

Getting the Word Out About the National Assessment

The National Assessment Governing Board and Changes in NAEP Reporting

John H. Stevens

March 2009

Paper Commissioned for the 20th Anniversary of the
National Assessment Governing Board
1988–2008

John H. Stevens was a member of the National Assessment Governing Board from 1998 to 2006 and served as chair of the Reporting and Dissemination Committee for 6 years. He retired in 2007 after 15 years as Executive Director of the Texas Business and Education Coalition. Previously, he worked for the American Federation of Teachers for more than two decades.

DEDICATION



In memory of **Luis Ramos**, a business executive and civic leader from Shickshinny, Pennsylvania, who served on the Reporting and Dissemination Committee from the time of his appointment to the National Assessment Governing Board on October 1, 2004, until his tragic death in an automobile accident on January 1, 2008.¹

The National Assessment of Educational Progress (NAEP) was started in 1969 to provide information about one of the most important enterprises in the nation, the education of its young people. The assessment is conducted by the National Center for Education Statistics (NCES), an agency of the U.S. Department of Education.

In 1988, when NAEP expanded to include state samples, Congress established the National Assessment Governing Board as an independent entity to insulate policymaking for NAEP from political considerations and the influence of special interests.

From the beginning of their working relationship, NCES and the Governing Board knew that the integrity of NAEP depends on producing reliable achievement data based on a process that is technically correct. But they also recognized that the value of the National Assessment depends on getting word of its findings out to the American public as clearly and broadly as possible.

This paper describes the development of Board policy on the reporting, release, and dissemination of National Assessment results and other relevant information. Through selected scenarios, it also illustrates the kinds of issues that required action by the Board. These issues have included concerns about the testing and reporting of English language learners and students with disabilities, the challenges of assessing performance in Puerto Rico, reporting on private schools, and producing a special study of public charter schools.

Establishing Policy on Reporting and Dissemination

On November 17, 1990, at a meeting in Atlanta, Georgia, the Board adopted its first *Policy Statement on Reporting and Dissemination of National Assessment Results*. That statement established “guiding principles” that have been the foundation of those activities ever since. Based on those principles, the statement laid out requirements and procedures for (1) report preparation and content, (2) public release of NAEP results, and (3) dissemination. Those provisions included:

Report Preparation and Content

- Results shall be issued by the Commissioner of Education Statistics in accordance with NCES clearance and review procedures.
- Interpretation of results and conclusions as to cause and effect relationships shall be limited and must be strongly supported by NAEP survey results.
- Achievement levels adopted by the Board shall be the primary means of reporting.

Public Release of NAEP Results

- Public reporting and release procedures shall be apolitical.
- The release date and manner of release shall be determined by the Commissioner with the approval of the Board. Any initial public release shall be made only in accordance with a specific plan and timetable approved in advance by the Board or one of its duly constituted committees.

- At press conferences, the Commissioner of Education Statistics and/or designee shall present major data findings. The Board shall select members to comment on the results. Either the Commissioner or the Board may invite other officials or experts to comment according to the approved plan.
- Press conferences and press releases for NAEP reports shall not be used to release or distribute any reports that do not pertain to NAEP.

Dissemination

- Consult with education, business, and civic organizations to gather ideas for making NAEP results more useful.
- Work with education organizations to disseminate NAEP results to their members.
- Presentations and briefings on NAEP shall be given to civic groups, organizations, and policymakers.²

Although there was occasional tension, the Governing Board, NCES, and contractors generally worked together well to report and disseminate results and information about NAEP. From time to time, they contracted with outside experts to bring a different perspective to bear on their work.

The Widmeyer Group Report

In March 1994, The Widmeyer Group, a Washington, D.C. public relations firm hired as consultants to the Board, delivered a report on *Dissemination Strategies for the National Assessment of Educational Progress*. The research found that NAEP was not reaching its intended audiences, partly because of presentation barriers. It said that:

- NAEP data are confusing to readers. Many believe the data are written for a highly technical audience.
- Users are intimidated and overwhelmed by the sheer volume of NAEP reports, which were said to be too comprehensive and poorly organized.
- Survey respondents wanted graphics, narrative descriptions, and an executive summary.³

The typical “comprehensive report” used to release NAEP results was not user friendly. For example, the 1992 Reading report was 310 pages long. It included a lengthy table of contents and technical information about several aspects of NAEP before discussing the results. Information was presented mostly in data tables with technically worded footnotes,⁴ and was interspersed with long gray summarizing paragraphs.

Another major problem for users was relevance. According to the Widmeyer study, the public found it difficult to connect the assessment to their lives. They said it was not localized and not interpretive or prescriptive.⁵

Educators said they would like to see information that would help them make changes in materials, curriculum design, and instructional delivery to improve student achievement.⁶

The business community found the data largely irrelevant because they did not measure qualities important to them such as teamwork, maturity, dedication to a task, and the desire to learn.⁷

However, much of the “relevant information” desired by these audiences cannot be provided by NAEP because of the nature of the assessment itself and the kinds of data collected. By law and by design, NAEP does not report scores for individual students or schools. It is a snapshot assessment with no longitudinal data. Thus, it cannot show the relative effectiveness of various educational practices, although it does offer a sound overall measure of academic achievement.

The Widmeyer report laid out an ambitious outreach plan. At least one of its suggestions was visionary in the early days of the Internet: “Hook into the new national information highway and make data and results available on-line.”⁸

Although many of the Widmeyer ideas were not possible to implement, the report provided a sounding board that reinforced many actions already taken and stimulated discussion about future activity to strengthen the ongoing informational activities of the Board and NCES.

The Board and NCES – A Settled Working Relationship

For nearly a decade, the reporting and dissemination activities of NAEP proceeded as envisioned in the initial policy statement adopted by the Board. More than two dozen National Assessment reports were released according to plans developed by NCES and approved by the Board. The working relationship between the two agencies remained stable through changes in the Administration, the composition and leadership of the Governing Board, and in the leadership and staff at NCES.

Over the years, one of the Board’s ongoing frustrations was the protracted time between testing and the release of reports (typically 12 months or longer). To many users, particularly policymakers, the achievement results seemed dated as soon as they were released. NCES agreed with the concern, but explained carefully that it was very difficult to shorten the timeline.

NAEP is a large, complex, highly technical and time-consuming enterprise involving many people doing different things according to a tight schedule, and most of the time it works as planned. Occasionally, however, the flow of work has been disrupted by unforeseen occurrences, some of which take extra time. Now and then a technical glitch has to be fixed before results can be released. Sometimes matters have to be dealt with to satisfy requirements stated in federal law. Suggestions made by respected individuals are taken seriously and given careful consideration. And, once in a while, issues demand attention because they challenge the integrity of NAEP itself.

The Governing Board and NCES dealt with known challenges and unforeseen developments effectively and without attracting much outside attention.

1998 Reading Release Had Far-Reaching Consequences

Although there was every reason to believe that the release of the 1998 Reading results would be routine, it proved to be anything but that.

As provided in policy, NCES Commissioner Pascal Forgione presented a release plan to the Reporting and Dissemination (R&D) Committee at the November Board meeting. The 1998 NAEP Reading results would be released at a news conference in Washington, D.C., in early February. Commissioner Forgione would present the main findings followed by comments from Board Chairman Mark Musick.⁹

In late January, NCES detected a technical error related to the results for the states. In a teleconference with the R&D Committee, the decision was made to release the national results in early February as planned and delay the release of state results until March, giving time for the errors to be corrected.

