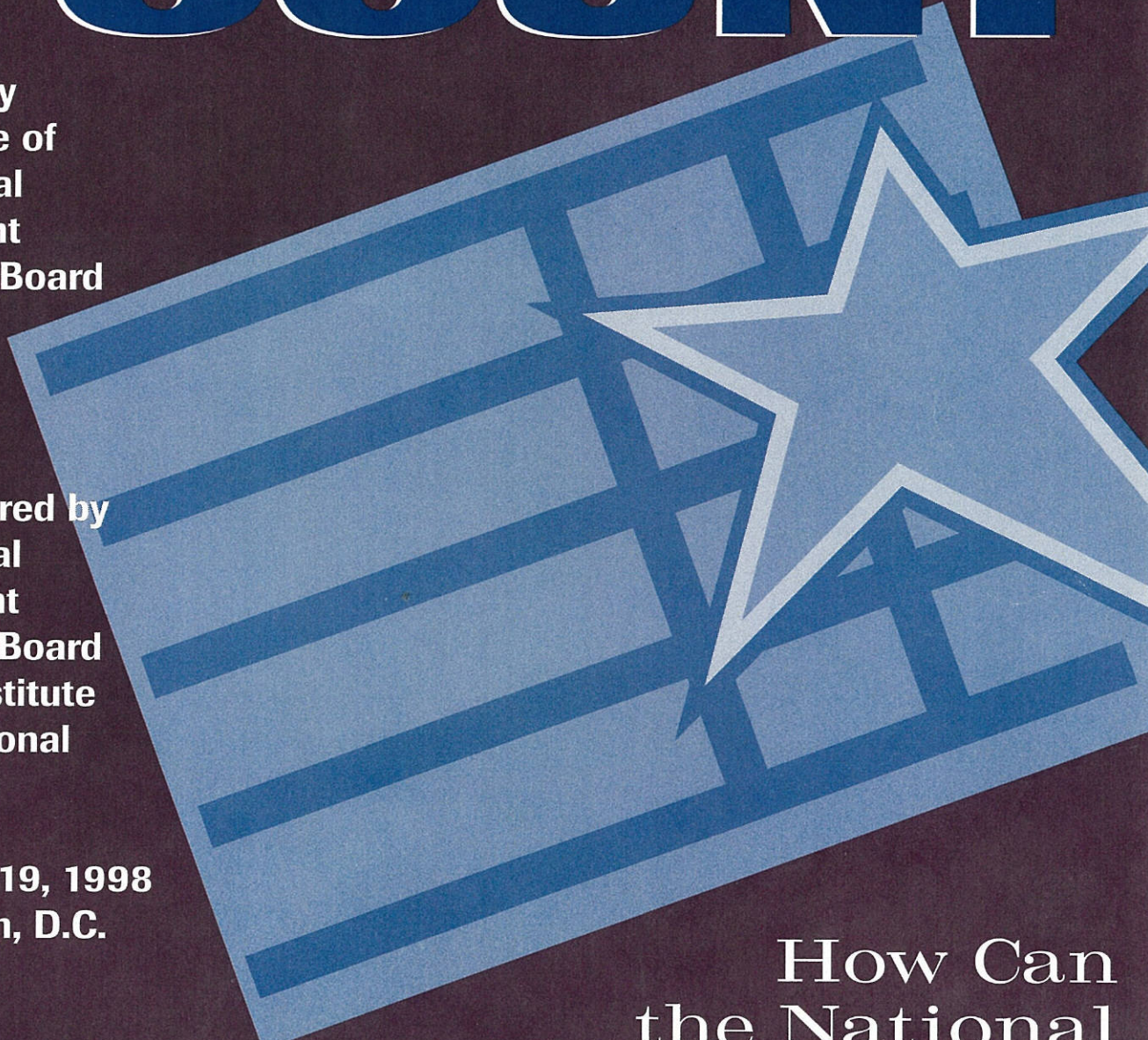


Standards **COUNT**

The Tenth
Anniversary
Conference of
the National
Assessment
Governing Board

Co-sponsored by
the National
Assessment
Governing Board
and the Institute
for Educational
Leadership

November 19, 1998
Washington, D.C.



