1988–2008
NATIONAL ASSESSMENT GOVERNING BOARD
NAEP AND THE PROGRESS OF EDUCATION IN AMERICA
20th Anniversary Conference Proceedings
MARCH 4, 2009
What is the National Assessment of Educational Progress?

For more than four decades the National Assessment of Educational Progress (NAEP) has been the only nationally representative and continuing assessment of what American students know and can do in a range of academic subjects. Authorized by Congress, the survey has become the Nation’s Report Card, providing reliable, independent information on student achievement in reading, mathematics, writing, science, U.S. history, geography, the arts, and other elementary and secondary school subjects. NAEP state-by-state assessments—also on a representative-sample basis—began in 1990, and provide the only comparable data on student achievement in different states. The Trial Urban District Assessment (TUDA), which began in 2003, tests representative samples of students in many of the nation’s largest cities.

NAEP’s national samples include both public and private schools. The state and TUDA samples are of public schools only, including both regular and public charter schools. All states are required to participate in NAEP reading and mathematics assessments in the fourth and eighth grades every two years. Participation in all other NAEP subjects and grades is voluntary.

The National Assessment tests a broad framework of material, and includes both multiple-choice and open-ended questions. Results are reported in terms of both scale scores and Basic, Proficient, and Advanced achievement levels. The assessments are given by NAEP’s own trained administrators to ensure standardized conditions and security. NAEP is forbidden by law to maintain information on individual students and schools. The assessment is carried out by the National Center for Education Statistics in the U.S. Department of Education.

What is the National Assessment Governing Board?

An independent, widely representative panel, the National Assessment Governing Board (NAGB) was created by Congress in 1988 to set policy for the National Assessment. The 26-member, bipartisan Board includes two governors and two state legislators—from different political parties; two chief state school officers; and a broad array of teachers, local school officials, testing and curriculum experts, and business and public representatives. Members are appointed by the Secretary of Education from lists of nominees in various statutory categories.

After a nationwide consultative process, the Board determines the content and performance standards for each NAEP assessment. It approves all test questions and is responsible for the methodology of the assessment and the initial public release of NAEP results. Its major projects in recent years have included improving NAEP testing and reporting of students with disabilities and English language learners, developing the framework for a new assessment of technology and engineering literacy, and conducting research on using 12th grade NAEP to report on preparedness for college and workforce training.
The 20th anniversary conference and the papers that accompanied it were the culmination of a year-long effort by a planning committee of past and present members of the National Assessment Governing Board, of Board staff, and of our authors and consultants.

The event itself was lively and thought-provoking. The conference theme, “NAEP and the Progress of Education in America,” was meant to signal its purpose—to look back, of course, but also to analyze the present and look forward to NAEP’s future. We hope what we heard will have an impact not only on those who attended but also on those who “listen-in” through the podcasts on our Web site and through this proceedings document. Sheila Byrd, the author of the proceedings, has captured the key points of what transpired.

I would like to thank the members of the 20th anniversary planning committee whose service on the Board has spanned its entire history—Richard Boyd and Mark Musick, who were Board members at the very beginning in 1988; Edward Donley, Michael Guerra, and Christine Johnson, who were active later on; and David Driscoll and Eileen Weiser, who are current members. It was an honor and a pleasure to serve as chair of this group.

Also, I wish to express appreciation to the Board staff who helped make all the arrangements—particularly, Mary Crovo, Interim Executive Director, and Lawrence Feinberg, assistant director for reporting and analysis, who served as principal staff to the planning committee and provided editorial support for these proceedings and the many other papers prepared for the event. I am also grateful for the efforts of Mary Ann Wilmer, the Board’s operations officer for its first 12 years, who came back as a consultant to assist with planning the anniversary events.

I wish to thank the many people who set aside time in their busy schedules to share their insights and show their support for education reform, NAEP, and the work of the Board by speaking at the conference and the anniversary dinner, including Education Secretary Arne Duncan and three of his predecessors—Richard Riley, Rod Paige, and Margaret Spellings.

The purpose of the Board’s 20th anniversary commemoration was to look at the role of the National Assessment in educational change, how that role itself has changed over the past two decades, and how it may change again. Certainly, the uses of NAEP and the visibility of NAEP have become greater and greater over the past 20 years. And this has made the work of the Board even more important as we move into our next 20 years. We hope these proceedings will preserve and extend the productive discussions we had.

Amanda Avallone
Chair, 20th Anniversary Planning Committee
Vice Chair, National Assessment Governing Board, 2006-2010
NAEP and the Progress of Education in America
20th Anniversary Conference Proceedings

By Sheila Byrd Carmichael
Education Policy Consultant

The views expressed and synopsized in this publication do not necessarily reflect the views of the National Assessment Governing Board, the National Center for Education Statistics and its contractors involved with the National Assessment of Educational Progress, or the U.S. Department of Education.

The representations of the events and statements of the presenters have been provided as accurately as possible. The summary of these proceedings is the responsibility of the author and any discrepancies with actual events is unintentional.

SUGGESTED CITATION

The National Assessment Governing Board wishes to thank all conference participants, the 20th Anniversary Planning Committee, and Board staff members Mary Crovo, Lawrence Feinberg, Stephaan Harris, Munira Mwalimu, and Tessa Regis for their work in making the program a success.

