

**CONTACT PERSON FOR MORE INFORMATION:**

Jean A. Webb, 202-418-5100.

**Jean A. Webb,***Secretary of the Commission.*

[FR Doc. 99-12869 Filed 5-8-99; 1:22 p.m.]

BILLING CODE 6351-01-M

**COMMODITY FUTURES TRADING COMMISSION****Sunshine Act Meeting****AGENCY HOLDING THE MEETING:**

Commodity Futures Trading Commission.

**TIME AND DATE:** 11:00 a.m. Friday, June 25, 1999.**PLACE:** 1155 21st St., NW., Washington, D.C., 9th Floor Conference.**STATUS:** Closed.**MATTERS TO BE CONSIDERED:** Surveillance Matters.**CONTACT PERSON FOR MORE INFORMATION:**

Jean A. Webb, 202-418-5100.

**Jean A. Webb,***Secretary of the Commission.*

[FR Doc. 99-12870 Filed 5-18-99; 1:22 pm]

BILLING CODE 6351-01-M

**COMMODITY FUTURES TRADING COMMISSION****Sunshine Act Meeting****TIME AND DATE:** 2:00 p.m., Monday, June 28, 1999.**PLACE:** 1155 21st St., N.W., Washington, D.C., 9th Floor Conference Room.**STATUS:** Closed.**MATTERS TO BE CONSIDERED:**

Adjudicatory Matters.

**CONTACT PERSON FOR MORE INFORMATION:**

Jean A. Webb, 202-418-5100.

**Jean A. Webb,***Secretary of the Commission.*

[FR Doc. 99-12871 Filed 5-18-99; 1:22 pm]

BILLING CODE 6351-01-M

**DEPARTMENT OF EDUCATION****National Assessment Governing Board****AGENCY:** National Assessment Governing Board; Department of Education.**ACTION:** Notice of request for comments.**SUMMARY:** The National Assessment Governing Board requests public comment on two draft documents it has prepared for submission to Congress and the President. The first document, required under section 305(c)(1) of the FY 1999 Omnibus Budget Act (the Act),

provides a suggested statement of the purpose, intended use, definition of the term "voluntary," and the means of reporting results for the proposed voluntary national tests in 4th grade reading and 8th grade mathematics. The second document, entitled "National Assessment of Educational Progress: Design 2000-2010," describes how improvements in the National Assessment of Educational Progress will be implemented during the 2000-2010 period. Interested individuals and organizations are invited to provide written comments to the Governing Board.

**Written Comments:** Written comments must be received by June 9, 1999 at the following address: Mark D. Musick, Chairman (Attention: Ray Fields), National Assessment Governing Board, 800 North Capitol Street NW, Suite 825, Washington, DC 20002-4233.

Written comments also may be submitted electronically by sending electronic mail (e-mail) to Ray\_Fields@ED.GOV by June 9, 1999. Comments sent by e-mail must be submitted as an ASCII file avoiding the use of special characters and any form of encryption. Inclusion in the public record cannot be guaranteed for written statements, whether sent by mail or electronically, received after June 9, 1999.

**Public Record:** A record of comments received in response to this notice will be available for inspection from 8 a.m. to 4:30 p.m., Monday through Friday, excluding legal holidays, in Suite 825, 800 North Capitol Street, NW., Washington, DC, 20002.

**The Voluntary National Test: Purpose, Intended Use, Definition of Voluntary and Reporting***Background**Purpose*

The purpose of this report is to fulfill one of the requirements of the FY 1999 appropriation act for the Department of Education (the Act). Specifically, with respect to the proposed voluntary national tests in 4th grade reading and 8th grade mathematics, the Act requires the National Assessment Governing Board to

\* \* \* determine and clearly articulate the purpose and intended use of any proposed federally sponsored national test. Such report shall also include—(A) a definition of the meaning of the term "voluntary" in regards to the administration of any national test; and (B) a description of the achievement levels and reporting methods to be used in grading any national test.