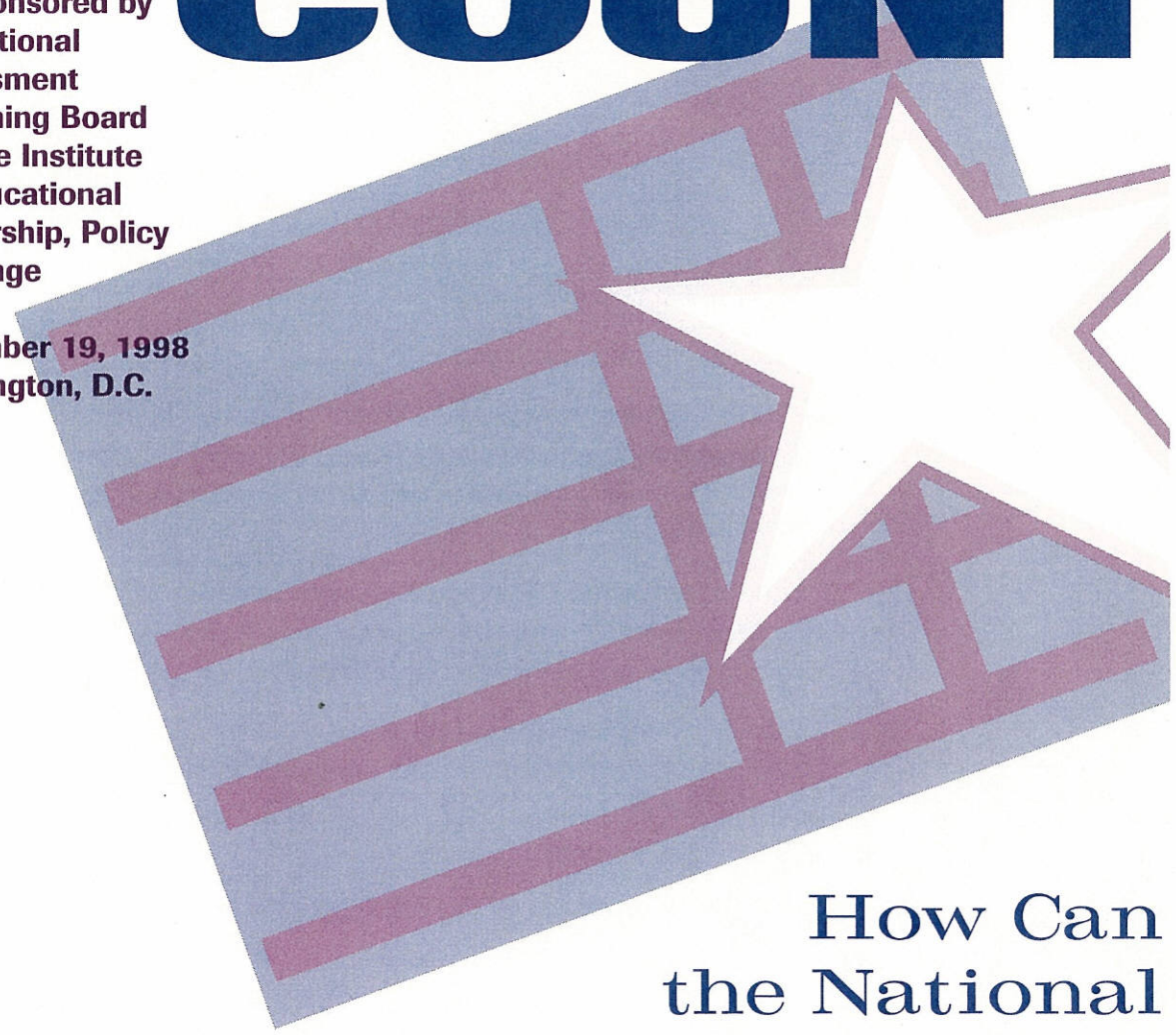
How Can
the National
Assessment of
Educational Progress
Make a Difference
in the Next Ten Years?

Papers Prepared for
the Tenth
Anniversary
Conference of
the National
Assessment
Governing Board

Co-sponsored by
the National
Assessment
Governing Board
and the Institute
for Educational
Leadership, Policy
Exchange

November 19, 1998
Washington, D.C.

Standards COUNT



How Can
the National
Assessment of
Educational Progress
Make a Difference
in the Next Ten Years?

**A REPORT
ON THE PROCEEDINGS OF A
CONFERENCE ON
THE FUTURE OF THE
NATIONAL
ASSESSMENT OF
EDUCATIONAL PROGRESS**

The Tenth Anniversary Conference
of the National Assessment Governing Board

By Harriet Tyson

Co-sponsored by
the National Assessment Governing Board
and the Institute for Educational Leadership

November 19, 1998
Washington, D.C.

ACKNOWLEDGMENTS

On behalf of the National Assessment Governing Board (NAGB) and the Institute for Educational Leadership (IEL), we would like to express appreciation to veteran education commentator Harriet Tyson for this cogent summary of the tenth anniversary conference of the Governing Board.

The Conference was co-sponsored by our two organizations on November 19, 1998. It was held in the Hart Senate Office Building in Washington, D.C., and drew nearly 200 participants, including members of Congress and Congressional staff, former NAGB members, and business and education policy leaders from Washington and across the country.

We would also like to express particular thanks to Larry Feinberg of the NAGB staff for his editorial assistance and note again the special contributions made by Margaret Dunkle and Lauren Handel of IEL's Policy Exchange in ensuring the success of our partnership.

Once more, may we express our gratitude to the funders of the tenth anniversary conference whose generosity permitted us to conduct the meeting, publish **Standards Count** (a collection of papers prepared for the conference), and, finally, issue this summary of the highlights of the meeting itself. Thanks again to:

The AT&T Foundation
The Donely Family Foundation
The Ewing Marion Kauffman Foundation
The Lockheed Martin Corporation
The Milken Family Foundation
The Procter & Gamble Company
Texaco, Inc.

We hope you find the discussions summarized in this report as interesting and valuable as we did.

Mark D. Musick
Chair
National Assessment Governing Board

Michael D. Usdan
President
Institute for Educational Leadership

**A REPORT
ON THE PROCEEDINGS OF A CONFERENCE ON
THE FUTURE OF THE NATIONAL ASSESSMENT OF
EDUCATIONAL PROGRESS**

November 19, 1998
Washington, D.C.

By Harriet Tyson

Introduction

The National Assessment Governing Board (NAGB) celebrated its first decade of policy responsibility for the National Assessment of Educational Progress (NAEP) by holding a conference, jointly sponsored by the Institute for Educational Leadership (IEL), in Washington, D.C. on November 19, 1998. The title of the conference, **Standards Count: How Can the National Assessment of Educational Progress Make a Difference in the Next Ten Years?**, was chosen to underscore the complementary and dynamic relationship between standards and testing. The conference had two main purposes: (1) to discuss the background, history, and future of the Governing Board and the National Assessment it oversees and (2) to examine the significance and future of the education standards movement and the role that NAEP and NAGB might play in its development in the years ahead. Panelists and participants represented a wide variety of stake-holders in American education, as well as an array of viewpoints on educational policy issues. A noteworthy feature of the conference was the appearance on videotape of three secretaries of education - the current Secretary Richard Riley, and two of his predecessors, Lamar Alexander and William Bennett—all of whom have been deeply involved with NAEP. The secretaries addressed questions posed by the conference planners.

This document complements the booklet, **Standards Count**, a collection of essays prepared before the conference discussions. While the brief essays by the panelists set forth their various ideas, the following account of conference proceedings shows how their ideas were tested in interactions with fellow panelists of differing views and with probing moderators and participants.