Planning Committee Members

Amanda Avallone
Planning Committee Chair and Board Vice Chair
Assistant Principal and Eighth-Grade Teacher
Summit Middle School
Boulder, CO

Richard Boyd
Board Alumnus
Former Mississippi State Superintendent of Education
Oxford, MS

David Driscoll
Board Member
Former Massachusetts Commissioner of Education
Melrose, MA

Edward Donley
Board Alumnus
Former Chairman
Air Products and Chemicals, Inc.
Allentown, PA

Mark Musick
Board Alumnus
James H. Quillen Chair
East Tennessee University
Johnson City, TN

Michael Guerra
Board Alumnus
President Emeritus
National Catholic Education Association
Washington, DC

Eileen Weiser
Board Member
General Public Representative
Ann Arbor, MI

Christine Johnson
Board Alumnus
Education Consultant
Morrison, CO
# Table of Contents

Preface ........................................................................................................................................................................... i

Introduction ..................................................................................................................................................................... 1

Conference Highlights .................................................................................................................................................. 1

Conference Welcome and Keynote Speakers .............................................................................................................. 2

Panel One: NAEP and the States ................................................................................................................................. 3

Panel Two: Achievement Gaps .................................................................................................................................. 7

Panel Three: Grade 12 NAEP – Measuring Preparedness for College and Work .................................................... 9

Panel Four: Testing Student with Disabilities and English Language Learners .................................................... 13

Plenary ......................................................................................................................................................................... 15

Remarks by Arne Duncan, U.S. Secretary of Education .......................................................................................... 15

Panel Five: NAEP and Standards for American Education ....................................................................................... 17

Conclusion .................................................................................................................................................................. 21

Anniversary Dinner: Former Education Secretaries and Governing Board Executive Directors .......................... 22

Postconference Presentations to Board by Former Board Chairs ......................................................................... 22

Summary of Conference Issues ................................................................................................................................ 24

Materials .................................................................................................................................................................... 25

Recognition of NAEP ............................................................................................................................................... 25

Letter from Senator Edward M. Kennedy .................................................................................................................. 26

House Resolution 222 .............................................................................................................................................. 27

Historical Papers ..................................................................................................................................................... 28

Conference Papers .................................................................................................................................................... 28
To commemorate 20 years of setting policy for the National Assessment of Educational Progress (NAEP), the National Assessment Governing Board convened a conference in Washington, DC, on March 4, 2009. In attendance were more than 200 national, state, and local education experts and policymakers, including current Board members and almost 20 of the Board’s 130 alumni.

In the first part of the program attendees participated in a lively set of panel discussions on:

- NAEP and the States
- Achievement Gaps
- Grade 12 NAEP: Measuring Preparedness for College and Work
- Testing Students with Disabilities and English Language Learners

U.S. Secretary of Education Arne Duncan then offered an overview of the Obama Administration’s educational priorities and expressed strong support for the National Assessment and for the independence of the Governing Board. A final plenary panel session reunited former Board members and federal policymakers to discuss the role of NAEP in state efforts to develop academic standards and assessments. The panel also discussed NAEP’s possible future role in the context of renewed interest in establishing common, voluntary, national education standards and assessments.

The conference was held at the National Press Club.

Prior to the conference, the Board commissioned six papers that examine the history of the Board and the National Assessment, illuminating some of the issues discussed by the panels. These papers are available at http://www.nagb.org/who-we-are/board-anniv.htm#conf-papers.

In the evening, the Board hosted a 20th anniversary dinner. The featured speaker, former Education Secretary Richard Riley, was a member of the Board (as Governor of South Carolina) at its founding in 1988. Two other former Education Secretaries, Rod Paige and Margaret Spellings, also addressed the gathering. In addition, the 125 guests heard from Representative Michael Castle (R-DE), who had been a Board member as Governor of Delaware in the early 1990s, and from the Board’s two former executive directors, Roy Truby (1989–2002) and Charles E. Smith (2003–2008).
Conference Welcome and Keynote Speakers

The conference began with welcoming remarks by Board Chair Darvin Winick and Vice Chair Amanda Avallone, who headed the 20th anniversary planning committee. There was also a greeting from Interim Executive Director Mary Crovo.

There were two conference keynote speakers: Senator Lamar Alexander (R-TN), former Governor of Tennessee, who cochaired the Alexander-James commission that proposed establishing the National Assessment Governing Board in 1987, and Senator Edward M. Kennedy (D-MA), Chair of the U.S. Senate Committee on Health, Education, Labor, and Pensions, who sponsored the legislation creating the Board in 1988. Both speakers praised the Board as a consistently independent, bipartisan voice in education policymaking, which is what the commission proposed and the legislation provided for.

Appearing in a DVD recording, Senator Alexander highlighted the significant role NAEP has played in providing policymakers with “clear and usable” facts about student achievement. As one result, he said, the nation is now focused on disaggregated data “like never before.” Senator Alexander suggested that NAEP be expanded to include testing in U.S. history that would yield state-level data.

As we look ahead to the future, we know that the role of NAGB and The Nation’s Report Card will not diminish. In fact, it will become even more important.

—Senator Lamar Alexander

NAEP data is an important tool for identifying and addressing the achievement gap.

—Senator Edward M. Kennedy

In his remarks, Senator Kennedy reminded the conference that NAEP data is an important tool for identifying and addressing achievement gaps, which have now become an issue of major national concern. Senator Kennedy’s message was read to the conference by Vice Chair Avallone.
It is an interesting time when you celebrate 20 years . . . to look back at the progress and relationship between NAGB, NAEP, and the states, and for that matter, the American public.

—David Driscoll

David Driscoll began the panel discussion by noting the progress states have made on educational standards, as well as the evolving relationship between the Governing Board, the states, and the general public. Mark Musick, former head of SREB, underscored the significance of the support that governors of southern states were willing to give in the late 1980s to a state-level NAEP assessment. They embraced an opportunity, he said, to receive and publish transparent information about actual student achievement in their states, despite the fact that the news might not be good. At one time, he noted, the state NAEP program was “illegal,” being officially proscribed by law. In just a generation, state NAEP has shifted from a trial to a voluntary program and then to a requirement under the No Child Left Behind (NCLB) law of 2001. State NAEP remains the best “truth-teller,” Musick declared, where achievement gaps are concerned, since states all set their own standards, and the rigor of proficiency levels can vary considerably.

Gene Wilhoit, of CCSSO, himself a former state school superintendent, related the story of state education chiefs who resisted the idea of participating in the state NAEP program at its inception. He recounted the heated debate in which many chiefs expressed trepidation about revealing potentially poor student performance in their states and recalled that the resolution endorsing state NAEP passed by only one vote at a CCSSO conference in 1986. Since then, Wilhoit said, state NAEP has had a dramatic impact on state policymaking, even more than on national education policy. In many states, NAEP has been a major driver of efforts to address the achievement gaps between white and minority students, which
NAEP revealed before state tests even acknowledged them. In difficult political environments, he said, NAEP has stood as a healthy guidepost for states trying to “whittle away” at racial disparities. In addition, the National Assessment has served as a model for the content and organization of many state standards and assessments.