This report addresses the four required areas: purpose, intended use,

definition of "voluntary," and reporting. Although the legislation states that the Governing Board shall "determine" these matters, the Governing Board recognizes that this report is advisory to Congress and the President. Any final determination on these matters will be made in legislation enacted by Congress and signed by the President.

The Act contains other provisions related to the voluntary national test. One provision amends the General Education Provisions Act, creating a new section 447, prohibiting pilot testing and field testing of any federally sponsored national test unless specifically authorized in enacted legislation. However, another provision permits the development of voluntary national tests, giving the National Assessment Governing Board exclusive authority for such test development.

In order to carry out the congressional assignment to prepare this report, the Governing Board had to envision a situation in which there was authority to conduct voluntary national tests, while recognizing that the Act prohibits such tests at this time. Further, the Governing Board had to envision how national testing could work, given that schools in the United States are governed by states, localities and non-public authorities. The Governing Board attempted to answer the question: If there are to be voluntary national tests, what is a feasible, coherent plan that would be beneficial to parents, students, and teachers? Thus, while not advocating for or against the voluntary national test initiative, the Governing Board interprets the congressional assignment to be to present a sound and logical case for the potential purpose and use of the voluntary national tests.

The Act sets September 30, 1999 as the deadline for submitting this report to Congress and the President. However, to assist Congress and the President in deliberating on the future of the voluntary national test, help promote a timely decision, and avoid a full year's delay in pilot testing should Congress and the President decide to proceed with the project, the Governing Board is submitting its report in June.

*Report Preparation Process*

In November 1998, the Governing Board established a special ad hoc committee to assist in drafting the report. The committee was composed of both veteran and new Board members. Chaired by Michael Nettles, the committee included Wilmer Cody, Thomas Fisher, Michael Guerra, Nancy Kopp, Debra Paulson, Diane Ravitch, and John Stevens.

The committee developed a plan for preparing the report, engaging the Governing Board in related policy deliberations, and obtaining public comment. At the March 1999 Board meeting, the committee presented materials that were developed for public comment. These included an explanatory statement; two possible scenarios addressing purpose, use, definition of voluntary, and the methods for reporting; and a set of questions related to the scenarios. The purpose of these materials was to provide a framework for public comment. They did not represent the positions of the Governing Board at the time.

The Governing Board discussed these materials at length, made several changes, and authorized the committee to proceed to obtain public comment. The materials and an invitation to provide written comments and/or oral testimony at four public hearings during March and April were disseminated.

Taking the comments received into account, the committee then prepared a draft report for review at the May 1999 Governing Board meeting. The Governing Board discussed and revised the draft report and authorized the committee to obtain comment on the draft report. The draft report was disseminated by mail, on the Governing Board's web site, and in the **Federal Register**. A hearing on the draft report will be conducted June 12 at the annual Large Scale Assessment Conference with state and district testing experts.

After taking the comments received into account, the committee will prepare a draft report for presentation to the Board at a special meeting on June 23. At the June 23 meeting, the Governing Board will discuss the draft and approve a final version for submission to the President and Congress.

#### *Overview*

This report is in three sections. The first section is in the form of a story. It is intended to put a "human face" on the details in the section that follows. The second section describes the Governing Board's recommendations on purpose, intended use, definition of "voluntary," and reporting for the proposed voluntary national tests. The third section is a summary with recommendations.

#### **The Voluntary National Test: A Story**

It is March 18; the year is 2006. Fourth grader Maria Johnson, along with her classmates and many other 4th graders across the nation, will be taking the voluntary national test in reading

tomorrow. Eighth graders will be taking the mathematics test.

Maria started kindergarten in September 2001; the first voluntary national test was administered the following March. That year and each year since, *Parade* magazine devoted an early April article to the test. The test questions were published, along with the answers. For questions that require students to write their own answers, samples of student work from the national tryout of the test the year before were included to illustrate different levels of student performance. These levels of student performance are based on the achievement levels set for the National Assessment of Educational Progress (NAEP). Similar materials were made available following each year's tests in newspapers, magazines aimed at parents and teachers, on the Internet, and on the Public Broadcasting System. Reading and mathematics achievement levels posters are displayed in pediatrician's offices across the country. January through March of each year, McDonald's, Burger King, Wendy's, and KFC print sample test questions on placemats and food containers.

