

NAEP: An Invaluable Resource in Need of Improvement

By John F. (Jack) Jennings

In American society, there are few issues more discussed than those involving education. The National Assessment of Educational Progress (NAEP) is an invaluable resource for making those discussions well informed and grounded on solid information.

For that reason, our deliberations today, on how to ensure the integrity of NAEP--and how to improve it--are crucial. It would be good for the country if these meetings resulted in broad, bipartisan agreement on how to make NAEP better, since the law authorizing the assessment will come under review in 1999 by the administration and Congress.

Although NAEP is invaluable, it needs improvement in three major areas--setting the student achievement levels in the main assessment, more frequently administering the long-term assessment than is currently planned, and clarifying the administration of the entire program.

Setting the Student Achievement Levels

The performance levels, which are the standards for pupil achievement that are based on score thresholds established by the National Assessment Governing Board (NAGB) for the use in the main assessment, must be beyond reproach. Those achievement levels increasingly are being used by policymakers, the news media, and others to measure the state of American education, but controversy continues as to how accurately the levels gauge student academic performance.

When these achievement levels first were established in the early 1990s, they were criticized by testing experts and the General Accounting Office for being too difficult and yielding results different from those of other large-scale assessments. In 1996, a NAEP technical review panel looked at the levels in use at that time and concluded: "The achievement levels (unlike the anchor points) reflect judgments about how students should perform, and different panels of judges (or different methods of setting the levels) would likely have produced different standards.

In September 1998, a panel of experts convened by the National Research Council (NRC) of the National Academy of Sciences concluded that the process used to determine these achievement levels was "fundamentally flawed" and should be overhauled. As reported by *Education Week*, the panel said that the performance levels often are too rigorous, are the product of a process riddled by inconsistencies, and do not reflect the results of similar large-scale assessments. Echoing the 1996 technical panel, NRC stated that NAGB must clarify that the standards are based on judgments by experts and that they do not "reflect some deeper scientific truth."

NAGB first must be straightforward and acknowledge that these performance levels as currently constituted have major limitations. Further, the Board must do a better job of accurately describing what the levels really mean. Then, the process for establishing these levels must fundamentally be rethought. NRC recommends a new method by which the standards and the test questions would be written simultaneously, and by which the results in field tests would be compared to the results of other large-scale assessments. The NRC panel also recommends creating a broad array of education indicators that provide a complete picture of what and how well children are learning, rather than relying on single test scores from one administration of NAEP.

More Frequently Administering the Long-Term Assessment

The long-term assessment should continue to be administered every two years. NAGB has adopted a plan under which this assessment would be administered less frequently--every four years beginning in 1999--while increasing the frequency of the main assessment.

The long-term assessment measures trends in mathematics, science, reading, and writing achievement, and has done so since 1969. This assessment is indispensable in understanding how well American students have done over time in these basic subject areas. If the plan to administer this assessment less often is implemented, long-term trend

results will be available less frequently and therefore will be used less often by policymakers, the news media, and others in judging the state of American education.

Concomitantly, by administering the other test--the main assessment--more often than is currently being done, the results of that assessment will be used more frequently by decisionmakers and the public. That main assessment has all the problems with the student performance levels that have just been described.

NAGB should revise its plan to downgrade the importance of the long-term trend assessment and continue to administer it every two years. The long-term trends are an essential counterweight to the "fundamentally flawed" student-performance levels used in the main assessment.

Clarifying the Administration

The roles of NAGB and the National Center for Education Statistics (NCES) must be clarified. The law as currently written has created a situation in which neither entity is entirely clear about its responsibilities in setting policy for NAEP and its administration.

In 1996, KPMG Peat Marwick, Inc., conducted a management analysis of NAEP and concluded that there is an undesirable level of confusion over the extent to which NAGB and NCES have authority over NAEP. In particular, these management experts said that efforts to avoid confusion and potential conflict through consensus frequently lead to delays in decisionmaking and to increased costs for conducting NAEP. Their recommendations were that NAGB general should limit its decisionmaking to the subjects on which pupils are to be tested, frequency of tests, and performance levels, and should simply provide advice to NCES on other issues.

The administration and Congress should adopt the recommendations of this management study and change the law to clarify the responsibilities of NAGB and NCES. The law also should be amended to require, in a few years, another management review, and likewise another study by NRC on how NAGB has improved the student-performance levels used in the main assessment.

Conclusion

NAEP is invaluable. It provides information to the country on how well students are progressing and highlights the problems that must be addressed. NAGB and NCES are to be commended for producing such an important resource. However, NAEP could benefit from major changes to improve its usefulness and reliability, and that is the task before us.