Wilhoit suggested that the Board could help define with clarity a new set of national standards, which many policymakers are beginning to ask for again. He added that we should try to determine how well aligned NAEP expectations are with those in other countries. He wondered if we can bring the best elements of international standards together with our own, and yet not overtest. Wilhoit urged the exploration of new kinds of assessments that would employ new technologies.

Reflecting on his years as associate superintendent in North Carolina, Henry Johnson discussed the impact that NAEP had in helping the state define its own curriculum and assessments. Johnson said that later when he became state superintendent of Mississippi, some members of the public were reluctant to embrace the reporting of student achievement. He said “telling the public the truth” is a necessity, and noted that the media in particular appreciate the willingness of public officials to address achievement gaps openly and honestly after the release of state NAEP reports on student achievement. Johnson added that international assessments are becoming important points of reference for the United States. He suggested that the Board explore ways to conduct useful comparisons with international achievement data to inform the development of assessments in the United States.

West Virginia Superintendent Steven Paine recounted the process of developing state standards in his state. He noted the difficulty of moving to more rigorous standards but said there is considerable pressure to do so as the general public becomes more familiar with the state comparisons provided by NAEP. He commended the deliberate effort in West Virginia to use NAEP as a model and to embrace transparency in reporting as a way to inform state and local policy. He expressed interest in the possibility of developing a common set of standards across the states and urged participants to focus especially in the next few years on educator development and strengthening the corps of teachers.

Telling the public the truth is a necessity.
—Henry Johnson
Questions and Answers

Former Colorado Governor Roy Romer, who is also a former superintendent of Los Angeles public schools and a former member of the Governing Board, asked panelists to discuss the issue of agreeing on “a common, rigorous set of standards voluntarily arrived at by states.” He asked specifically why the assessment frameworks or standards on which NAEP is based have not been used as a model set of standards for the nation. If a new effort emerged to develop voluntary national standards, Romer asked, who should be designated to oversee that work, and what role should the Governing Board have in such an effort?

Gene Wilhoit acknowledged that it has taken states some time to develop their own standards and raise the bar for student achievement. He urged states to work with the Board on any new set of voluntary national standards so that they would be aligned with those assessed by NAEP. Mark Musick reminded participants that the NAEP frameworks would be a useful tool for this purpose. Wilhoit observed that educators’ original fears about NAEP reporting had been transformed into opportunities for leadership in raising standards and stated that another opportunity to do so should be welcomed. Steven Paine discussed the difficulty of defining achievement levels and the importance of providing support to teachers and schools even if common standards are developed. Former Board member Phyllis Aldrich suggested that policymakers need to experiment with value-added assessments. Henry Johnson said there is a need to develop more powerful diagnostic and formative assessments.

One participant asked about the potential advantages and pitfalls of using the 12th grade NAEP assessment to report on college and career readiness. Panelists noted the need to define “preparedness” carefully before any reports can be made and reminded participants that the Board is studying the issue now. Potential pitfalls mentioned by panelists include the possibility that too many students would be identified as unprepared; that 12th grade might be too late to inform students that they are not prepared; that if 12th grade NAEP were able to perform this function, states would need to examine and alter state policies so that postsecondary institutions would value 12th grade NAEP performance; that states might need much stronger intervention programs prior to 12th grade; and, finally, that regardless of the ability of 12th grade NAEP to measure readiness, the “motivation” problem would still have to be addressed, i.e., how 12th graders could become invested in taking and caring about the 12th grade NAEP assessment.
Peggy Carr, associate commissioner of NCES, presented an analysis of the gaps in mathematics course-taking between white and black students from 1990 to 2005. She observed that important gaps (for example, in the percentage of students taking algebra I) have been closed, but added that white students are now taking more higher-level math, which has opened a new gap with blacks. Carr noted that the timing of when a student takes algebra I appears to be the most significant indicator of the ability to succeed in advanced mathematics courses, such as calculus. (The earlier a student takes algebra I, the more likely it is that the student will succeed in advanced mathematics classes.) She added that progress in mathematics course-taking among black students has not resulted in the closing of performance gaps. She offered data on the persistent gap in mathematics achievement between white and black students on NAEP, even when accounting for the highest-level math course taken. She said NCES is currently analyzing the actual content of two key mathematics courses, algebra I and geometry, by examining syllabi and textbooks at 650 schools. NCES anticipates that the study may reveal whether course content varies significantly from school to school and for students from different racial groups.

Kati Haycock, director of the Education Trust, discussed how NAEP data has helped her organization prompt states and school districts to address achievement gaps. She called NAEP an effective instrument of “information, inspiration, and torture.” She spoke first of the simple power of being able to document student achievement to “complicate people’s thinking” and help them understand the contours of the achievement gaps. She said NAEP data has helped reformers not only to “alarm” educators with documentation of stunning gaps, but also to spread the good news of significant progress where it has occurred.
Haycock said NAEP provides information about what kinds of reform efforts are working for all students and which ones are not so that schools can refocus on more effective programs. Parents have become much more aware of NAEP data and of the disparities between student performance on state tests and their performance on the NAEP. Thus, NAEP has constructively “shamed” states into becoming more motivated to raise their expectations for all students. Finally, Haycock reminded participants that NAEP provides corroborating evidence that low achievement is not inevitable, that what schools do matters, and that “demography is not destiny.” She cited the ability of the Trial Urban District Assessment (TUDA) program, for example, to illuminate substantial differences in performance among similar populations of at-risk students in different cities.