Background

The National Assessment Governing Board was authorized by P.L. 100-297, passed by both houses of a Democratic-controlled Congress and signed by President Reagan, a Republican, in April of 1998. The Board was designed as an independent, bipartisan agency to set policy for NAEP, which had been operating as a low-key federal program since 1970. The creation of the board was proposed in a 1987 report by a blue-ribbon study group headed by Lamar Alexander, then Governor of Tennessee and later Secretary of Education under President Bush, and H. Thomas James, President-emeritus of the Spencer Foundation.

The Alexander-James Commission, whose 22 prestigious members included Hillary Rodham Clinton, recommended that NAEP be expanded from an obscure national survey into a vehicle for collecting reliable state-by-state data on educational achievement, which had previously been unavailable. To govern this controversial new enterprise, the panel urged the creation of a strong, independent, and broadly representative Board to gain widespread state participation and credibility, ensure fair comparisons, and allay fears of federal control.

The charter creating NAGB was approved by Education Secretary William Bennett on September 5, 1988. The Board included 24 state and local officials, educators, and members of the public, and held its first meeting on November 18 and 19, 1988. Since 1990, 49 states have voluntarily participated in state-level NAEP assessments. Those assessments include fourth and eighth grade tests in reading, mathematics, science, and writing. Also, representative national samples have been tested in grades 4, 8, and 12 in those four subjects, plus U.S. history, geography, civics, and arts education. The Board has determined the content of all exams through a national consensus process.

NAEP results are presented both as scale scores (based on distributions of performance nationwide) and by achievement levels (set by the Board). The levels--called basic, proficient, and advanced--are based on judgments of what students should know and be able to do in each grade and subject tested.

The Conference

Prior to the conference, most of the invited panelists submitted brief essays setting forth one or several points they considered to be most important. Their collected papers are published in a booklet, **Standards Count: How Can the National Assessment of Educational Progress Make a Difference in the Next Ten Years?**, which can be obtained from either NAGS or IEL.

Each conference session was organized around a series of related questions, and panelists were asked to address those questions in their remarks. Invited panelists and participants included all present and former members of NAGS, members of the original Alexander-James Study Commission, former assistant secretaries for the Office of Educational Research and Improvement who have served as ex-officio NAGS members, the commissioners' of education statistics, members of Congress, several chief state school officers, heads of major education associations, business leaders, elected officials, teachers, members of the media, and as noted already, three secretaries of education.

Session I

How is the National Assessment of Educational Progress Important to the Nation, States, Schools, and Parents of Students?

What difference has NAEP made in education?

Why should policy makers, business leaders, educators, and parents care about NAEP?

Moderator: John Merrow, Host, The Merrow Report (PBS/NPR)

Panelists:

Senator Jeff Bingaman (D-New Mexico), Member of the Committee on Labor and Human Resources

Linda Bryant, Former NAGB member; Principal, Westwood Elementary School (Pittsburgh, Pa); and Pennsylvania Teacher of the Year in 1987

Representative Michael N. Castle (R-Delaware), Former NAGB member and
Vice-chairman of the Early Childhood, Youth and Families Subcommittee

Michael Ward, Superintendent of Public Instruction in North Carolina

Senator James Jeffords (R-Vermont), Chairman of Labor and Education Committee

Following greetings and introductions by **NAGB Chairman Mark Musick** and **IEL President Michael Usdan**, **Senator Jeff Bingaman** of New Mexico launched the session by saying that his state does poorly on the National Assessment. Nevertheless, he believes that the good information that NAEP provides will help New Mexico focus on the areas where students need to make progress. "NAEP has become the closest thing we have to a national system for measuring educational achievement. It provides comparable results nationally and across states. It uses rigorous academic material, and performance levels are described in terms I think folks can understand."

Also, Bingaman said, NAEP helps to counter the idea that it is acceptable for children in different states to learn different amounts. "I cannot see how we can justify a student growing up in New Mexico learning less math than a student in Boston or anywhere else...and clearly that's my view of where most people in American are on that issue."

Bingaman said that standards and accountability are the key to improving education, and that NAGB has a key role in this effort. Although he supports the continuing development of voluntary national tests of individual students in reading and mathematics, he noted the "national testing stalemate." He said that the erosion of affirmative action programs has rekindled attacks on all forms of standardized tests because of their impact on low-income and minority students. He also acknowledged that some critics want to do away with NAEP, as well as the National Education Goals Panel and its annual report.

In order to sort out the conflicting claims and misunderstandings about NAGB, NAEP and the National Goals Panel, Bingaman proposed holding a national education summit for governors and the President, patterned on the meeting in Charlottesville, VA in

1989. He said the summit would revamp the goals and recommend the changes necessary to achieve them over the next decade. Also, he advocated more frequent administration of NAEP subject-matter tests, reporting results of those tests not only at the national and state levels, but also at the school district level, and a more formal statement on the independence of the NAGB.

Superintendent Michael Ward believes that NAEP has been "awfully useful" in North Carolina. Although the state is proud of its own curriculum and tests, "what they don't generate is a gauge by which to compare the progress of our state and our students nationally. That's where NAEP comes in." Noting that many people in and out of his state believe that North Carolina trails the nation because it ranks 48th in SAT scores, Ward has been pleased to publicize the state's NAEP scores because they show that a representative sample of students score near the middle of the pack on all the NAEP tests, thus countering the state's negative image. Ward sounded a theme that recurred in subsequent sessions: parents don't pay much attention to NAEP results because there is no apparent link between NAEP results and what is happening with their own children. Therefore he favors a national, all-pupil assessment in reading and mathematics that would provide parents a basis for determining the progress of their children on national measures.

Linda Bryant observed that few teachers know or care about NAEP test results because the scores cannot be tracked to their students or even to their school. But she has seen the beneficial influence of NAEP on her school district as it moves toward standards and standards-based assessments. Despite the data yielded from local and state tests, Bryant still doesn't think she knows whether her children are learning enough to be "productive citizens in an ever-changing society." She therefore favors reporting NAEP data by school.