Maria's school district decided to volunteer to participate in the national test in 4th grade reading. The school district administration had examined the test framework, specifications, and sample test and determined that they were consistent with the district's reading program. They knew that the results would belong to the district and the families. The federal government would not report or maintain any of the data resulting from testing nor require the district to report any of the data to the federal government.

Maria's school provided copies of the *Parade* article to each of the families. In the school district, the policy is for all students to participate in testing unless a parent specifically objects. When Maria's parents finished reading the article, they had a clear picture of what a proficient reader in the fourth grade should know and be able to do. They understood that proficiency would not come overnight, but with many small steps and that each year of school would mark progress toward the goal of reading proficiency. Maria's parents decided that having a clear goal and following progress toward that goal are good things to do and wanted their child to participate.

Having this initial knowledge, the Johnsons wanted to learn more and did their homework. They attended a school-sponsored seminar on the reading program. They learned what they could do at home to reinforce what Maria was learning in school. The

Johnsons obtained a special version of the NAEP framework, written for parents, to deepen their understanding of the material covered by the test. The Johnsons now had a frame of reference for talking with Maria's teachers in specific terms about the reading program and for monitoring Maria's progress each year toward 4th grade reading proficiency. Maria, with her parents' encouragement and teachers' support, has worked hard in school and at home on her reading assignments and enjoys reading on her own.

With this shared understanding and common language about reading proficiency, the school was helped in its efforts to involve parents. The school had developed its own testing program to track the reading progress of each student each year toward 4th grade reading proficiency. Thus, needs for extra help were identified early, in-depth diagnosis was provided when needed, and remediation occurred before it was too late.

The school liked using the achievement levels. They were consistent with the state's performance standards for reading. They helped keep the school staff focused as they worked day-by-day, making hundreds of decisions about materials, instruction, and curricula to achieve the many incremental steps needed for each student to progress.

Parents and teachers also like the fact that the test booklet is returned. This permits parents and teachers to review with the student all of the test questions and the student's answers. The student gets reinforcement on what was done well. Parents and teachers can see which questions were answered well and which were missed, probe the reasons why with the student, and, from the student's response and other knowledge of the student, explore whether advanced activities, diagnostic testing, or any other intervention should be considered.

Together with the on-going assessment program and the state's standards and assessments, the school and parents found that the voluntary national test adds in a unique way to the range of methods for monitoring individual student progress. The teachers and principals found that the achievement levels used to report voluntary national test results were much easier for parents to understand than percentiles, stanines, or mean scores. Also, the voluntary national test provides parents and schools a single basis of comparison for individual student performance across states that is generally not available from classroom developed tests or state-wide

assessments. Most of all, parents have a clear and very specific understanding of how their child has performed in comparison to rigorous standards.

Although the test was designed to provide individual results, the school district has decided that it will compile the individual student results that were provided by the voluntary national testing program. The district administrators want to know how the district overall compares with the students in the national sample who participated in the national trial run of the test the year before.

The district has joined a consortium of similar districts that have agreed among themselves to follow the guidelines for compiling and reporting voluntary national test data developed by the National Assessment Governing Board (NAGB). Following these guidelines ensures that the data analyses are done properly, comparisons between and among districts and schools are fair, and inferences about achievement are defensible. When the district reports these results to the public, it makes a big point of saying that it has followed these guidelines to the letter and spirit, as a means for establishing credibility and trust.

The story presents one plausible scenario for how the voluntary national test might be implemented in public schools, but other scenarios are possible as well. The story is focused on the future because effects of the proposed voluntary national test would not be fully achieved in its first year. But two things are clear. If there is to be such a test, it should be made available to all who would find value in it, whether state, public school, private school, home school, or individual parent. And, while the federal government would provide resources to make the tests available, there should be no federal coercion, sanctions, or rewards for participating.