The NAEP Reading Release News Conference

However, the actual format of the news conference was changed shortly before the release date by the Secretary of Education, Richard Riley, who by law has final authority over the release of all Education Department reports, including those by NAEP. The news conference on February 10, 1999, featured Vice President Al Gore and was attended by several hundred Education Department employees and representatives of education associations. Vice President Gore opened with a prepared statement: “Today I am very proud to report to you new evidence that our efforts are beginning to pay off. For the very first time, reading scores have improved for each of the three grades measured by the National Assessment of Educational Progress: 4th grade, 8th grade and 12th grade. This is great progress and we are proud to report it.”

“These results are encouraging and show that we are beginning to move in the right direction,” said the Vice President. “But we must work faster because there is much more room for improvement.”¹⁰ He commented on several Administration initiatives and encouraged Congress to enact the President’s education agenda, particularly those items related to reading. Secretary Riley made a short statement that was followed by remarks from a suburban Washington teacher, who also praised Administration policies.

After Vice President Gore and most of the audience left, the event continued with a data presentation by Commissioner Forgione. His explanation of the results differed from that of the Vice President:

“The 1998 results show some improvement in reading achievement nationally, particularly at grade 8,” said Forgione. “However, the increases between 1994 and 1998 for students in grades 4 and 12 showed *no net gain over the 1992 average scores.*”¹¹

The comments by Governing Board Chairman Mark Musick were in line with the Commissioner’s comments, indicating no progress from 1992 to 1998 at grades 4 and 12 and only slight gains at grade 8.

The Board's Response

The appearance by Vice President Gore was not the way the release had been planned. After consulting with the Executive Committee, Musick wrote a letter to Forgione stating his concern that the “format, tone and substance of the news conference was not consistent with the principle of an independent, non-partisan release of National Assessment data, an important and long-standing board policy.” He went on to state that the plan approved by the Board was not followed and that the arrangements put more focus on the Vice President and the Administration’s education policy than on the NAEP report. He commented that the presentation of the data by the Commissioner took place toward the end of the program, after the Vice President and much of the audience had left, and observed that the press was placed in a roped-off area at the back of the room while education association representatives and Education Department staff were at the front.¹²

At the March Governing Board meeting, the R&D Committee reviewed the release. Chairman Musick described what had occurred at the February 10 news conference and noted that it was not in agreement with Board policy. He reported that the March 4 release of state results did follow the approved plan. On that occasion, Commissioner Forgione’s official release of the results was followed by commentary from Board member Wilmer Cody and Secretary Riley.

The committee discussed a redraft of existing Board policies that had been prepared by Board staff. Commissioner Forgione and committee members made several suggestions and Acting Chairman Mitsugi Nakashima instructed Larry Feinberg, a member of the Board staff, to incorporate those changes into a draft for consideration at the May Board meeting.¹³

A Challenge to the Integrity of State Results

Vice President Gore was not the only source of controversy related to the 1998 NAEP Reading assessment. An airline pilot from Kentucky stirred up a hornet’s nest after the release of state-level results. Only 2 days after the March 4 news conference, Richard G. Innes, a former Air Force fighter pilot whom *Education Week* describes as “a persistent critic of the state’s education policies,” distributed an e-mail in which he questioned the gains in some states because they had excluded a higher percentage of students with disabilities or limited English proficiency in 1998 than they had in previous years.

Innes wrote, “After doing some careful checking, I have discovered that the NAEP excluded 10% of our (Kentucky’s) IEP/Student with Disability kids from the testing sample in 1998. In 1994, the last time the test was given, only 4% of IEP/SD was excluded!”¹⁴ Innes concluded that the gains reported for Kentucky and other states were realized because of the increased exclusion of special education students.

The stakes were raised on March 12, when the *Cincinnati Enquirer* featured an article with the headline, *Omitting Special Ed Kids May Have Aided Ky. Scores.*¹⁵

The Board's Response: Part II

In the opening general session of the May 13 Governing Board meeting, Commissioner Forgione discussed a report by the Educational Testing Service (ETS), a NAEP contractor, suggesting that comparisons of 1998 NAEP Reading results to those in 1994 and 1992 were probably affected by changes in the rate at which states excluded students from participating in the assessment.

In reviewing results for all 36 states with student samples for both 1994 and 1998, ETS assumed that students excluded from the 1998 assessment would have scored at about the same level as students who scored the worst in previous years. Based on that assumption, analysts found that the gains reported for Kentucky and Maryland would not have been statistically significant.

Understandably, Board member Wilmer Cody, Kentucky Commissioner of Education, was upset that his state's gains were being questioned. "We would challenge that assumption... that the students who were excluded were the lowest-performing students," said Cody.¹⁶

After the general session recessed, the R&D Committee reviewed the draft policy statement and comments sent by NCES. It decided that the official press release about new NAEP reports would be issued by NCES rather than the Department of Education to strengthen its independence and separation from political considerations. The committee retained the goal of reporting NAEP results in either 6 or 9 months even though NCES said those timelines were unrealistic.

On May 15, the Board unanimously adopted the R&D Committee's recommended policy, making it very clear that the type of statement made by Vice President Gore was not an acceptable way to release NAEP results: "The official press statement or release announcing NAEP results shall be issued by the NCES and shall be a straight-forward presentation of significant data without political commentary or program advocacy." Another key change was that a new, brief, popular-style summary report would be the primary way to release NAEP results in initial public releases.¹⁷

Addressing the Exclusion Issue

Peggy Carr of NCES reviewed recommendations from NAEP contractors ETS and Westat on reporting the exclusion of students with disabilities and English language learners, and on testing with accommodations.

The two NAEP contractors warned that "One of the most basic issues that must be dealt with is any supposition among users of NAEP data that there is systematic wrongdoing in the conduct of the assessment. We must be prepared to deal with claims that states or jurisdictions are purposefully excluding students from NAEP to improve scores." The memo included recommendations for the Board's consideration, suggesting that state results should be flagged if they fail to meet certain standards and that policies should be established regarding results that include data from assessments administered to students using accommodations.¹⁸

Following the Board meeting, the national media resumed its coverage of the exclusion issue:

- The *Associated Press* reported on May 14 that the chief of the Education Department's statistical branch, Pascal D. Forgione, Jr., said that neither the states' test scores nor their rankings would be affected by the discovery of the exclusions, but that "the federal study of the 1998 test results could lead test-givers to change some policies."¹⁹
- On May 17, *USA Today* reported that the Department of Education is planning "more investigations into the accuracy of recent national reading test scores."²⁰
- On May 18, the *Christian Science Monitor* made the stakes very clear in an article that directly questioned NAEP's credibility: "Is there a reliable way to find out how well American kids are reading? Until now, the National Assessment of Education Progress (NAEP) seemed the best route. It's as close as the U.S. has come to a national report card. But a Kentucky parent has raised an issue that the national testing establishment is having a tough time answering: Can a state's scores be accurate when they don't include large numbers of low-scoring students?"²¹

The Congressional Hearing

The issue of the Vice President's role in the February news conference had come to the attention of Congressman William Goodling (R), Chairman of the House Committee on Education and the Workforce. He was not satisfied with Secretary Riley's response to his inquiry into the matter, and Commissioner Forgione and Chairman Musick were summoned to appear at a hearing of the Subcommittee on Oversight and Investigations.