Representative Michael Castle confessed that he didn't know what NAEP was when he was Governor of Delaware, but learned something about it when he became a member of NAGB. He now favors offering NAEP results for individual students, but "at a minimum, breaking them out by the school and school district level." He

acknowledges that most members of Congress don't know much about NAEP, much less whether it should be applied to school districts and schools. But the people in Congress who are familiar with NAEP and NAGB are "very respectful" of them. The real issue in Congress, he said, "is not what to do," but rather, "should anything be done at the federal level?" The names of the test and the Board are difficult for people to understand, and should be changed, Castle said.

When NAEP scores became available, Castle discovered that "Delaware did less well than I thought we should. I thought Delaware should be in the top ten states based on income data, and I learned that we did worse. That was important news to learn... Even if my state had come in 50th, I would have been a total believer that you should have the best comparisons you can get. I would take NAEP down to the district and school levels. I think that information is helpful."

Superintendent Michael Ward, reacting to Castle's position, said it would be hard to carry the national assessment back to the local level without adding some sort of national governance structure and some sort of central national body to develop these sorts of assessments. He was concerned about the "schizophrenia" in Congress: "Everybody's in favor of accountability for the nation's schools, but many are reluctant to create the mechanism by which we can provide useful comparisons across the nation."

Representative Castle observed that critics of national testing say there is already a multitude of tests. "The critics tend to ignore the reality that testing companies compete fiercely with each other, and as a result, different areas give different tests." He strongly supports NAEP testing "because it is a fair comparison done fairly."

Participant, **Mary Blanton**, former Vice-Chair of NAGB, asked whether allowing district and school level testing and reporting would build more support for NAEP because more constituents would know about it. **Representative Castle** answered that the expansion of NAEP might have a counter-effect because it could arouse more charges of federal meddling in local schools.

Senator James Jeffords of Vermont affirmed his belief in standards and in the importance of having some measuring stick to know whether standards are being

achieved. He sees great political value in testing at the school district level because you can "wake people up" and motivate parents and educators to take action. The issue of standards, he said, will be considered by Congress "at considerable length" as part of the re-authorization of the Elementary and Secondary Education Act in 1999.

A participant observed that many states are not very adept at using NAEP data to improve instruction, and proposed that NAEP expand its responsibilities to include the training of state education department personnel and teachers on how to analyze and make use of NAEP test results.

Moderator **John Merrow** listed some of the commonly-heard reasons for not making better use of NAEP: delay in reporting scores, infrequent administration of some of the tests, the fact that students don't have much motivation to do well on NAEP because the tests have no consequences for them, and finally, the problem that some critics think NAEP achievement levels are set too high, or else that the discrepancies between state and NAEP standards of proficiency create unsupportable confusion. Panelists nevertheless agreed that many states have used the NAEP model to upgrade their state tests, and that some states have made good analytic use of NAEP data to shore up weaknesses in their curriculum and instruction.

Session II

What State Issues Must NAEP Address in the Future?

What impact should NAEP test frameworks and achievement levels have on state standards?

Should states adopt NAEP standards, or would that produce an undesirable loss of variety and local autonomy?

Should NAEP be used as a "gold standard" to judge whether state standards are high or low?

What impact should state standards have on NAEP?

Should states control NAEP?

What impact should NAEP results have on state policies, curriculum and funding?

Moderator: William T. Randall, Project Director, Connect; Former
Commissioner, Colorado State Department of Education; Former NAGB
Chairman

Panelists:

Ronald R. Cowell, Member of the Pennsylvania House of Representatives and
Chairman of the Committee on Education

Timothy P. Kelly, Education Policy Director, Office of Governor
John Engler of Michigan

Henry Marockie, State Superintendent of Schools in West Virginia

Robert Schwartz, President, Achieve, Inc.

Moderator William Randall described "All States 2000," launched by the National Center for Education Statistics. The goal of the project is to get all states and territories to participate in the year 2000 national assessment so that there will be a comprehensive baseline for the new millennium.

Ron Cowell expressed strong support for the continuation of NAGB, NAEP, and the National Education Goals Panel. If Congress is to be persuaded to let these efforts continue, he said, the duties assigned to each must be more specific and more significant. Referring to the "Value to States" section in the most recent Goals Panel Report, he cited the following uses of NAEP: (1) to monitor educational progress over time, (2) to benchmark state performance against the best in the nation and the best in the world, and (3) to monitor whether all groups of students in a state are achieving at high levels. Many states, he said, use NAEP scores for these purposes. But Cowell thinks state policy makers could also use NAEP as an analytical tool to determine what is working and what is not. At present, he said, the analysis of program effectiveness in the states is not one of NAEP's duties. In the future, he thinks that assignment should be given to some group at the national level, whether it is NAGB, the Goals Panel, or some other entity.

Cowell, echoing Superintendent Ward, noted the "schizophrenia" among policy makers: they want information, accountability, and comparisons, "but in the name of

local control, we often will not embrace and actually oppose some of the tools...that would serve those purposes that we say we're all interested in."

Tim Kelly, of Governor Engler's office, believes the challenge facing the states is to determine whether their standards are high enough and if their assessments measure the knowledge and skills that children truly need to succeed. "The NAEP test provides a necessary and trustworthy external yardstick...of reliable data on how students are doing in relation to the rigorous standards of other states, the nation as a whole, and more importantly, other countries." Michigan uses NAEP standards, well-regarded standards from other states, and Third International Mathematics and Science Study (TIMSS) data to test the rigor of its own standards and assessments. As a result, said Kelly, it has produced some significant improvements in mathematics achievement.

Superintendent Henry Marockie made a vigorous case for the many uses of NAEP in state education policymaking and instruction. West Virginia used NAEP's "rigorous and visionary" curriculum frameworks to develop its own state standards, and selected the SAT 9 test, "the most rigorous of the commercial, norm-referenced tests, which contains many NAEP-like items." Also, West Virginia has used NAEP's frameworks as statewide textbook selection criteria, and has incorporated them into all the state's staff development programs. Marockie believes that this close alignment of all elements of its education program with NAEP standards and assessments is one of the reasons why West Virginia has scored so well in reading -- 13th place in the nation, while third from the bottom in per capita income.