Michael Casserly, executive director of the Council of the Great City Schools, focused his remarks on the effectiveness of the Trial Urban District Assessment (TUDA) program. A collaborative effort involving the National Center for Education Statistics, the National Assessment Governing Board, and the Council of the Great City Schools, TUDA measures the performance of public school students at the school district level. He noted that achievement gaps have closed in many of these cities since TUDA began, and disputed the notion that such gaps are immutable. Casserly said NAEP has been used at the district level to track where reforms are working and also as an important tool to make big-city schools more effective. “We are not passively witnessing gains,” he declared, and listed a series of efforts in TUDA districts to use NAEP data to inform professional development programs, to analyze the effect of resource levels, to understand family and school differences, and to identify potential issues regarding the differences between state standards and NAEP frameworks. He noted that TUDA is expanding from 11 to 18 school districts in 2009.

The final panelist, Michael Nettles, senior vice president of Educational Testing Service and a former Vice Chair of the Governing Board, spoke of NAEP’s ability to “inform and reform.” He asked what the Board could do differently to help drive reform and close achievement gaps, and suggested that two current limitations on NAEP might be changed to help the assessment become more of a reform instrument. The first limitation is “the problem of anonymity of schools”—that under current law, the names of those participating in NAEP cannot be disclosed. But Nettles said reporting school-level data might be helpful. For example, we might learn where African-American students are outperforming their peers, which is something we do not know much about now.

The second limitation cited by Nettles is that no student ever takes an entire NAEP exam and no student score is ever reported. He said generating individual scores for NAEP might spark more meaningful interest on the part of parents and students. Nettles suggested that the Board consider developing a parallel system of individual NAEP-like tests, similar to those used by the Southern Regional Education Board for reporting on workplace readiness under the High Schools That Work program.

Nettles called for improvement in the background or noncognitive data collected by NAEP on student, teacher, and school characteristics and academic-related programs and behavior. He encouraged the Board to make greater use of these data to improve explanations of student performance. Finally, Nettles encouraged the Board to consider having NAEP conduct longitudinal surveys to help understand student achievement beyond high school.

The panel session ended too late for audience discussion and questions and answers.
George Thornton, professor of industrial and organizational psychology at Colorado State University, summarized the report of the technical panel on 12th grade preparedness research on which he served. The panel was appointed by the Board to consider the potential of the 12th grade National Assessment for reporting on college and workplace readiness. The panel recommended that a variety of research be conducted to report preparedness benchmarks on the 12th grade reading and mathematics assessments. Recommended research includes:

- content alignment studies to examine the correspondence between NAEP and widely used college admissions and placement exams
- statistical-linking studies between NAEP and these exams, using empirical data and representative samples
- postsecondary surveys to gather information about cutoff scores for placement in college courses
- judgmental studies asking experts to examine NAEP items and recommend the score levels needed to qualify for job training programs or credit-bearing college courses

Thornton concluded that it would be possible to use 12th grade NAEP to report on preparedness, but that substantial research must be conducted first. He added that the question of setting different levels of preparedness for different college and career pathways should be studied. He cautioned against reporting preparedness state-by-state unless strong representative samples are obtained.
Speaking as a former high school principal and community college instructor, Christine Johnson spoke of the severe and ongoing need for remediation in reading, writing, and mathematics at early postsecondary levels. She lamented the increasing costs associated with that need. Johnson said she supports the administration of 12th grade NAEP assessments, but noted that states do not often test at that level. Charles Kolb reiterated Johnson’s call for good assessments and spoke about the need for the United States to be more concerned about our “connectedness,” both positive and negative, with the rest of the world. The juncture point of 12th grade preparedness is a critical one, not only for our economic vitality, but also for American democracy. Twelfth grade NAEP could become a “positive contagion,” he suggested, and eventually might be adapted to have meaning for the student as well as for policymakers.

John Stevens also decried the lack of preparedness among high school graduates, despite many years of standards-based education reform. He said there was a disconnect between what subject-area specialists sometimes insist on including in standards and what graduates need to be successful in the “real world.” “Educated” may mean something different than “job-ready,” he suggested. He wondered if having high expectations necessarily means having the same standards for everybody. Stevens asked if having different educational pathways would necessarily discriminate against disadvantaged students. Who are the winners and losers, he asked, in the one-size-fits-all systems we have now? He suggested that more attention be paid to career and technical education, especially programs that are well articulated with job certifications and with programs at local community colleges and universities. Twelfth grade NAEP may need to be grounded more in “the world of work” before it can be determined whether the assessment actually can measure readiness. He expressed support for the longitudinal studies proposed in the conference paper by Paul Barton.

Moderator Jay Mathews asked each panelist to answer the question: “Is there something we can change in NAEP that will help it to assess college and career readiness, and if not, who should be doing it, and how?” George Thornton said that NAEP standards need to be more closely aligned with what the workforce needs (versus postsecondary demands) if we are going to make reasonable statements about preparedness. Christine Johnson recommended that more attention be paid to writing and that the assessments should be mandatory. Charles Kolb suggested that the Board look carefully at how to translate 12th grade preparedness to the American people in order to “stir people up” so they will support educational improvement efforts. John Stevens suggested that the Board reduce the influence of subject-matter experts on NAEP assessment frameworks. In reading, he said, NAEP does a good job by using a variety of texts, but the focus in math is more strictly academic and limited. He recommended that the NAEP mathematics assessment emphasize the thinking skills needed as a consumer and in one’s personal life.

Questions and Answers

Pat Forgione, superintendent of the Austin Independent School District in Texas, asked whether the Board had considered embedding a subset of NAEP items within the high-stakes state tests that students already have to take to get a good estimate of 12th grade preparation. George Thornton responded that
reporting at the individual level would be difficult. Christine Johnson wondered what would happen in states that do not have tests beyond 10th grade. Concerning the issue of motivation and the possibility of making a 12th grade NAEP assessment mandatory, Charles Kolb suggested that such a policy would face serious political hurdles. John Stevens suggested that technology could be brought to bear to make the assessment more engaging for students. Kolb added that 12th graders must recognize that 12th grade is an important starting point and not the endpoint in their education.

Robert Wise, president of the Alliance for Excellent Education and former Governor of West Virginia, asked about the level of interaction between the Board and the ACT and SAT examinations, particularly since some states are using one or the other of these tests as part of their state testing systems. George Thornton said the Technical Panel has recommended statistical-linking studies between NAEP and both tests.