Subcommittee Chairman Congressman Peter Hoekstra, (R) Michigan, called the hearing to order and announced, "We are here today to address the potential politicization of the National Assessment of Educational Progress (NAEP) and the trustworthiness of the scores that the states received during the last reading assessment."²²

In his opening statement, Hoekstra detailed the Vice President's role in the February 10 news conference, observing that he was the featured speaker and that it was the first time any official higher than the Secretary of Education had participated in a NAEP release: "The Vice President's participation was not an apolitical release of educational data, but rather an orchestrated media event to promote a political agenda." He cited e-mails from the Office of the Vice President that described the event as "a VP education event" that was "good for the VP, for the Department and for education" and referred to suggestions from the Secretary's office about the program's dimension, stating, for example, "It is a reading day, why not push for reading dollars?"²³

Hoekstra brought up the other issue before the committee. "In addition to questions about the release of the scores, I have questions about how well the new NAEP inclusion policies are working. Why did the exclusion of students with disabilities and limited English proficiency increase since the 1994 test? Was NAGB pressured into making decisions that have brought us to where we are today, questioning the way in which the results were released and the possibility that the scores themselves might have been clouded by politics?"²⁴

Chairman Hoekstra ended his statement and recognized Representative Tim Roemer, (D) Indiana, the ranking minority member of the subcommittee. Representative Roemer opened his remarks by agreeing with the Chairman's statement "that politics should not influence this process." He added, "...whether it be a President or a Vice President." He called attention to a letter written on June 3, 1992, by Richard Boyd, then Chairman of the Governing Board, to Emerson Elliott, Acting Commissioner of NCES.

Mr. Boyd had written, "I was surprised and disturbed to learn that on the day before the scheduled press conference the main results in the report were released prematurely by President Bush. According to press reports, the President described the NAEP findings and discussed them during two appearances in Georgia: at a private school and at a Republican fund-raising dinner. A summary of the results was distributed by the White House."²⁵

Representative Roemer continued, "Now, we can have a hearing on how President Bush released these prematurely and speak about that. Quite frankly, I am not interested in President Bush talking about these NAEP scores at a private fund-raiser or school the day before." He concluded his remarks by again echoing the Chairman's earlier comments, "...it is in everyone's interest to have clear boundaries. The situation of ambiguity in releasing reports invariably leads to awkward situations that do not serve the Nation well."²⁶

After the opening remarks, Commissioner Forgione and Chairman Musick made their statements and responded to questions. Forgione explained the sequence of events leading up to the news conference:

- In November the Governing Board approved the Commissioner's proposed plan for the release of the 1998 NAEP Reading results.
- In early January, after briefing Education Department staff about the release, Forgione and Roy Truby, Executive Director of the Governing Board, were informed that the White House was interested in participating.
- During the week of February 1, the Commissioner received a request from the Secretary's Office to add a practicing teacher to the agenda. After he consulted with Truby, the proposal was accepted.
- Late in the week prior to the release, Forgione was alerted by the Department that the Vice President wished to be involved in the release.
- On February 8, only 2 days before the event, Forgione participated in a teleconference with staff from the Vice President's office, the Secretary's office, and the Office of Public Affairs of the Department of Education. He recommended that the event should proceed as originally planned, with the Commissioner opening and releasing the results followed by Chairman Musick, who would speak for the Board. Then the Vice President, the Secretary, and others (possibly the teacher) could comment on the results. When Forgione left the meeting, he understood that this would take place and he called Roy Truby to inform him that the original plan was still in effect.
- Later that day, Forgione learned from a member of his staff that the Department had changed the agenda. There was a delay until 11 a.m., when the Secretary and Vice President would open the event. Since the announcement had been made about the

In his testimony, Musick observed that a release of the 1998 state Reading results in March was conducted in accord with the NCES plan approved by the Board; he then made a point that went to the heart of the matter, “The Commissioner implements Board policies, but he is an officer of the Education Department, appointed by the President and subject to confirmation by the Senate.”²⁸ The obvious implication of his statement was that the Commissioner can be and, in fact, had been overruled by higher-ups in the Department of Education and the Executive Branch. Musick concluded his remarks by reporting that the Board had strengthened the Policy on Reporting and Dissemination at its May meeting.

Although not a member of the subcommittee, Representative Mike Castle, (R) Delaware, attended the hearing because of his interest in NAEP. He recalled that the relationship between the Governing Board and NCES was “a little bit of a troubled area when I was on the NAGB Board.” His questions had to do with the independence of the agencies and the interrelationships between them.²⁹

Musick and Forgione both said the current relationship was very good because the leaders worked well together, but that the institutional structures could be improved. Musick observed that the Board had very little real power because its authority was limited to deciding on the appropriateness of individual test items to be used in the Assessment. Representative Castle asked them to submit their thoughts about those matters to him in writing.³⁰

The discussion shifted to the exclusion issue; Representative Castle stated his concern that “on this round of tests there were some serious questions raised, at least in certain states, about students who might have achieved at lower levels. Learning disabled programs were not given the test in the same percentage as they might have been in other states or other years. That really bothered me.” Forgione replied that “the study we just completed finds that this had modest impact on the results.” Castle replied, “You may argue it is not significant, but because of what you just went through there are differences in who is taking the test and who isn’t.”³¹

Responding to a question from Representative Ernie Fletcher, (R-KY), Musick explained that the NAEP exclusion policy was changed so that the national assessment would be consistent with the efforts of states to include more students in their testing programs. He said the challenge was to do this while maintaining our trends and ties to what went on before.³²

Congressman Ron Kind, (D) Wisconsin, stated, “What concerns me is that we have got what appears to be some manipulation of NAEP test results going on right now. I would like to know what is being done so that we don’t have a state-by-state manipulation of future test scores and we can get some accurate and quantifiable data that are reliable in the future.”

Forgione replied, “The answer to the question is there was no manipulation. We administered the tests independently. We picked the schools in every state. We picked the roster of fourth graders and we drew 30 names of those students to be tested. Only at that point could

the principal or administrator say that certain students should not be tested because their IEP (Individual Education Plan) says so.³³

At the end of the hearing, Chairman Hoekstra thanked the witnesses and made a promise. “As we go through the reauthorization of these programs next year, you can be sure that we will give you greater independence to do the job as you best see fit. That is the commitment I make to you.”³⁴

Aftermath

Forgione resigned following a White House decision not to nominate him for a second term as NCES Commissioner. He left Washington, D.C., to become Superintendent of the Austin (Texas) School District and, because of his continued belief in the value of NAEP, secured his district’s participation in the Trial Urban District Assessment (TUDA) beginning in 2005.

Summer 1999

The exclusion issue continued to be a concern. A committee of the Council of Chief State School Officers (CCSSO) recommended that 1998 NAEP Writing reports include data about students assessed under nonstandard conditions, but that no data be flagged because of the number or percentage of students excluded from the assessments.³⁵

Steve Lazer from ETS sent a memo to Peggy Carr observing that there was a precedent for flagging the results for jurisdictions whose participation rates failed to meet certain standards. In the early 1990s, the Board adopted a policy of not reporting data for jurisdictions with school participation rates below 70 percent of the original sample. States that meet that standard but have other problems with school or student participation are flagged in NAEP reports. ETS offered thoughts to consider in developing similar policies related to exclusions:

- States might resist because of limited ability to influence inclusion decisions at the school level.
- Any flagging rules must avoid penalizing states with large English language learner populations.
- States must be informed before participating in an assessment in which new rules are applied.³⁶

The August 1999 Governing Board Meeting

Peggy Carr made a presentation about the issue to a joint meeting of the Design and Methodology, and R&D Committees. She reviewed previous discussions and reported that full data on exclusions and accommodations would be included in all 2000 NAEP reports. However, state testing officials expressed concern about introducing new rules regarding “cautions” and “flagging” at this time since states had already signed agreements to participate in NAEP 2000.

The committees agreed that it would be unwise to change the rules after states had agreed to participate and decided against presenting a resolution on the issue for the Board's consideration.³⁷

Fall 1999: The Second NCES Study

At the request of the Kentucky Department of Education, Laress Wise, of the Human Resources Research Organization (HumRRO), imputed the scores of Kentucky fourth graders who were excluded from the 1998 National Assessment, based on the results for similar students who had also taken the state's own reading test (which permitted the read-aloud accommodation). Wise concluded that Kentucky's gains were as large as NAEP had originally reported. His report stated, "Even in the worst-case scenario, the rise in Kentucky's scores from 1994 to 1998 was clearly significant."