Robert Schwartz, of Achieve, Inc., advanced a contrasting view of the ideal relationship between NAEP and the states. "NAEP was designed to give us periodic snapshots of trend lines over time of what students at a particular age level know. Its virtue is to provide an independent check on the effects of the ups and downs of various education reform "fads, trends, and state policies, which change all the time. This was not a test that was designed to be taught to or to have consequences for students or schools."

State tests, based on state standards, Schwartz noted, have now been developed by 39 states, and ten states are in process of development. "At their best, these are assessments that are deliberately designed to be taught to and are in fact worthy of being taught to. State assessments that are customized to that state's standards can provide feedback that is really useful in guiding instruction, and individual and school level data can really drive parental involvement."

"States need both kinds of information -- state tests that measure how well students are learning the state program, and NAEP tests, which have an arms-length relationship to curriculum and instruction and provide periodic information on how well students can apply what they have learned." NAEP tests, argued Schwartz, should not yield scores at the district or school or individual level. "Let NAEP be NAEP," he said. "Don't try to make it do anything it wasn't intended to do. I think we ought to take very seriously what Congressman Castle put on the table, warning us of the political consequences of pushing NAEP to produce district-level or school-level achievement data."

Superintendent Marockie disagreed, saying that NAEP is highly respected, becoming better known, and shouldn't be a threat to anyone. Having each state develop its own standards and tests makes it impossible to compare student performance across state lines, and the attempt to get consortia of states to agree on common standards is like reinventing the wheel when strong standards and excellent assessments –NAEP--already exist. He favors publicizing NAEP in order to raise its visibility with the public and policy makers.

Schwartz, however, believes that NAEP has been useful because it has not been intrusive. "If you look at the stalemate that surrounds a relatively innocuous proposal for voluntary national tests, it suggests to me a certain degree of caution about having NAEP provide district and sub-district data... It seems to me it is very important to make sure that NAEP gets re-authorized and is able to continue playing the very important function it has played in these last 20 years."

Ron Cowell said that having a national debate on a mandatory national test for all

school districts in all states would focus on the fact that there was a "federal mandate" rather than on the enhancement of understanding that the test information might provide.

Participant **John Stevens**, Executive Director of the Texas Business and Education Coalition and a new NAGB Board member, said that Texas had found NAEP very valuable in validating the results on its own state tests at a time when higher student performance on those tests had been questioned by academic experts. The fact that Texas students also improved their performance on NAEP provided independent verification of progress in student achievement.

Session III

What International Assessment Issues Must NAEP Address in the Future?

What are the implications, if any, of an increasingly global economy and international competitiveness for educational assessment and NAEP?

Looking to the future, are international comparisons of student achievement feasible and desirable? How and why?

Should NAEP play a role in making these comparisons?

Should students in other nations be invited to take NAEP assessments?

Are comparisons made by linking different tests—such as NAEP and TIMSS—satisfactory, or is it necessary to use a common exam?

Moderator: John Merrow, Host, The Merrow Report (PBS/NPR)

Panelists:

Albert Beaton, Professor of Education and Director, Center for the Study of Testing, Evaluation and Educational Policy, Boston College and Director of TIMSS

Pascal D. Forgione, Jr., Commissioner, National Center for Education Statistics

Iris Rotberg, *Research Professor of Education Policy, Graduate School of Education and Human Development, George Washington University*

Harold Stevenson, *Professor of Psychology, University of Michigan*

The session on international assessment unfolded more as a debate than a series of opinions. Panelist **Albert Beaton**, noted that while NAEP had never been authorized to do international assessment, "the TIMSS looks a lot like NAEP and produces credible results."

In contrast, **Iris Rotberg** disputed the credibility of international comparisons on the ground that there are too many uncontrolled variables. She charged that TIMSS included in its rankings many countries that failed to meet the sampling standards established by the organizing committee of TIMSS. She noted that the average age of students taking the advanced level test ranged from 17 to 21, and their grade levels ranged from 10 to 14. Across countries, the percentage of the age cohort taking the test ranged from 2 percent to 75 percent. She listed a number of other variables in the populations taking the test, most notably, poverty. The gap between rich and poor is especially problematic in the case of the U.S. because it is wider than in other rich countries, and that affects our rankings, she said. "We really don't know what the results and rankings mean."

Pascal Forgione countered that TIMSS was really three studies, not one. In addition to the advanced level test that Rotberg criticized, there is also a fourth grade (9 yrs old) study and an eighth grade (13 years old) study, and "I have not heard one criticism of those two studies." The supposed advantage of countries whose students stay in school longer did not turn out to be an advantage after all; they didn't do better on the tests, he noted. Also, the view that comparisons are unfair because the U.S. tests its masses while other nations test only their elite (which some critics proffer as an explanation of the poor performance of our 13 year-olds in mathematics relative to other nations) doesn't hold up under scrutiny. "America's elite students are far behind other

nation's elite students, and that discovery points to the real value of international assessments." "TIMSS has provoked America into looking at curriculum, textbooks, teaching, and instruction. We're beyond the horse race—who's better than whom. We're about what are the factors that make a difference."

Professor Harold Stevenson said, "If you're going to ask general questions about the status of the U.S., it is very clear: We're not doing as well as we should and we have room for improvement. That's all we need to know."

Moderator John Merrow prodded the panel to explore another contentious issue: whether or how NAEP should be connected to international assessments.

Professor Stevenson believes that NAEP should not attempt to influence TIMSS, or vice versa. "The NAEP is constructed as an American test to evaluate the status of American children across time. The TIMSS is purposely constructed by an international board, trying to make a fair test that will give us information across countries."

Commissioner Forogione, however, took the view that NAEP and international studies should be linked, at a minimum, and eventually aligned. "There is a lot of overlap already." He believes that the discrepancies between NAEP scores and international assessments help us to raise expectations for what American students should know and be able to do, and to benchmark the productivity of American education relative to that of other nations.