Roy Romer, former governor of Colorado who recently was superintendent of Los Angeles public schools, asked about the value of the 12th grade NAEP test for informing and improving grade-by-grade instruction. How can the 50 states use a 12th grade benchmark on NAEP as an anchor for back-mapping to earlier grades? John Stevens responded that NAEP frameworks might help with this process. George Thornton said NAEP should “stick to its knitting” and not become involved in any attempt to create grade-by-grade curriculum standards.
Daniel Domenech opened the panel by discussing his experiences as an English language learner (ELL) when he was a Cuban-immigrant child in New York City and as superintendent of schools in Fairfax County, VA, where 147 languages are spoken. He stressed the importance of validity in testing and said that for English language learners, the only way to make testing valid is to test in the student’s own language. He suggested that ELL students be given more time (and perhaps be screened for proficiency in English) before testing them in English. The testing of ELL students is an important issue for NAEP because these students are excluded from taking the assessment in different proportions in different states and districts.

Noting the importance of the Board devoting a panel to these issues, Miriam Freedman discussed special education law as it relates to NAEP administration. She suggested we “get back to basics” in testing and not allow accommodation policies to detract from NAEP’s ability to administer valid tests. She said to preserve its validity, NAEP needs meaningful, representative samples of students taking the assessment as it was designed to be taken. Tests can become invalid when the testing is altered severely to allow unspecified accommodations for students with disabilities or when students are improperly excluded from taking the test.

Sharif Shakrani reminded participants that NAEP is a survey administered to a sample of students. It is critical that the sample be representative of the full population it is designed to sample. Since 1992, the percentage of special education and particularly ELL
students in the samples has increased significantly. They are concentrated in certain states and large urban districts. Shakrani said inaccurate information about these populations and the exclusion of these students from NAEP testing can skew results. In California, for example, 25 percent of fourth grade students are English language learners. About 78 percent of those students participate in NAEP. In Texas, less than half of the ELL population participates in the NAEP assessment. For students with disabilities, local Individualized Education Program committees often determine whether students participate. These panels sometimes misinterpret the federal regulations for the Individuals with Disabilities Education Act, which should not be used to exclude students from NAEP’s matrix sample assessment that does not give results for individual students.

Martha Thurlow noted the progress that has been made in administering NAEP to students with disabilities. She said the NAEP program has been doing research on increasing participation and exploring accommodations. NAEP has provided training to state NAEP coordinators and changed its decision-making guidelines over time, exploring options for obtaining better representation. She cautioned that we should not assume that students with disabilities are unable to participate in NAEP. She recommended that the Board closely study which students are excluded and why. She also suggested that the Board reconsider its accommodation policies. She said it may want to look at state policies in this area and consider how they might inform the development of new policies for NAEP.

Questions and Answers

A participant asked whether allowing more time would enable more students with disabilities to participate in NAEP. Sharif Shakrani responded, saying that extra time is already the most common accommodation provided for students with disabilities. Miriam Freedman reiterated the importance of deciding clearly what a test measures, and that it is the test—not the student—that should determine whether something is an appropriate accommodation for the student or an impermissible modification to the test.
Secretary of Education Arne Duncan spoke to a plenary session of conference participants. He said the education priorities of the new Obama Administration are focused on:

- college-ready, career-ready, and internationally benchmarked standards
- great assessments
- great data systems
- teacher quality
- turning around poorly performing schools

Duncan called the $5 billion Race to the Top fund, part of the recently-enacted $786 billion stimulus package, a way to work with “a small number of states who are going to lead the way in all of these efforts.” He stated that the end goal of the new initiatives is “to dramatically increase” high school and college graduation rates. Doing so will require both expanding access to college and raising academic standards.

The Secretary praised the Board and NAEP for having “fundamentally changed the conversation in this country” by staying the course as “truth-tellers” where student achievement is concerned:

> If it were not for this, we would have states continuing to do their own thing willy-nilly. But when you see states where 80, 85, 90 percent of students are meeting their state standards [and] you have NAEP scores …at 15, 18, 20 percent, something is wrong. You guys have been the truth-tellers….And I’m just so, so appreciative….

What I think NAEP has done is expose the state-by-state differential, the state-by-state gap, and that it doesn’t make sense, and that it’s not fair for children. And I would go so far to say that we’ve lied to children and families historically….After 20 years, we are really at a crossroads. Are we going to continue to deceive? Are we going to continue to take the easy road, or are we going to start to tell kids the truth?

Over the past 20 years, Secretary Duncan said, NAEP data and NAEP standards have forged a “confluence of folks”—unions, business leaders, the Department of Education, states, school districts, and governors—“who are saying we have to do something fundamentally different.” He urged the Board to “continue to honestly report the facts about education in America.”
Michael Cohen commended the contribution of NAEP and the Governing Board to the development of standards and assessments in the United States, particularly in the area of transparency in testing, or “telling the truth.” NAEP has also helped us understand what “proficiency” means. Looking forward, he identified two policy areas with which the Board might reasonably become engaged. First, Cohen suggested that the Board consider the possibility of trying to benchmark NAEP to international assessments, citing recent and renewed calls for international benchmarking to inform the development of standards and assessments in the United States. He acknowledged the technical and political challenges associated with international benchmarking, but recommended that the Board consider doing it nonetheless.

Second, Cohen explained that a growing number of states are interested in defining common standards and possibly administering common assessments. He wondered what the role of NAEP might become in this environment. Cohen said we would still need a check on what state tests tell us by continuing to give NAEP, even if we were to have common standards and tests. Answers to questions about how and by whom national standards could be developed are not simple and straightforward, he said. Finally, Cohen remarked that the Board’s technical studies on what “college- and career-readiness” means are a very welcome addition to this important discussion. He said the Board can play an important role in clarifying these definitions by bringing empirical data to the fore.