However, that did not satisfy Innes, the airline pilot. He said "the basic, fundamental premise is absolutely out to lunch." Innes estimated that three-fourths of the Kentucky students with disabilities who were excluded from the National Assessment received such accommodations as having questions read to them when they took the state reading test. Innes argued that such extra help, which is not allowed on NAEP, skewed the state results upward.

Gary Phillips, the acting Commissioner of NCES, said the agency would not underwrite any more outside research into the issue. Peggy Carr, Associate Commissioner in charge of NAEP, said that NCES would examine possible policy changes in reporting scores.³⁸

A New Testing Cycle – Exclusions Again

The issue of exclusions and accommodations came before the R&D Committee again as it prepared for the release of 2000 NAEP reports on mathematics, science, and reading. To study the effects of accommodations, the national and state samples for the 2000 assessment had been divided in half. Half of the students completed the NAEP assessment under standard conditions. The other half was allowed to use appropriate accommodations.

At its August 2000 meeting, the Board decided that the NAEP 2000 reports would include data from both samples—those tested under standard conditions and those with accommodations. To maintain trends, results for students tested without accommodations would be designated as the "official reporting sample," with results for students tested with accommodations presented later in the report.

After being alerted by NCES and ETS staff that the proportion of students excluded from the non-accommodated sample had increased substantially over the past 3 years and that there were wide variations among states, the R&D Committee asked NCES to develop recommendations on how to handle this issue in the NAEP 2000 reports.³⁹ On November 17, 2000, Peggy Carr said that NCES recommends that significant changes in exclusion rates be noted in NAEP reports. She proposed that all jurisdictions with a change of 3 percentage points or more from a previous assessment should be indicated in the data tables with a flag or mark.

In response to a question from Ed Donley, Acting NCES Commissioner Gary Phillips said 3 percent was chosen as the point to make a notation because that is where the change is statistically significant. Phillips said the flagging would affect 10 of the 40 jurisdictions that participated in the math assessment.

Acknowledging that states had not been advised of this possibility when they agreed to participate in NAEP 2000, Ms. Carr said some notations and discussion were still necessary to report results fairly and avoid charges that NAEP was trying to hide the situation.⁴⁰

After discussion, the R&D Committee voted unanimously to recommend that the Board adopt guidelines for the presentation of data on exclusion rates and accommodations in NAEP 2000 reports:

- Exclusion and accommodation rates for all state and national samples, including data for all previous years used in trend comparisons, would be presented in an appendix.
- A change in exclusion rates of 3 percentage points or more from a previous assessment would be noted with an explanatory footnote written in objective, nonjudgmental language.
- No notations or flags would accompany current-year comparisons.⁴¹

The resolution was adopted by the full Board at its closing session. However, those who believed that action brought the difficult issue to a close were to be disappointed.

Déjà Vu – Flagging Revisited

On March 21, 2001, NCES presented new recommendations on presenting differences in exclusion rates in NAEP 2000 reports to a joint meeting of the Committee on Standards, Design, and Methodology (COSDAM) and the R&D Committee.

NCES had continued its data analysis after the November 20 meeting and concluded that flagging would be inappropriate because it was not possible to determine with sufficient accuracy a particular point at which the exclusion rate was significant. NCES recommended that the flagging portion of the November 2000 Board resolution be rescinded. Instead, it proposed to include explanatory language and a footnote indicating that changes in results over time may be affected by changes in exclusion rates. An appendix would be attached to reports describing several analyses of the relationship between changes in exclusion rates and state scores.

COSDAM Chairman Edward Haertel stressed the importance of explaining how the research had been done. Michael Ward said the report must be open and forthcoming to avoid any appearance of a cover-up. When the committees voted unanimously to recommend the NCES plan to the full Board, he abstained because of the possible impact on how results might be reported for his state.⁴²

The Board adopted the recommendation and the NCES plan was used in the NAEP 2000 reports.

Although more protracted than most issues, the exclusion and accommodation situation is an example of the diligence and perseverance with which the Governing Board and NCES approach any matter that might compromise the integrity of the National Assessment and its value to the American public.

A Recommendation from a Friend

Before serving as Secretary of Education in the Clinton Administration from 1992 to 2000, Richard Riley had been Governor of South Carolina and a member of the Governing Board. In comments to a Board meeting on February 19, 1999, he asked the Board to consider the achievement levels because they were “not very useful in showing where improvement is taking place or not taking place.” Riley observed that in the last Writing report a great many students were at the Basic level and NAEP was not able to convey that many more of them were in the top half of that achievement level than in the bottom half.

“The Basic and the below Basic categories are very broad and the Advanced category is so very narrow that it’s hardly useful.” He encouraged the Board to consider establishing another achievement level to be able to report better on movement within the range of the assessment.⁴³ Subsequently, the Board contracted with ETS to explore possible enhancements to the reporting of results by achievement levels.

The Governing Board

Using data from the 1994 NAEP Geography assessment, ETS developed achievement level descriptions based on what students know and can do at the various levels as demonstrated by their performance on the assessment. This differs from the current achievement levels developed in the standards-setting process that describe what students should know and be able to do at each level.

COSDAM Chairman Edward Heartel said that using descriptors based on actual student performance on actual NAEP questions put the achievement levels on a much firmer basis than using descriptors based on test frameworks. Board staff suggested that the ETS recommendation be tested by applying the procedure to the 2001 Geography results and the committees adopted a resolution to that effect for the Board’s consideration.⁴⁴

The second aspect of the ETS study explored new ways of reporting on the performance of students whose scores are approaching the Basic and Proficient achievement levels. The study explored three methods of defining “approaching.” The first was a statistical method that used the median score of students below Basic and between Basic and Proficient. The other two were based on a judgmental process—one in which a lenient judge may have placed the achievement levels and the other in which a lenient panel of judges may have set the levels.

Regardless of the method chosen, an analysis of the 1994 Geography results based on the new definitions resulted in greater differences between high-scoring student groups and low-scoring student groups than the ranges in the current reporting system.⁴⁵

Governing Board staff recommended that the “approaching” concept be defined as students having attained most of the knowledge and skills needed for Basic level work, rather than on the midpoint of the distribution (which would be difficult to explain) or on the degree of agreement among judges in the level-setting process (whose recommendations the Board did not always accept).

Staff also recommended that the study be continued based on the 2001 Geography assessment, but only for students “approaching Basic” since the Board’s policy definition for Basic (“partial mastery of prerequisite knowledge and skills that are fundamental for proficient work”) already embodied the concept of “approaching Proficient.”⁴⁶

There was no shortage of opinion among the committee and other Board members in the meeting. Michael Ward said that most state assessments, including those in his own state of North Carolina, focus on Proficient, so it would be more useful to have information about students approaching that level than Basic.

Chairman Haertel said that either of these two new “approaching” achievement levels would detract from the Proficient level, which for many years had been the goal for all students to attain.

Michael Nettles, Vice Chairman of the Board, said it was necessary to get more information on how the new “approaching” concept might work for both levels.

Diane Ravitch said that no matter what the Board called the new reporting concepts they would be regarded by the press as two new achievement levels, thus making Proficient seem high up the ladder rather than something all American students should aspire to.

R&D Committee Chairman John Stevens echoed Ravitch’s sentiments, adding that it might be wise not to proceed with the “approaching” idea at all.

The committees developed a resolution combining the recommendation of ETS and the Governing Board staff. “Approaching” would be defined close to the cut-scores for Basic and Proficient at each grade (one-fourth of a standard deviation or 10 points, for example) and the descriptions of what students know and can do would be developed using the methodology suggested by ETS.