Professor Stevenson spoke to the value of the studies that accompanied TIMSS tests: video-tapes of teachers in three nations, case studies, and data about variations in schools and at home. "All these things go together to give you helpful information." **Dr. Rotberg** agreed that case studies are helpful, but when rankings are reported, "people focus only on the test scores, and policymakers leap to conclusions: for example, more testing, more calculus, more specialized schools, the same curriculum for everyone, vouchers, or doing away with high schools altogether...Any one of these policies can be examined on its merits, but not on the basis of flawed test scores."

Commissioner Forogione said that what you need from TIMSS is to be able to say, for example, that in geometry, Oregon is above the international average, but it is weak

in parallelograms and congruent triangles. That's why you take the TIMSS test--to get curriculum information--to be able to look deeply at what you're teaching, when you're teaching it. It's not a scoreboard; it's about content.

Session IV

What Classroom Instruction Issues Must NAEP Address in the Future?

How can NAEP frameworks, achievement levels, and results be made more useful to teachers and schools?

Can NAEP be useful for revising curricula and instructional methods? If so, how?

What aspects of NAEP are important for communities, parents and teachers to understand—and why?

***Moderator: Warren Simmons, Director, Annenberg Institute for School Reform,
Brown University***

Panelists:

***Phyllis Aldrich, Former NAGB member and Coordinator for Gifted Education,
Washington, Saratoga, Warren, Hamilton, and Essex Board of Cooperative
Educational Services (New York)***

***James E. Ellingson, Former NAGB member and Retired Teacher, grades 4,5, and
6 (Moorhead, MN)***

***Christine Johnson, Former NAGB member and Vice President of
Educational Services, Colorado Community College and Occupational
Education System***

***Grant Wiggins, President, Center on Learning, Assessment and School
Structure (CLASS)***

Christine Johnson addressed the value of NAEP frameworks (as opposed to the tests themselves) and of NAEP achievement levels (basic, proficient, and advanced) to classroom instruction. The NAEP frameworks establish "what constitutes...quality in the areas that NAEP tests." The achievement levels "help to define our society's

expectations for young people." She believes that NAEP will not realize its full potential until students in teacher education schools, curriculum leaders in school districts, and teachers study NAEP curriculum frameworks and learn how to use NAEP achievement levels to judge student performance.

At this point, **Moderator Warren Simmons** reviewed the components of NAEP: (1) a set of frameworks that define objectives or standards that NAEP exams will attend to, (2) a set of examinations, and (3) descriptions and examples of characteristic variations in the levels of achievement. The achievement levels are determined by teachers, administrators, psychometric experts, and national experts on the subject matter, who establish both the content and the proficiency levels.

Participant **Elton Jolly**, a former NAGB member, asked whether there is a correlation between the curricula being taught and the assessment instrument itself. How can you have an instrument measuring effectiveness if you don't know the extent to which the information has been imparted?

James Ellingson answered Jolly's question by showing the interplay between state standards and assessment and NAEP standards and assessments. "NAGB can't be all things to all people who are interested in assessment and assessment results. It must stick to its knitting--which is to report what students know and can do, to monitor that over time and through the states and the nation, and to stimulate important discussion and debate throughout the nation."

As an example, Ellingson recalled a workshop he did with K-4 teachers on the writing process. The teachers noticed that they were teaching only narrative writing, while the NAEP framework said that students should be expected by 4th grade to do persuasive writing as well. "The teachers turned around and have made some changes in curriculum based on that," he said.

Moderator Simmons then asked **Phyllis Aldrich** to answer the question: Why can't state and local assessments serve the same kind of purposes that Ellingson just mentioned? What value is added by a national assessment?

Phyllis Aldrich said that what is good about NAEP data is that "it gives you a definition of what 'good' is across the states. I say it is critical for us to be able to tap into other people's definition of the good because if there is one central problem I see, it's flab. We have flab in the time that kids and teachers spend together on real learning, and we have flab in our definition of rigor. I have used NAEP as a kind of North Star."

Christine Johnson answered the "value added" question as well. "NAEP is truly a high quality, highly secured exam, and no other test has that kind of quality and depth. Only NAEP gives you longitudinal information about student achievement."

James Ellingson also responded, "I really think that the (NAEP) achievement levels are one means of avoiding what has often been called the Lake Woebegon Effect, where everybody thinks his village is above average."

Moderator Simmons asked how NAEP can become an instrument to improve practice. The final panelist, **Grant Wiggins**, addressed that question. "We are getting better and better at measuring things but we are not a whit closer to understanding how to use results to improve performance." He emphasized that teachers are not trained to interpret test results, and until they are, tests are of little use, whether the test is NAEP, a state test, or a commercial test. "Professional development has been construed as learning things from other people, but it should be construed as the learning that takes place from analyzing student performance and making adjustments on the basis of it."

Panelists disagreed on whether NAEP should expand its reporting. For **Wiggins**, "the expansion of NAEP would be a disaster because the value of NAEP is as a model, and you can't get high quality assessment if you go to district level assessment." For **Phyllis Aldrich**, the expansion of NAEP, at least to the district and school level, is necessary. "If there is no relevance to the parents, parents tune out. I think we have avoided legitimate comparisons and standards in our parochialism about local control."

Session V

What Should NAEP and NAGB Do to Meet Future Challenges?

How should NAEP and NAGB change to meet future challenges and maintain strong bipartisan support? What should not be changed?

What should be done-and by whom-to continue to ensure the integrity and independence of NAEP and NAGB? What are the most important issues that Congress and the Administration should address in the reauthorization of NAEP and NAGB? Moderator: Mary Blanton, Former Vice-chair of NAGB and attorney

Panelists:

Wilhelmina R. Delco, Former Vice-chair of NAGB and Adjunct Professor, Community College Leadership Program, University of Texas at Austin

Chester E Finn, Jr., Former Chairman of NAGB and President, Thomas B. Fordham Foundation

John F. Jennings, Director, Center on Education Policy

Maris Vinovskis, Professor of History, University of Michigan

In her opening remarks, **Moderator Mary Blanton** set the stage for a panel discussion of issues that might arise during the impending re-authorization of NAGB.