The story emphasizes that, while having widely recognized standards and assessments can provide focus for planning and a common language for students, parents and teachers, what is most important is what parents, students, and educators actually do with that knowledge. The story, implicitly, also suggests that a wide voluntary mobilization of private resources in society reinforcing the value and importance of learning (e.g., Parade and McDonald's) would be important.

#### **The Purpose of the Voluntary National Test**

As the Governing Board worked on this report, it became evidence that

purpose, intended use, the definition of voluntary, and means for reporting are, to a large degree, interdependent. A change in any one of these could affect the others. Therefore, it is important that these four areas be coherent.

In addition, the test should serve a unique purpose. If the same purpose is already being fulfilled by another testing program, there is no need for the voluntary national test. If the same purpose could easily be fulfilled by another testing program, it would be prudent to consider that possibility in weighing the pros and cons before proceeding with full implementation.

The National Assessment Governing Board suggests that Congress and the President consider the following as the purpose of the proposed voluntary national test:

To measure individual student achievement in 4th grade reading and 8th grade mathematics, based on the content and rigorous performance standards of the National Assessment of Educational Progress (NAEP), as set by the National Assessment Governing Board (NAGB).

#### **Rationale**

The legislation giving responsibility for voluntary national test development to the Governing Board does not specify or limit the subjects and grades to be tested. However, the accompanying conference report does direct that the tests be based on NAEP content and NAEP performance standards and be linked to NAEP to the maximum extent possible. The Governing Board in August 1996 adopted a policy on NAEP redesign. The redesign policy provides for testing at grades 4, 8, and 12 at the national level is 11 subjects and, based on the needs and interests expressed by states, at grades 4 and 8 at the state level in reading, writing, mathematics and science.

Grades 4, 8, and 12 are transition points in American schooling. Consistent with the National Assessment redesign policy and the congressional directive that the voluntary national tests be designed to parallel NAEP, the Governing Board limited the test development contract to cover grade 4 reading and grade 8 mathematics. Proficiency in these subjects, by these grades, is considered to be fundamental to academic success.

Most importantly, measuring individual student achievement based on the National Assessment affords this proposed testing program a unique niche among K-12 academic testing programs in the United States. For 30 years, the National Assessment has reported the status and progress of student achievement on nationally

representative samples of students. It has done so with credibility, technical competence, and widespread acceptance. For the last ten years, the National Assessment also has reported on state-representative samples of students in volunteering states, providing participating states with the only available comparable measure of student achievement.

However, the National Assessment, by law, does not provide individual student results. It provides only group-level results (e.g., for students overall, by sex, by race, by type of school, etc.). The NAEP state-level assessments represented a watershed event. Ten years ago, state-level assessments were begun with fears of encroachment on state and local autonomy and worry that a national curriculum would result. The promise that the NAEP state-level assessment program would serve a unique function—to provide comparable state results, trends over time, and an external validity check for state standards and assessments—has been realized. The fears have not. This is because there are checks and balances built into the governance of the program.

Today, similar fears of federal encroachment and the emergence of a national curriculum are being expressed about the voluntary national test and must be addressed. As with the NAEP state assessments, checks and balances can be provided for in the governance and operation of the voluntary national testing program to prevent these reasonable concerns about federal encroachment and national curricula from becoming reality.

#### **Definition of the Term 'Voluntary'**

There are two dimensions to the definition of the term "voluntary" as it would apply in the administration of the voluntary national tests. The first dimension has to do with the role of the federal government. The second dimension has to do with who makes the decision to participate in the voluntary national tests.

#### *Federal Role*

The role of the federal government in the proposed voluntary national tests should be limited. The federal government should not make any individual take the voluntary national tests or require any school to administer the tests. The federal government should have no control or authority over any data resulting from the administration of the voluntary national tests, nor should participation in the voluntary national tests be a condition for receiving federal funds.