Chester Finn, Jr., the first Chair of the Governing Board, opened his remarks by highlighting two “historic shifts” during the Board’s twenty-year existence: (1) the administration of state NAEP, allowing comparisons that had never been allowed or made before, and (2) the setting of achievement levels for use in reporting NAEP results. The adoption of achievement levels, he said, was a controversial decision and remains so today, given
the range of opinion on the methodologies for defining them. The Board was willing to concede that setting achievement levels is fundamentally an act of judgment—not of science. The Board should be commended, he suggested, for sticking with its decisions about the cut-scores and the meaning of Basic, Proficient, and Advanced. He said the NAEP achievement levels comprise the closest thing to national standards that the United States has ever had.

Still, Finn cautioned, NAEP’s reporting by achievement levels does not provide help at the building, parent, or student level; hence, interest persists in establishing “national standards” that can be related to individual student achievement. If national standards were compulsory, he argued, it might make sense for NAEP itself to become the national standard. But as long as standard-setting efforts are voluntary, and likely to be multiple, the United States will need NAEP as an external auditor of the various efforts. Finally, Finn reminded participants that no institutional mechanism exists that could own, operate, manage, update, and audit national standards. The nation would also need some entity to manage the testing programs that would follow from the standards, ensuring that tests are administered properly and scores are accurately reported. These are issues that must be considered in moving forward with any efforts to establish national standards.

Jack Jennings recalled the reformist fervor of the late 1960s when he first came to Washington and observed that we may be entering a similar era now with the interest in national standards and assessments. He echoed the commentary of others that NAEP is an invaluable tool for educators and policymakers, but cautioned the Governing Board against getting involved in the development of national standards. Federal law forbids a national test, he noted, and he recommended that standards be developed outside of Washington, “bubbling up” from the states. The Governing Board can contribute to the effort, given all it knows about standard-setting, he said, and it could lead the country in a conversation about that process. He said policymakers should also have a discussion about what it means to have national standards, which would necessitate agreement on how to develop tests and cut-scores, and how to align curricula.

Diane Ravitch recalled past discussions of voluntary national tests and speculated about future meanings of the word “voluntary.” The United States once had national standards for education, she said. These were created by the College Entrance Examination Board (CEEB) in 1901 and set the academic framework of American education for many decades until they were replaced in 1941 by the Scholastic Aptitude Test (SAT), which she called “a curriculum-free IQ test.” A similar shift has occurred, she asserted, with the standards movement having been hijacked by “test-based accountability” under which students are tested in reading and math skills, with curriculum having become irrelevant.

NAEP has always upheld the importance of a wide range of academic subjects. It is a curriculum-based assessment, and content matters.

—Diane Ravitch

NAEP is a curriculum-based assessment, she said, but most state tests are curriculum-free. She said states are manipulating the cut-scores for proficiency on their exams so that their students appear to be doing better than they actually are. Therefore, NAEP is an important external auditor of state reporting. If and when the No Child Left Behind (NCLB) law is reauthorized, Ravitch recommended that states be required to report their scale scores in addition to proficiency rates. She lamented the dwindling emphasis in schools on content areas other than reading and math, but noted that NAEP, on the other hand, has always upheld the importance of a wide range of academic subjects. She also deplored recent efforts to stress so-called “21st century skills” at the expense of content, reminding participants that without knowledge, skills are of very little value. Ravitch urged that two principles be maintained in any effort to create national standards: (1) that content matters, and (2) that NAEP be retained as an independent barometer of student achievement, regardless of whatever else occurs.
Discussion Among Panelists

Michael Cohen agreed with Chester Finn’s description of recent “ad hoc” approaches to setting common standards, but added that he has also been struck by the fact that states are increasingly enthusiastic about the need for common standards. Cohen suggested that mistrust in the federal government is especially high now, and so we need to learn from the state-led efforts currently underway about the best ways to establish common standards and assessments. Finn added that mistrust in the federal government had been exacerbated by failed efforts in the past to create national standards and tests.

Jack Jennings commented that it is not always the case that state tests are bad and NAEP tests are good. Rather, the states make choices, and their schools are not always teaching to the content of NAEP, which is a test that does not count in their accountability systems. Therefore, he said comparing results on the two tests is not really fair. Diane Ravitch said that in some areas almost all standards and tests are quite similar. Early reading assessments, for example, actually do not vary much across states. But she said states have been pressured to manipulate the passing scores because of unrealistic federal goals. NAEP, on the other hand, has remained consistent. Michael Cohen speculated that, in retrospect, it might not have been a good idea to allow states to give flexibility to schools in meeting standards in exchange for accountability. Perhaps instead, he suggested, states should have given schools the curriculum and other instructional tools they need.

Questions and Answers

Richard Whitmire asked Jack Jennings about Diane Ravitch’s characterization that many schools are focusing on “bubble kids”—those close to reaching a test’s passing score—in a “curriculum-free environment.” Jennings said some teachers are concentrating on the bubble kids, who are close to meeting state standards, but not all teachers are doing so. He argued that NCLB has in fact caused schools to pay much more attention to the achievement gap, prompted schools to use data to improve instruction, and encouraged schools to employ tutoring and outside assistance. More problematic, he said, are the penalties in the accountability system created by NCLB. Diane Ravitch agreed with Mike Cohen that, under the law, the nation has adopted an approach to accountability that is solely punitive. She suggested that more effort should be made to assist failing schools rather than to punish them.

Mark Musick added a comment about the use of the word “proficient” and recommended that states be required to report results for both the Basic and Proficient achievement levels instead of Proficient alone as most do now. State standards generally come closer to the Basic level on NAEP, but they should be shooting continuously for the Proficient level and should be making that clear to the public. Historically, he added, the public has not necessarily understood what “proficient” means or why it is the goal.