Panelist **Wilhelmina Delco**, took up the contentious issue of the proposed Voluntary National Test. She argued strongly that there are negative effects on students when tests are used punitively, rather than as diagnostic tools to move students along. She thinks that the proposed Voluntary National Test would be a mistake in the current era, when states are using tests more to punish students, teachers, and principals than to improve teaching and learning. She questioned the use of standards in districts that "don't even have any hope, any funds, any personnel to deal with bringing those standards up."

Jack Jennings believes that testing helps schools focus on instruction. But he is concerned about NAEP's performance levels, which were "set ambitiously as goals," but which newspapers translate to say that "kids can't read." He cited the criticisms of the

performance levels by the General Accounting Office and by the National Academy of Sciences which says that the performance-setting process is "fundamentally flawed." Jennings charged that NAEP officials have dismissed the critics by saying they are "against NAEP."

He also believes it is a mistake to downgrade NAEP's long term assessment schedule from every two years to every four years. Despite these problems, Jennings still supports the re-authorization of NAGB and NAEP, and hopes that Congress will clarify NAGB's responsibilities and its relationship to the National Center for Education Statistics, which administers the NAEP program, in order to reduce overlap and conflict between the two organizations.

Chester Finn spoke to NAEP's value as an external audit of student performance in core subjects on a regular, predictable basis, reported according to standards. "Anything else that can be built on top of that would be nice, probably welcome, but it should be voluntary," said Finn. He thinks NAEP tests should be made available to school districts and states that want to use them, and to charter schools as they attempt to prove themselves. "The one thing it's really good for is not, unfortunately, reforming American education," and he urged the audience to be open to an entirely new formulation of NAEP.

Maris Vinovskis, Professor of History at the University of Michigan and the author of a history of the National Assessment and NAGB prepared for the conference, believes that poverty is "one of the real tragedies of American society." He noted that there was more public interest in the problem in the 1960s than there is today, and that NAEP is not helping us understand the role of poverty in education. "If poverty is an important variable," said Vinovskis, then we need to collect better data and use multi-variant analyses. He also made a strong plea for better civics and history education in the next decade which would reflect the growing diversity of the population, and urged NAEP to test those subjects more frequently and over time. Finally, Vinovskis urged NAEP to concentrate its energies on education reform by collecting better data on at-risk children in federal programs. He suggested that the monies spent on moving NAEP to the school

district level might be better spent on tracking Title I and Head Start over time, because those programs don't seem to be having much effect.

Session VI

What Role Should Education Standards Play in Our Country's Future?

What is the long-term value—and cost—of having clear and tough education standards nation-wide?

Who should set education standards?

What one or two issues are critical for NAEP and NAGB to address effectively in the next ten years? Why? How?

The final session began with videotaped presentations by **Secretary of Education Richard Riley**, and two former Secretaries of Education, **Lamar Alexander** and **William Bennett**.

Secretary of Education Riley observed that NAGB, although largely unnoticed, is "making a powerful difference in shaping and improving American education." He recalled that five years ago, "opponents questioned the idea of saying we should use standards for all students. Yet today, a new American consensus has developed around the idea of higher standards. Forty-nine states are raising their standards and Iowa's local commitment to challenging standards remains very strong... This effort commits us to ending our nation's two-tiered system of high standards for some and watered down curriculum for too many others."

"The American people," said Riley, "have come to the conclusion that smaller classes, better trained teachers, a solemn focus on mastering the basics and a core academic curriculum are all part of this new consensus that is definitely based on high standards for all. That's why the development of quality assessments by NAEP is so important. They tell the American people where we are and what we need to do to get where we want to go."

Riley gave his support to NAGB's efforts to develop performance standards linked to a strong assessment system. "However," he said, "I also strongly urge NAGB to actively continue its research to make sure that these performance standards are practical and attainable."

"NAEP provides a valuable service to the American people because of its independence and its integrity" said Riley, "...and gives the American people a reliable, national report card on the progress we're making in our national effort to improve education. I believe that NAGB should remain independent and bipartisan and should continue to foster this growing American consensus around high standards."

Secretary Lamar Alexander believes that we have a political problem, rather than an educational problem, in the United States. "We know exactly what to do; we're just not doing it." The three things that will create an environment in which education improvements can occur are, first and foremost, standards. Second on Alexander's list is getting rid of overhead—and "the union rules and government regulations that suffocate the schools." Third is to give parents more choices.

Alexander thinks that the value of NAEP (the nation's report card) is to answer everybody's questions about what kids should know. NAEP scores told us in 1996 that about 24 percent of eighth graders...were proficient in math and about 38 percent were below basic, "which means that they are basically incompetent in math." It is good to know that about the whole country, and about the whole state, he said. Although Alexander has reservations about the proposed Voluntary National Test, he believes that NAEP data should be available to "your own school district."

Over the next ten years, "it will become more and more important that policy be set by as independent and bipartisan a Board as possible," said Alexander, and that NAEP "should continue to have high standards, and continue to stay free of the interest groups and the political agendas."

Secretary Bennett said that NAEP "is probably one of the most important federal education responsibilities." He said that "a lot of state standards, as we have seen, are very inefficient, very unsatisfactory. We need a way to answer the question: 'How are

the nation's children doing educationally?' Now immediately one has to say this has to be insulated from political pressure, because a lot of government is subject to partisan political pressures."

"The Clinton national test idea burned and failed for good reason," he said. "People didn't think that this way of approaching things would be insulated from political pressure. We do not want the Department of Education doing this...If we let the thing become so politically partisan, then it's no use to anybody anymore and it's just a massive politically correct agenda." Nevertheless, Bennett affirmed that NAEP "is the best thing we've got." Bennett noted that NAEP needs "the wherewithal to develop the instruments we need. That means it needs money." Finally, he said, "we need the capacity to help states and local school districts, individual schools even, make the best use of it."

After the secretaries' messages, participants heard a final panel discussion.