The resolution clarified its intent by stating that “In no event will the approved cut-points of the achievement levels be changed. Approaching Basic and approaching Proficient are not to be regarded as new achievement levels, but as a means of reporting on and describing the skills of students who are close to the Basic and Proficient levels.”

The resolution was approved by a 9 to 2 vote, with the “nays” cast by the chairmen of the two committees, Ed Haertel and John Stevens.⁴⁷

On March 3, 2001, after extensive debate and having been advised of the committee Chairmen's opposition, the Board unanimously adopted a motion by Thomas Fisher referring the issue to its staff for additional work.⁴⁸

Staff reported to another joint meeting of the two committees on May 11, 2001. Sharif Shakrani, Deputy Executive Director of the Governing Board, called previous proposals problematic because they could be misinterpreted as creating additional achievement levels; that is, below the Basic and Proficient cut-points. He presented a new proposal that would overcome this problem by describing the knowledge and skills students need to move from below Basic to Basic and from Basic to Proficient.

Jim Carlson, also of the Governing Board staff, informed the committees that the reason for pursuing this topic is that many NAEP users have expressed a need for more detailed information about the performance of students within the achievement levels. He cautioned that, although a procedure like the one Dr. Shakrani described might be useful for subjects like mathematics, it would not work well for subjects where the skills and information to be learned are less distinct.

He suggested postponing action on the issue while continuing to study it, given the workload involved in preparing for the President's *No Child Left Behind* (NCLB) initiative and changes being considered for NAEP.

The committees passed a motion encouraging further discussion and analysis, and the new reporting initiative was postponed indefinitely.⁴⁹

A New Administration Brings Dramatic Change

During the first year of the new administration in Washington (2001), the Governing Board responded to legislative actions that had implications for the reporting and dissemination of NAEP results.

Anticipating that President George W. Bush's education reform initiative would require annual state-level testing of reading and mathematics, at a special meeting on June 28, 2001, the Governing Board adopted a policy statement providing that:

- Beginning in 2003, the national and state-level results for grades 4 and 8 reading and mathematics shall be released within 6 months after the completion of testing.
- The primary means of releasing initial grade 4 and 8 reading and mathematics results each year shall be a summary report intended for a general public audience.⁵⁰

These were two very significant and interdependent changes. NCES finally agreed that by changing the nature of the report used in the initial release, the Board's longstanding goal of issuing assessment results within 6 months of testing could be achieved.

By the November 17, 2001, Board meeting, it was clear that the legislation pending before Congress would require reading and mathematics to be assessed every other year rather

than annually. The Board adopted the R&D Committee's recommendation that results would be available on the Internet at the time of the scheduled NAEP release, including one block of test questions for each grade.⁵¹ The release of test questions was part of the Board's commitment to help NAEP users better understand the nature of the assessment and the learning represented by the achievement levels.

On January 8, 2002, President George W. Bush signed into law his major education reform initiative, the *No Child Left Behind Act*. The Governing Board's anticipatory actions had positioned the National Assessment to play an important role in this major education reform initiative. NCLB required all states to participate in the NAEP Reading and Mathematics assessments as a condition of receiving federal aid. Now the nation would have available a steady stream of National Assessment results in reading and mathematics for all states, the District of Columbia, Puerto Rico, and Department of Defense schools.

A Promise Kept

At the end of the May 27, 1999, hearing into the politicization and exclusion issues, Representatives Hoekstra and Castle, both Republicans, promised that Congress would act to end any confusion between NCES and the Governing Board about the responsibility for the release of NAEP results.

The promise was kept. However, with a new administration, Senator Ted Kennedy took the lead in preventing future NAEP releases from being politicized.

On November 2, 2002, President Bush signed a law creating the Institute for Education Sciences (IES). The law included a provision that the Governing Board shall "plan and execute the initial public release of National Assessment of Educational Progress reports." Other continuing duties included (1) to "develop guidelines for reporting and disseminating results," (2) to "develop a process for review of the assessment..." and (3) to "exercise... its responsibilities... independent of the Secretary and the other offices and officers of the Department."⁵²

The new law dramatically shifted the responsibility for reporting NAEP results to the Governing Board from NCES and made other responsibilities more explicit. Although it was not always easy, the Governing Board and NCES worked together to make the transition as nondisruptive and smooth as possible.

The Initial Release of Results

On May 17, 2003, the R&D Committee issued its first recommended policy statement addressing the exercise of the Board's new responsibility to "plan and execute the initial public release of National Assessment of Educational Progress reports."

The statement said that implementation of the law would require (1) the assumption by the Governing Board of certain activities conducted by the Department of Education, (2)

collaboration and cooperation with the Commissioner of NCES, and (3) additional resources to carry out required activities. The statement further clarified that:

- Responsibility for planning and executing the initial release “means organizing and conducting the actual release event for a NAEP report and the arrangements and activities leading up to and following, which are related to that event.”
- The Board determines the time, location, agenda, and presenters for the event. (The statement made a commitment that the Commissioner of NCES would be asked to present NAEP results.)
- “Planning the initial public release” includes setting policy for the content of NAEP reports.⁵³

Defining the Reports

At the next Board meeting in August, the R&D Committee presented a *Policy Statement on Reporting National and State-Level NAEP Results in Reading and Mathematics* to formalize policies related to NCLB that required those subjects to be assessed every 2 years. The statement provided that:

- The primary means for the initial public release shall be a summary report.
- Beginning in 2003, the initial reports shall be ready for release within 6 months of testing.
- The manner of release shall be determined by the Board.
- After the Board has received the final text and tables of each completed summary report, the Chairman of the R&D Committee, on behalf of the Board, shall determine the date of the initial release. The initial release shall be within 30 days after receipt of the completed report.⁵⁴

The statement also addressed the Board’s new responsibility to provide guidance for the presentation of the report and to determine its content, including:

- Reports must be written in a clear, jargon-free style with attractive charts, tables, and graphics.
- A limited number of contextual variables and subject-specific background information must be included to enrich and give perspective to the results.
- Sample questions must be used to illustrate the range of student performance and exemplify the NAEP achievement levels.

The statement also specified issues related to the reporting of achievement levels, average scale scores, percentile distributions, and trend data.⁵⁵

The Ogilvy Report

The Board sought outside expertise to provide critical advice as it took on its new responsibilities. Ogilvy Public Relations Worldwide delivered its report, *NAEP: Reporting Initial Results – Analysis and Recommendations for Improvement* on July 22, 2004.

The Board and NCES had worked together over the years to produce more user-friendly materials to support the initial release of NAEP results, and they were proud of their work. However, Ogilvy found many areas for improvement:

Policy guidelines were not functioning as stated and intended.

Guideline: The primary audience is “the general public.”

In practice, the audience has turned out to be education policymakers, advocacy groups, federal and state officials, educators, parents, and representatives of media outlets that reach them.

The “interested users” had difficulty locating the information they wanted, charts lacked necessary explanations, and there were too many complex phrases and terms, assuming technical knowledge they simply did not have.

A readability analysis found that NCES reports are written at a college graduate level, whereas the *Chicago Tribune’s* NAEP coverage is written for a high school graduate or college freshman.

Guideline: Reports are to be “relatively brief, written in a clear, jargon-free style with attractive charts, tables, and graphics.”

In practice, the reports have become increasingly lengthy. In just one year,

Highlights went from an 8.5 x 11, 25-page report to a 9 x 14, 35-page document.⁵⁶

Highlights reports need improvement.

A summary of important information should be at the front of each report.

A consistent data-reporting structure would help users find the information they want and the use of color would help with navigation.

Better graphics could help explain achievement level and subgroup results.

Technical explanations should be kept to a “necessary minimum.”