Former Colorado Governor Roy Romer asked about the issue of institutional structure that had been raised by Finn. He suggested that federal stimulus money be devoted to developing national standards that come from “the bottom, outside Washington.” He said the U.S. Department of Education could establish criteria for standards and tell states to come together to develop them. Romer noted that ACT, the College Board, the Governing Board, and Achieve have all been working in this area and could be helpful with the process. He said somehow we need to combine SAT and ACT standards with end-of-12th grade expectations that can be held out to students early in school so they understand their real end-of-school goals.
Anniversary Dinner: Former Education Secretaries and Governing Board Executive Directors

At a post-conference dinner on March 4, marking the 20th anniversary, former Secretary of Education Richard Riley delivered a keynote speech, highlighting some of the Board’s defining moments: adoption of achievement levels, gaining acceptance for state NAEP, and development of the Voluntary National Test (even though it was never deployed). Riley was a member of the original Governing Board in 1988 and worked with it for eight years as Education Secretary in the 1990s. He praised the Board for operating always as “a pragmatic, independent, problem-solving board.” Its members, he said, have been people “who come from different backgrounds—teachers, parents, superintendents, state board members, researchers, and who bring different points of view to the table when important decisions are made.” He reminded the audience of his own continuing interest in finding ways for NAEP to demonstrate “where improvement is taking place or not taking place.”

On the issue of national standards, Riley said, “I believe the states should determine their own education standards—after all, education is a state and local responsibility—but there should be a national model that the states can use.” He said NAEP provides a very good model and the Governing Board “has provided a very good example of the consensus-building that must be used when any national standards are created and set.” Riley said he is not sure exactly what the vehicle should be for developing national standards, but he said the standards should be voluntary and NAEP and the Board should play an important role.

Four other dinner speakers—former Education Secretaries Rod Paige and Margaret Spellings and former Board Executive Directors Roy Truby and Charles Smith—made a similar point, that the Governing Board has always stayed the course in its role as an independent and bipartisan voice in education. They said it should remain so. The Secretaries noted the collegiality with which the Governing Board Chairs had worked with them and expressed appreciation for their candor.

---

NAGB has operated always as a pragmatic, independent, problem-solving board.

—Richard Riley
Postconference Presentations to Board by Former Board Chairs

Chester E. Finn, Jr. (1988–1990)


At the March 6 Board meeting after the anniversary conference, former Board Chairs Chester Finn, Richard Boyd, and Mark Musick offered comments and suggestions on the future roles of the National Assessment and the Governing Board. They praised the Board for having transcended constituencies and for maintaining its independent voice in education policymaking.

Among their recommendations were the following:

- Explore ways to conduct some international benchmarking so that international reference points become part of NAEP reporting.
- Consider carefully the reauthorization of NAEP as part of the Institute of Education Sciences. Is this configuration working?
- Consider whether the delegation of authority from the Secretary of Education to the Board is working properly.
- Explore whether the Governing Board may be trying to do too many things and consider the possibility that doing fewer things might increase impact and efficiency.
- Revisit the process for development of NAEP frameworks, whether to continue the current subject-by-subject approach versus taking a longer-term, holistic approach across the curriculum.
- Continue to examine what “proficiency” means.
- Continue to play the role of truth-teller about trends and achievement gaps.
- Continue to maintain a strong relationship between the Board Chair and the Secretary of Education.
- Focus on the issue of communication. How well does the Board communicate with the general public? What could it do better?
Mark Musick recalled the first time NAEP ever assessed state samples in 1985–86. The program was sponsored by the Southern Regional Education Board and included just three states. Starting in 1990, state NAEP was authorized as a federal program, and 49 states volunteered to take part at one time or another before participation became mandatory in 2003. He said the work of the Board in developing consensus on the content of the assessment and responding to public concerns was an important factor in building support. He said it is important for the Board to continue to listen and respond. Musick said the motivation of 12th graders taking NAEP continues to be a problem. This point was also made by Richard Boyd, who said that low motivation may cause NAEP to understate the achievement of 12th graders.

Chester Finn suggested that Board members be wary in three particular areas: international benchmarking, 21st century skills, and national standards. On international benchmarking, he said the Board should look carefully at how international tests are designed and for what purpose they are administered. The Board should examine the relative merits of each potential assessment to which it considers linking NAEP. He urged the Board to “look under the hood” at what are called 21st century skills, such as critical thinking, information literacy, and cooperation, with special attention to whether they can be measured and their attainment solidly reported. He said NAEP should not try to focus on vague, immeasurable skills at the expense of real content, as Diane Ravitch had warned during the plenary panel. Finn also warned the Board to be attentive to burgeoning calls for national standards. Given the Governing Board’s historic and related role, he said it should carefully analyze what role if any it would play in such a process (e.g., as a continuing independent auditor, as a reconfigured agency to oversee new developments, or other roles). In any case, Finn cautioned, the Governing Board should not “sit on the sidelines and cross its fingers.”
Summary of Conference Issues

The 20th anniversary conference of the National Assessment Governing Board was convened in a policy-making climate ripe with renewed energy surrounding the potential development of national standards, or common standards, as they are most often referred to in their new configuration. Conference participants suggested various ways in which the Governing Board could be involved in such a process. As the only developer of an operationalized set of national education standards and assessments, the Board might be able to help define a rigorous standard-setting process, according to many participants. Much of what the Board has done over the years to improve and strengthen the National Assessment can and should inform any future efforts to develop national standards and assessments.

That said, most participants agreed that the Board should continue to oversee NAEP as it is—a reliable external auditor, given on a representative-sample basis under scrupulously uniform conditions—even for any new configuration of assessments that might be based on new voluntary standards. This is especially true, they pointed out, because the voluntary nature of the proposed new standards and assessments means there could easily be more than one set of national standards and assessments developed by various groups of states that decide to collaborate.

Panelists and participants observed, as others had at the Board’s 10th anniversary, how powerful state NAEP assessments had become for “truth-telling” about student achievement. They praised the NAEP frameworks and assessments for their role as a model for states that are trying to improve their standards. They expressed concern, however, about the motivation of 12th graders taking the assessment, especially as the Board seeks to develop 12th grade NAEP as an indicator of college- and career-readiness. The Board was praised for its efforts to address the difficult issues surrounding the testing of students with disabilities and English language learners.