Moderator: Mark Musick, *Chairman of NAGB and President, Southern Regional Education Board*

Panelists:

Michael Nettles, *Vice-Chairman of NAGB and Professor of Education and Public Policy, University of Michigan*

Mary Blanton, *Former Vice-Chair of NAGB and attorney*

Michael Cohen, *Special Assistant to the President for Education Policy*

James J. Renier, *Chairman, IEL Board of Directors, and Former CEO of the Honeywell Corporation*

Moderator Mark Musick issued a challenge to the panel: "If everybody's for high standards, then what's the problem?"

Michael Nettles observed that "when it comes to the setting of achievement levels, people want to believe we are above whatever the standards are." **Mary Blanton** said, "When kids do better on state tests than on NAEP tests, people want to know who has decided what proficient is. NAGB hasn't done a good enough job of publicizing

how the standards are set and who sets them."

Moderator Musick asked **Michael Cohen** to explain why the President wants to use NAEP achievement levels with his proposed Voluntary National Test. "NAEP is a credible assessment," he said. "Forty plus states use it. In some states, 35 percent of the kids were meeting NAEP's basic level, and 85 percent of the kids were proficient according to the state's standards. But in another state, the percentage of kids meeting the NAEP standard and the state standard were similar. If we were going to worry about who's setting standards right, I'd look at the states that depart dramatically from NAEP. The problem is convincing parents and teachers that standards are going to be used to lift kids up, not put them down."

James Renier, affirmed the strong support in the business community for standards. When he left Honeywell and spent some time as principal of a public middle school, "I found out this is a much more complicated thing. I'm a conservative, and my peers frankly constantly underestimate the complexity of this problem and try to solve this by talking about ideological principles which are just wonderful, but they don't solve the problem. We've got to understand what prevents various organizations from achieving the standards and go to work on that."

Concluding Thoughts About NAEP's Future

All who spoke at the conference support the continued existence of NAEP and see it as the most credible, objective measure of how American students are achieving in the core academic areas over time. Also, many said that Congress should clarify the respective roles and responsibilities of NAGB and the National Center for Education Statistics in order to reduce duplication and conflict. Moreover, a number of speakers stressed the importance of insulating NAGB from the education agendas of whatever party controls the national administration, and of assuring its integrity and professionalism. No one demurred.

There also appeared to be a nearly unanimous consensus that few state education

policymakers, and even fewer school district leaders and classroom teachers, are using NAEP results---to the extent that the data permit---to leverage improvements in the academic performance of their students. Some presenters argued that NAGB, or some other entity, should interpret NAEP results in ways that would help state and local policy makers and educators make better decisions about curriculum and instruction. Other argued that NAGB, or some other entity, should train state and local educators to interpret NAEP data for themselves.

A prominent division of opinion concerned whether NAEP should remain a fairly inconspicuous, periodic report on the nation's and states' educational progress, or should expand its operations in various ways: testing more subjects; administering tests more frequently; expanding its reporting to the school district, school, and individual level; and aligning NAEP tests more closely with various international educational assessments.

Those who favor NAEP's expansion to school districts and schools argued that the value of an independent, external audit of student performance, based on high standards, will be even greater when NAEP tests become relevant to parents, students, teachers, and school district leaders. Implicitly, they want NAEP to become an instrument of reform--a lever for raising expectations for what children can achieve and for improving student performance.

Those opposing the expansion of NAEP did so on two differing grounds. One camp argued that expansion might excite fears of federal meddling in local schools and thus harm the existing program. Another group said that expansion to districts and schools would compromise NAEP's high quality as a national survey and its value as a model to the states and localities.

Although the potential political costs of expansion were alluded to by several speakers, there were only brief mentions of the technical, organizational, and financial consequences of shifting from samples sufficient to make good generalizations about the nation and whole states, to the many additional samples required to report on school districts and schools, to the massive operation needed to make the test available to individual students throughout the country, even in just two subjects and grades as

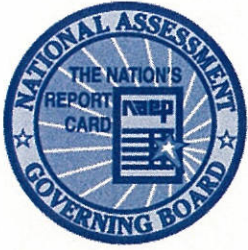
President Clinton proposed.

One other suggestion was that Congress should explore the feasibility of collecting more data aimed at defining the factors that enable poor children to perform well in school.

Whether NAGB's achievement levels are set too high or just right, whether very high standards stretch the schools or defeat them, and whether the level-setting process or its critics are "fundamentally flawed," are topics worthy of more sustained attention than was possible in a one-day conference. On these issues, a jousting contest among psychometricians will not suffice because the issues touch the highest aspirations and the deepest fears of parents, teachers, and school administrators across the nation.

Finally, it is clear that even many people closely involved in education do not understand the separate and interdependent functions of state tests and the National Assessment. NAGB must craft better explanations of why it is valuable to know how students perform on a test that represents a national consensus on what students should know and be able to do, and also to know how students perform on state tests keyed to a particular state's curriculum. NAGB must tell the public, simply and clearly, what the revealed discrepancies between NAEP and state test results signify, and why examining those discrepancies provides essential information for serious educational reform in the United States.

Harriet Tyson is a free-lance education writer and consultant.



National Assessment Governing Board

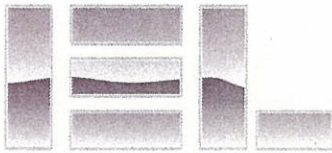
800 North Capitol Street, NW
Suite 825

Washington, DC 20002

Phone: (202) 357-6938

Fax: (202) 357-6945

Web site: www.nagb.org



Institute for Educational Leadership

1001 Connecticut Avenue, NW
Suite 310

Washington, DC 20036

Phone: (202) 822-8405

Fax: (202) 872-4050

E-mail: iel@iel.org

Web site: www.iel.org

The National Assessment Governing Board (NAGB) is an independent, bipartisan board setting policy for the Nation's Report Card.

The Institute for Educational Leadership helps people, organizations, and institutions improve educational results and the well-being of children and youth through leadership, policy, and bridging strategies. The IEL Policy Exchange improves policies and programs for children, families, and communities by connecting policies and policymakers across systems, levels of government, and ideologies.