The current 9 x 14 size is difficult to store because it does not fit on a common bookshelf.⁵⁷

NAEP Web access is not user friendly.

The Web site is difficult to locate.

Once on the NAEP site, it is difficult to find the results because there is a tremendous amount of information on a very large site.

The design and presentation of information should be different from that in the printed report.

Ogilvy recommended creating a separate *Nation’s Report Card* Web site for the initial release.⁵⁸

Ogilvy stated that “NAGB could itself take on the challenge of re-thinking and re-designing the initial reports and Web environment. NAGB could oversee (with NCES involvement) the creation of new format options for the report and the development of a new Web site architecture. Once developed and agreed upon, these would become a flexible design template for future NAEP releases.”⁵⁹

Over the next several months, the R&D Committee, Governing Board staff, and Ogilvy representatives worked extensively to develop recommendations for the Board's consideration. This intense activity culminated in several actions at the November 2004 Board meeting.

A Full Agenda for the November 19, 2004, Board Meeting

First, the R&D Committee brought forward an updated Policy Statement that clearly delineated the responsibilities of the Governing Board and NCES, and also addressed Part I of the Board's expanded role: Report Preparation and Content.⁶⁰

Second, the Board also approved the Committee's recommended "Specifications for NAEP 2005 Reports," laying out very detailed descriptions that left no questions unanswered. It also outlined detailed requirements for *The Nation's Report Card's* Web site.⁶¹

Third, the Board adopted the Committee's recommended "Reporting Schedule for 2005 NAEP Assessments" that covered 13 reports to be released beginning in September 2005 and continuing through Fall 2006 or Winter 2007.⁶²

The schedule reflected a tremendous amount of work to be done by the Governing Board, NCES, and contractors as they dealt with data analysis, report preparation, and release activities for assessments of Mathematics, Reading, and Science for the nation, states, and trial urban districts, plus several special reports.

More New Responsibilities Addressed

In the months before the May 2005 Board meeting, the R&D Committee completed recommendations for two more sections of the policy statement that addressed the expanded responsibilities given to the Board: the Public Release of NAEP Results, and Dissemination and Outreach.

Part II outlined the Board's responsibilities and authority for planning and conducting the initial public release of NAEP results and its relationship to NCES throughout that process.⁶³

Part III explained the Board's ongoing activities to disseminate information about the National Assessment and its results.⁶⁴

These new sections were adopted by the Board on May 20, 2005, completing a statement that included policy principles and guidelines for reporting, release, and dissemination of National Assessment results.

Puerto Rico – No Jurisdiction Left Behind

Meeting of the R&D Committee

“This is totally unacceptable,” said Luis Ramos, responding to a presentation by an NCES staff member who had explained carefully that it was impossible to report results of the 2005 NAEP for Puerto Rico because of technical reasons.⁶⁵

As a native of Puerto Rico, Ramos was intensely interested in this subject. Ramos was a business and civic leader from Pennsylvania who had served on the State Board of Education and the local school board in Allentown; the Governing Board had first come to know him when he served as a member of the Commission on the 12th grade NAEP.

A specially translated Spanish language version of the NAEP mathematics assessment had been administered to students in Puerto Rico in 2003 and 2005 to comply with NCLB requirements, but the results had not been reported. (The Secretary of Education exempted Puerto Rico from the NAEP reading assessment because that test is an assessment of reading in English.)

In a memorandum to the Board, NCES said the 2003 results for Puerto Rico in both grades 4 and 8 could not be placed on the NAEP scale because the percent correct on many items was low and omit rates were high, particularly on constructed-response questions. Suzanne Triplett of NCES said results for 2005 have presented similar problems that have not yet been resolved.⁶⁶

The tendency of the Governing Board had been to defer to NCES on such technical matters, but Ramos would not be deterred. He said if the results show low performance the data should still be released to the public, and argued that NAEP needs to be transparent and should not hold back information.

Chairman John Stevens said that the Committee would take up the issue of NAEP reports on Puerto Rico again in May. After the meeting, Ramos and Stevens met with staff from NCES and the Governing Board to express their determination that results for Puerto Rico would be finally released. At the full Board meeting on March 4, Mr. Stevens announced that he and Ramos would meet with NCES and Board staff to try to resolve the Puerto Rico reporting issue before the May Board meeting.⁶⁷

Joint Meeting of NCES and Governing Board Staff

On April 6, 2006, Ramos and Stevens attended a joint meeting of NCES and Governing Board staff to discuss the Puerto Rico situation. NCES staff had prepared a PowerPoint presentation to explain how the test had been administered in Puerto Rico.

For 2003, the translation had been done by ETS bilingual staff using a version of the language commonly found in Spanish-language textbooks. Using standard NAEP procedures, the test was administered to approximately 3,000 students in 100 schools for each grade. There was

a large amount of missing data. The average result was 25 percent of students getting each item correct, which was too low to give stable results on the NAEP scale.

For 2005, the translation was reviewed by native Puerto Ricans and about 35 percent of the items were revised for cultural and linguistic suitability. Two workshops were held in Puerto Rico for the purpose of increasing participation and response rates by reviewing test content and released materials. Testing procedures were changed to allow an additional 10 minutes for each block of items, to explain test item types, and to encourage students to attempt all items.

NCES staff presented their further analysis of the Puerto Rico data, concluding that some meaningful results could be derived from the assessment and discussed the reports that would be produced.⁶⁸

The Governing Board Meeting

At the May 19, 2006, R&D Committee meeting, Peggy Carr reported on the work that had been done since March. She explained that initial analyses indicated that results could not be reported for both 2003 and 2005 because of a high nonresponse rate and a poor fit of the Puerto Rico response patterns with the national model. However, more recent analysis found that results could be reported reliably.

Accordingly, Carr announced that plans are in place to report the Puerto Rico data in fall 2006. She said that, to explain the delay in the release and describe the research that has taken place, three reports will be issued:

1. A “Highlights” report will include the 2003 and 2005 scale scores and achievement level results.
2. A technical report will explain the background of the Puerto Rico assessment, how the assessment was conducted in 2003 and 2005, changes made in test administration, and the research that now makes it possible to report results for both 2003 and 2005.
3. An item analysis report for Puerto Rico will show the types of questions and content areas students in Puerto Rico tended to answer correctly and incorrectly.

Luis Ramos expressed his appreciation to Ms. Carr and her staff for the extra efforts they made to produce reports on the Puerto Rico results. He said educators in the Commonwealth should pay close attention to the NAEP data. The Committee recommended that the reports be released at a press conference in Puerto Rico.⁶⁹

Education Week reported on the progress that had been made in reporting results for Puerto Rico. It quoted the NAEP Coordinator for Puerto Rico’s education department, José A Rivera, who said he believes the lack of alignment between the test and the curriculum could be a factor in students’ answering questions incorrectly or leaving them blank. In response to e-mail questions, Rivera wrote, “The Puerto Rico Department of Education and their students take NAEP seriously. This is not a factor driving down scores.”⁷⁰

Releasing the Results

On March 29, 2007, NAEP results for Puerto Rico were released by NCES Associate Commissioner Peggy Carr and Board member Luis Ramos at a news conference in San Juan. They released two reports, *The Nation's Report Card: Mathematics 2003 and 2005 Performance in Puerto Rico—Highlights* and *The Nation's Report Card: Mathematics 2005 Performance in Puerto Rico—Focus on the Content Areas*.

CommunicationWorks, the Board's public information contractor, worked closely with a public education agency in Puerto Rico to prepare for the event. A major effort was made to inform reporters about NAEP because this was the first time the National Assessment had tested students in the Commonwealth. The NAEP reports and all materials were available in Spanish and the entire news conference was conducted in Spanish, with Ramos speaking his native language and Carr using an interpreter.⁷¹

In announcing the results, Carr said there were some problems with the translation of the assessment used in 2003, but the problems were resolved for the 2005 test with the involvement of native Puerto Ricans.