Of particular note at this 20th anniversary was the widespread support for NAEP’s Trial Urban District Assessment (TUDA) program, which began in 2002. TUDA is a collaborative effort involving the Governing Board, the National Center for Education Statistics, and the Council of the Great City Schools. It uses district-level NAEP samples to measure the performance of public school students and allows for comparisons among similar groups of students in large urban districts across the country. TUDA is expanding from 11 to 18 school districts in the 2009 administration. As is the case with state NAEP and state policymakers, the TUDA seems to have been very helpful to district leaders in understanding student achievement. They are using the TUDA data, among other things, to inform curriculum and professional development. The achievement gaps in many districts have begun closing.

With renewed calls for international benchmarking, due to increased interest in international assessments such as TIMSS, PISA, and PIRLS, a number of participants suggested that the Board support the reporting of international benchmarks on NAEP. By linking NAEP to similar international assessments, the Board might foster a better understanding of what other countries expect students to know and be able to do, and of what their students have attained. In theory, this would increase the usefulness of NAEP to the public and enhance the NAEP standards-setting process. Former Board Chairs cautioned, however, that the purpose and design of such tests should be carefully considered.

Speakers and panel members expressed appreciation for the role that the National Assessment and the Governing Board have played as truth-tellers about student achievement in the United States. They stressed the importance of maintaining the independence of the Governing Board to ensure that the content and reporting of NAEP are not distorted by partisan considerations. The role of NAEP and the Board in any new national standards effort is unclear. But, no matter what eventually is decided, the need to maintain NAEP as a reliable, independent measure of student achievement would be unchanged.
Recognition of NAEP

In a March 3, 2009 letter to Board Chair Darvin Winick, Senator Edward M. Kennedy commended NAEP as an indispensable resource for the nation’s schools, teachers, and policymakers, and the Governing Board for ensuring that NAEP maintains the highest standards for its assessments.

In addition, the U.S. House of Representatives officially recognized the Governing Board’s 20 years of service in measuring student achievement by passing a House Resolution 222, introduced by U.S. Representative Michael Castle of Delaware, on March 9, 2009.
Dear Darv,

I’m grateful for the invitation to address all of you today, and I wish very much that I could be with you in person. The National Assessment of Educational Progress is an indispensable resource for our nation’s schools, teachers, and policymakers, and for two decades the Governing Board has been ensuring that NAEP maintains the highest standards for its assessments and for our nation’s youth.

Our current economic crisis reminds us that education is essential for the economic transformation our nation will need to stay strong for the future. Before the creation of NAEP it was difficult for policymakers to reliably assess how well American students were learning, and to set an effective strategy for the nation in this vital area. Thanks to NAEP, we now have an objective yardstick for the performance of students over time in a wide range of subjects.

The Governing Board’s work is more vital than ever today to ensure that accurate and dependable information is available about the achievement of the nation’s children. In providing it, you draw the attention of all of us in Congress to the persistent achievement gaps that mean too many of our low-income and minority students are still being left behind.

The Board was founded to help raise our goals for the future. By maintaining NAEP’s high standards, you have an essential role in keeping our eyes firmly on our goal of the best possible education for all students.

I look forward very much to working with you in the new Congress to guarantee that NAEP remains strong and that we continue to use the important information you provide as the foundation for improving education nationwide.

Today’s thoughtful conference is a fitting tribute to all you do so well, and I commend you for your distinguished service to America.

With great respect and my warmest wishes for the years ahead,

Sincerely,

Edward M. Kennedy
House Resolution 222
In the U.S. House of Representatives
March 9, 2009

Whereas the National Assessment Governing Board (the Governing Board) is an independent, bipartisan board created by Congress in 1988 to set policy for the National Assessment of Educational Progress (NAEP), commonly known as 'The Nation's Report Card';

Whereas the Governing Board is made up of 26 members, including Governors, State legislators, local and State school officials, educators, researchers, business representatives, and members of the general public;

Whereas when Congress established the Governing Board to oversee The Nation's Report Card, it ensured that the NAEP would be conducted independently and free from inappropriate influences and special interests;

Whereas in overseeing The Nation's Report Card, the Governing Board identifies subjects to be assessed, determines the content and achievement levels for each assessment, and approves all assessment questions;

Whereas The Nation's Report Card is conducted as a representative sample and currently includes National NAEP assessments (which assess the performance of students in grades 4, 8, and 12 in reading, mathematics, writing, science, U.S. history, geography, and other subjects), State-by-State assessments (which are administered to students in grades 4 and 8 to access performance in reading, mathematics, writing, and science), Trial Urban District assessments (which report on the achievement of 4th and 8th grade students in 18 urban school districts that participate in reading, mathematics, writing and science assessments), and long term trend assessments (which are administered nationally every 4 years to students ages 9, 13, and 17 to assess performance in reading and mathematics);

Whereas State participation in NAEP assessments is voluntary with the exception of reading and mathematics assessments, which States are required to administer to public school students in grades 4 and 8 every 2 years in an effort to measure student performance in reading and mathematics;

Whereas all students who participate in NAEP do so on a voluntary basis and NAEP is forbidden by law to maintain or report information on individual students or schools;

Whereas the Governing Board works to inform the public about The Nation's Report Card by communicating its results to a wide range of Americans, including educators, the media, and elected officials and policymakers at the National, State, and local levels; and

Whereas the Governing Board has served an important role in evaluating the condition and progress of American education for 20 years: Now, therefore, be it

Resolved, That the House of Representatives
(1) congratulates the National Assessment Governing Board on its 20th anniversary in measuring student academic achievement; and
(2) recognizes past and present members of the National Assessment Governing Board for their service to the Nation in improving elementary and secondary education.
The following papers were prepared for the 20th anniversary of the National Assessment Governing Board. They are available at the Board’s Web site, http://www.nagb.org:

**Historical Papers**


**Conference Papers**

Paul E. Barton, *Twelfth Graders and All Their Futures*, http://www.nagb.org/who-we-are/20-anniversary/barton-12th-graders-futures.pdf


Martha Thurlow, *Back to the Future for NAEP: NAEP and Students with Disabilities and English Language Learners*, http://www.nagb.org/who-we-are/20-anniversary/thurlow-on-SD-NAEP.pdf