80% of the variance in mathematics, reading, and writing scores was due to a common, underlying factor. Multiple regression analyses confirmed that scores contribute little information that is unique to a particular subject (6% or less). Although different assessments may create the illusion of providing unique information, they may be tapping into generic cognitive abilities that cut across content areas. These results raise suspicions about the value and validity of interpretations based on school-level subject area scores. [Author's abstract]

Simmons, C., & Mwalimu, M. (2000). What NAEP's publics have to say. In M. L. Bourque & S. Byrd (Eds.), *Student performance standards on the National Assessment of Educational Progress: Affirmation and improvements. A study initiated to examine a decade of achievement level setting on NAEP* (pp. 184-219). Washington, DC: National Assessment Governing Board.

Sopko, K., & Reder, N. (2007). *Public and parent reporting requirements: NCLB and IDEA regulations*. In Forum. Alexandria, VA: National Association of State Directors of Special Education.

Trout, D. L., & Hyde, B. (2006, April). *Developing score reports for statewide assessments that are valued and used: Feedback from K-12 stakeholders*. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Westin, T. (1999). *Reporting issues and strategies for disabled students in large scale assessments*. Washington, DC: Assessing Special Education Students, SCASS, CCSSO.

Ysseldyke, J., & Bielinski, J. (2002). Effect of different methods of reporting and reclassification on trends in test scores for students with disabilities. *Exceptional Children*, 68(2), 189-200.

State education agencies are now required to report on the educational performance and progress of all students, including students with disabilities. States are beginning to report trends, and to compare trends in performance of students with and without disabilities. We compare the effects of different methods of analyzing trends to illustrate how failure to account for changes in classification of students will lead to misinterpretation of data on the performance and progress of students with disabilities, and inappropriate policy decisions. We compare three ways of looking at trends over time, and use data from 5 years of assessment in a large state to illustrate the effects of students who change classification. We discuss how accounting for changes in classification of individual students will lead to more appropriate decisions and help avoid negative consequences for students with disabilities. [Authors' abstract]