In 2003, 9 percent of fourth grade students in Puerto Rico scored at or above Basic compared to 76 percent of students in the nation. Four percent of eighth grade students in Puerto Rico scored at or above Basic compared to 67 percent in the nation.

In 2005, 12 percent of fourth grade students in Puerto Rico scored at or above Basic compared to 79 percent in the nation. Six percent of Puerto Rico's eighth grade students scored at or above Basic compared to 68 percent in the nation.

Ramos said, "These results suggest that there is considerable room for improvement, but they also provide a good understanding of where student performance in Puerto Rico is strongest and what areas may need extra attention." As an example, he pointed out that island students scored relatively well in geometry, but they scored considerably lower on "word problems" that required them to show their work and explain their answers.

"It is important to keep in mind that the results we are talking about today represent a starting point to which we can compare future scores in order to measure growth in achievement over time," he said.

In reacting to the release, Puerto Rican officials urged caution in interpreting the results because of "linguistic and cultural bias." But regardless of the language issue, Waldo Torres, the island Education Department's Deputy Secretary for Academic Affairs, said the test results would serve as an important "instrument of analysis" to make improvements in the math. He said more math classes would be offered and there would be more emphasis on solving word problems, rather than strict memorization and solving formulas.⁷²

The NAEP results were a major story in Puerto Rico. Not surprisingly, mainland coverage was limited:

- There were stories in every daily newspaper on the island and several front-page articles, including an article in the *San Juan Star*.
- The release was featured in nearly 20 TV and radio stories, including coverage by Univision TV and Univision Radio, Telemundo, and the San Juan Public Broadcasting System.
- Combined coverage both in Puerto Rico and elsewhere in the U.S. reached an audience of more than five million people.⁷³

Release of the NAEP 2007 results was delayed until December 2008 because of the same kind of difficulty as that stated earlier in analyzing the data to produce useful results. Meanwhile, officials in Puerto Rico were feuding about whether or not the island should even participate in NAEP.

Puerto Rico’s Secretary of Education, Rafael Aragunde-Torres, wrote to U.S. Department of Education officials requesting that the island be permanently exempted from participating. However, that effort was rejected by recently elected Governor Luis G. Fortuno, who wrote to Secretary Spellings that removing Puerto Rico from the assessment would “do a terrible disservice to students, parents, and teachers” on the island.⁷⁴

Reporting Standards Applied to R&D Policy Statements

Following the enactment of legislation in 2002 that expanded the responsibility and authority of the Board, the R&D Committee worked with Board staff, NCES, and contractors to establish new policy and guidelines for the reporting, release, and dissemination of National Assessment results. Realizing that the separate statements it had developed on reporting, release and dissemination were rather cumbersome and not very user friendly, the R&D Committee decided to look at them with the same “critical eye” it had applied when improving the presentation of information in NAEP reports.

The culmination of the long-term collaborative effort and the critical review was represented in two documents recommended for adoption by the Governing Board at its August 6, 2006, meeting.

The first, an eight-page document titled *Policy Statement on Reporting, Release, and Dissemination of NAEP Results*, delineates the general responsibilities of NCES and the Governing Board and lays out the policy principles for (Part I) Report Preparation and Content, (Part II) Public Release of NAEP Results, and (Part III) Dissemination and Outreach. The appendix describes the materials to which the principles apply.⁷⁵

The second document, *Guidelines for The Nation’s Report Card*, includes detailed specifications for the two primary mechanisms for use in the initial release of NAEP results—the printed *Highlights* report and the dedicated Web site.⁷⁶

The continuing impact of this 3-year effort is significant:

- The Nation’s Report Card is the official “brand name” of the initial report of NAEP results in print and on the dedicated Web site.
- The printed report is very different, as envisioned in the guidelines developed by the Committee.
- The new Web site exceeds everyone’s expectations. It presents the results clearly and allows users easy access to information in which they are interested.
- The report content has also changed. It features new graphics and shows both Basic and Proficient achievement level results. It emphasizes changing demographics and their powerful influence on results (a favorite topic of Chairman Darwin Winick). Technical information is presented clearly for the interested user.

The Board has also put its stamp on the release events:

- The events are often held in different cities around the country in venues that are relevant to the subject of the report being released.
- Interaction with the media has been enhanced with briefings, the provision of embargoed information in advance, and the “Peggy Carr Web Chat” immediately after the news conference, which gives interested parties an opportunity to ask questions and to obtain more information.
- Former Governing Board Executive Director Charles Smith began the practice of providing a short video clip commenting on the release, for use in televised coverage.
- The materials and comments at the event are entirely free of politics and unwarranted commentary.

Fittingly, this meeting brought to a close the service of R&D Chairman John Stevens on the Governing Board as he completed the second of two 4-year terms.

Closing Thoughts

This brief chronicle of events focused on reporting and dissemination gives some sense of how the National Assessment must constantly respond to the changing environment in which it operates. Both the Governing Board and NCES are aware that, as a report on educational progress, NAEP must remain fundamentally constant so that it measures real changes in achievement and other important factors. Yet, NAEP cannot be mired in the past. If it is to remain “the gold standard of educational assessment” it must embrace new testing technologies, it must figure out new ways to deal with the exploding demographic diversity of our nation, and it must constantly strive to provide the greatest value possible to its users.

As the R&D Committee and the full Governing Board dealt with the topics in this paper and other matters, a good deal of research and thought took place that never manifested itself in policy or practice for NAEP. Some of them warrant further exploration because they could help make NAEP more than just a “scorekeeper or referee.”

The enhanced description of achievement levels presented by ETS promised to report what students actually demonstrated they knew and could do at various levels on the assessment. Done well, that would give greater meaning to the assessment results. The report could also

show, as NCES staff suggested, what students need to learn to move up to the next level, an obvious benefit to those working to improve student achievement.

The racial/ethnic groupings have become a challenge as the country becomes dramatically more diverse and of mixed ethnicity. The National Assessment needs to help users understand the powerful impact of demographic differences between locations and over time.

The discussion of achievement gaps is critical, but perhaps they might better be explained in terms of percentile distributions rather than ethnic groupings. After all, the real purpose of education improvement should be to strengthen the performance of all students while reducing the difference between the lowest performers and the highest performers.

For example, about a decade ago, the media reported that there had been no change in reading achievement over the time span a NAEP report covered. This conclusion was reached because the average scale score for the nation was exactly as it had been for the earliest assessment included in the report. However, over that time span the weakest readers (those at the 10th percentile) were reading worse than they had been reading previously, and the strongest readers (those at the 90th percentile) were reading considerably better than they had been reading previously. That very significant and undesirable trend should have been a major and important finding, yet the National Assessment report had not called attention to it.

Finally, it is time for the knowledge and skills that are taught in our schools to be defined not by educators but by those who are aware of what is needed to participate and succeed in the job market, in personal life, and in other aspects of community life. By measuring and reporting on the knowledge and skills that are necessary to succeed in life, the National Assessment could help the nation's education system become more responsive to the nation and world in which we live. The "gold standard of assessment" would be even more of a "national treasure" if it could not only remain the best and most trusted instrument of educational measurement, but also inform and even motivate more purposeful actions for improvement.

Author's Notes

Serving on the National Assessment Governing Board was one of the most satisfying experiences of my professional life. I especially appreciated the opportunity to get to know and work with so many interesting, talented, smart, and dedicated people. Without sounding too much like an Oscar recipient, I must acknowledge a few of them by name. They are representative of many others.

I am grateful to Secretary Riley for first appointing me to the Board in 1998, and to my friend Rod Paige for reappointing me. It was a pleasure to serve under the leadership of Chairmen Mark Musick, who became my friend, and Darvin Winick, who already was my friend. The Governing Board Executive Directors, Roy Truby and Charles Smith, provided expert leadership for the staff and they were always helpful and supportive in all of the activities in which I was engaged. The Governing Board staff, especially Larry Feinberg, with whom I worked so closely and upon whom I depended so much, always gave to every endeavor the very best they had without reservation. The NCES Commissioners, especially my friend Pat Forgione, and staff, especially Peggy Carr, never held anything back in working with us both on routine and on the most difficult matters. The staff of the NAEP and Governing Board contractors provided essential technical expertise and the benefit of experience that were essential to our work.

It was remarkable to see this diverse collection of people, who came from so many different backgrounds and perspectives, work together effectively because of their common commitment to make the National Assessment the best that it could be. In writing the paper, I was frustrated because it was not possible to capture the diversity of the characters, the intelligence often displayed in the discussions, the passion and sometimes the tension evident in the room, and the way in which agreements were finally reached that drew on both technical expertise and unifying vision.

Then, one night at about 3 a.m., it struck me that this story really should be written as a short historical novella, capturing the personalities, the settings, the dialog, and everything else that made the experience so rich for me. I asked myself, "If lawyers and doctors and medical examiners could be action heroes, why not the Governing Board and NCES?"

Perhaps some other time...

Endnotes

1. Luis Ramos, *The Times Leader* (Wilkes Barre, PA), January 3, 2008, Article ID: 11E5C7F534F9B8C8.
2. *Policy Statement on Reporting and Dissemination of National Assessment Results*, Governing Board, adopted November 17, 1990, pp 3–4.
3. *Dissemination Strategies for the National Assessment of Educational Progress*, The Widmeyer Group, p. 7.
4. *NAEP 1992 Reading Report Card for the Nation and the States*, NCES, Office of Educational Research and Improvement, U.S. Department of Education, September 1993.
5. Widmeyer, p. 8.
6. Ibid, p. 9.
7. Ibid, p. 9.
8. Ibid, p. 26.
9. *Release Plan for the 1998 NAEP Reading Report Card*, Governing Board, November 17, 1998.
10. Subcommittee on Oversight and Investigations Hearing, U.S. House of Representatives, May 27, 1999, Educational Resources Information Center (ERIC), Document ED440373, p. 28.
11. “National Reading Scores Show Some Improvement,” *The Heartland Institute*, April 1999.
12. Letter from Mark Musick, Chairman, Governing Board, to Dr. Pascal D. Forgione, Commissioner of National Statistics, February 18, 1999.
13. Minutes of R&D Committee meeting, Governing Board, March 4, 1999.
14. "Richard G. Innes" 70224.434@compuserve.com, Subject: [education-consumers] RE: NAEP 1998 Reading -- CAUTION!!!, March 6, 1999.
15. Andrea Tortora, “School Test Results Questioned – Disabled Students Excluded,” *Cincinnati Enquirer*, May 10, 1999.
16. David J. Hoff, “Board Won’t Revise State NAEP Scores,” *Education Week*, May 19, 1999.
17. Minutes of R&D Committee meeting, Governing Board, May 13, 1999.
18. Memo from Nancy Allen et al. (ETS and Westat) to Peggy Carr, NCES, April 26, 1999.

19. "NAEP Test Results Questioned," *The Education Reporter*, July 1999.
20. Ibid.
21. Gail Russell Chaddock, "What's New," *The Christian Science Monitor*, May 18, 1999.
22. Subcommittee on Oversight and Investigations Hearing, U.S. House of Representatives, May 27, 1999, Educational Resources Information Center (ERIC), Document ED440373, p. 7.
23. Ibid, p. 9.
24. Ibid, p. 9–10.
25. Ibid, pp. 101–102.
26. Ibid, pp. 10–12.
27. Ibid, pp. 76–79.
28. Ibid, pp. 84–85.
29. Ibid, pp. 23–24.
30. Ibid.
31. Ibid, pp. 24–25.
32. Ibid, pp. 29–30.
33. Ibid, p. 35.
34. Ibid, p. 57.
35. Memo from Wayne Martin, CCSSO, to Holly Spurlock, Peggy Carr, and Suzanne Triplett at NCES, June 18, 1999.
36. Memo from Stephen Lazer, ETS, to Peggy Carr, NCES, June 20, 1999.
37. Minutes of COSDAM and R&D Committee joint meeting, Governing Board, August 5, 1999.
38. David J. Hoff, "Ky.'s 1998 NAEP Gains Declared Statistically Significant," *Education Week*, October 6, 1999.
39. Minutes of R&D Committee meeting, Governing Board, November 17, 2000.

40. Ibid.
41. Ibid.
42. Minutes of COSDAM and R&D Committee meeting, Governing Board, May 2, 2001.
43. Transcript of Governing Board meeting, November 19, 1999, p. 4.
44. Minutes of COSDAM and R&D Committee meeting, Governing Board, March 2, 2001.
45. Ibid.
46. Ibid.
47. Ibid.
48. Minutes of Governing Board meeting, March 2001.
49. Minutes of COSDAM and R&D Committee meeting, Governing Board, May 11, 2001.
50. *Policies and Procedures for Reporting State-Level NAEP Results in Reading and Mathematics*, adopted by Governing Board June 28, 2001.
51. *Policies and Procedures for Reporting State-Level NAEP Results in Reading and Mathematics*, adopted by Governing Board November 17, 2001.
52. *Policy Statement on the Initial Public Release of National Assessment of Educational Progress Reports*, Governing Board, May 17, 2003.
53. Ibid, pp. 2–3.
54. *Policy Statement on Reporting National and State-Level NAEP Results in Reading and Mathematics*, Governing Board, August 2, 2003, p. 2.
55. Ibid, p. 3.
56. *NAEP: Reporting Initial Results*, Ogilvy Public Relations Worldwide, July 22, 2004, pp. 9–16.
57. Ibid, pp. 20–25.
58. Ibid, pp. 27–30.
59. Ibid, p. 5.

60. *Policy Statement on Reporting, Release, and Dissemination of National Assessment Results*, Governing Board, November 19, 2004.
61. *Specifications for NAEP 2005 Reports*, Governing Board, November 19, 2004.
62. *Reporting Schedule for 2005 NAEP Assessments*, Governing Board, November 19, 2004.
63. *Policy Statement on Reporting, Release, and Dissemination of National Assessment Results*, Governing Board, May 20, 2005, pp. 6–7.
64. Ibid, pp. 9–10.
65. Minutes of R&D Committee meeting, March 3, 2006.
66. Ibid.
67. Ibid.
68. *Analysis of Puerto Rico 2003 and 2005 Results – And Plans for Reporting*, PowerPoint Presentation, NCES/Governing Board meeting, April 6, 2006.
69. Minutes of R&D Committee meeting, May 19, 2006.
70. “Puerto Rico Still Awaits NAEP Math Results,” Sean Cavanagh, *Education Week*, May 3, 2006.
71. Minutes of R&D Committee meeting, May 18, 2007, p. 3.
72. Rosario Rajardo, “Island students score low on math tests,” *San Juan Star*, March 30, 2007, pp. 5–6.
73. Minutes of R&D Committee meeting, May 18, 2007, p. 4.
74. Sean Cavanagh, “Puerto Rican Officials Feud over NAEP Participation,” *Education Week*, November 21, 2008.
75. *Policy Statement on Reporting, Release, and Dissemination of NAEP Results*, Governing Board, August 4, 2006.
76. *Guidelines for The Nation’s Report Card*, Governing Board, August 4, 2006.