The National Assessment Governing Board suggests that Congress and the President consider the following as part of the definition for the term "voluntary":

The federal government shall not require participation by any state, school district, public or private school, organization, or individual in voluntary national tests, make participation in voluntary national tests a specified condition for receiving federal funds, or require participants to report voluntary national test results to the federal government.

#### Rationale

It is fundamental that the definition of the term "voluntary" include limits on the role of the federal government. The limits on the federal role should be specified in legislation and designed to insure against any encroachment on state, local, and private school autonomy. Several witnesses in the Governing Board's public hearings argued that the 55 mile-per-hour speed limit was voluntary, too, but became universally implemented by states (and in that sense was "mandatory") because it was a specified condition required to receive federal highway funding. The definition of "voluntary" provided here would foreclose such an outcome. However, it would not foreclose any federal grantee from using the voluntary national test to meet a general reporting requirement if other options are available as well and could be fulfilled validly and appropriately by the voluntary national tests. On the one hand, it is not fair to require that the VNT be used. On the other hand, it is not fair to foreclose its use if doing so is done without coercion and solely at the participant's discretion.

#### Who Decides To Participate

Since the federal government will not coerce participation, it will be up to others to decide whether to participate. Education governance for public schools in the United States, about 88 percent of K-12 school enrollment, is vested in state and local public authorities. Responsibility for the remaining 12 percent of K-12 school enrollment resides with private school authorities and parents.

The definition of "voluntary" needs to accommodate a wide range and diversity of governance authority. For example, there is great variation among state laws in the degree of central authority and responsibility for education and the degree of local district autonomy. Similarly, there are differences among private schools in how they are governed as well as among state laws regarding the oversight of

private schools and home schooling. While provisions for who decides to participate should accommodate this range and diversity of authority, such accommodation must be made in a manner that does not conflict with state and local law and policy.

With respect to who decides to participate in voluntary national tests, the National Assessment Governing Board suggests that Congress and the President consider the following:

Public and private school authorities should be afforded the option to participate in the voluntary national tests. For public schools, state and/or local law and policy should determine whether the initial decision to participate is made at the state level or at the local district level. Where state law or policy provides that the initial decision be made at the state level, and the state decides not to participate, school districts should be afforded the opportunity to decide whether to participate, to the extent permitted by state and local law and policy.

For private schools, the decision to participate should be made by the appropriate governing authority.

Parents may have their children excused from testing as determined by state and local law and policy in the case of public schools. In the case of private schools, parents may have their children excused from testing as determined by the policy of the appropriate governing authority.

Parents whose schools are not participating but want their children to take the voluntary national tests should have access to the tests either through a qualified individual or testing organization before the tests are released to the public or through dissemination procedures at no or minimal cost (e.g., public libraries and the Internet) after the tests are released to the public.

#### Rationale

The definition of "voluntary" adopted by the Governing Board is intended to align with state and local law and policy regarding the authority to make decisions about testing. The definition is designed to allow for choice in providing the opportunity to participate, but without exceeding the authority of the federal government in this sensitive area, without coercion by the federal government, and without intruding on the prerogatives of states, school districts, private schools, and parents.

Typically, if not universally, determinations about testing are made by school authorities, whether state, local, or private (including home schools). They determine what should be tested, what grades should be tested, the time of year for testing, the content of reports on test results and the use of the results. These authorities decide whether tests will be taken by all students or by a sample of students. Therefore, the definition of "voluntary" is designed to account for the fact that

schools are the most likely venue through which the proposed voluntary national tests would be administered and that school authorities decide which tests will be given. At the same time, the definition of "voluntary" recognizes and accommodates the variation in responsibility and authority for education governance that exists across state boundaries among states and schools.

School authorities also decide the extent to which official policies will provide for parental intervention to have their children excused from testing. The definition of "voluntary" intends to accommodate this variability as well, again, without intruding on local prerogatives.

Finally, the definition of "voluntary" recognizes that there could be instances in which school authorities decide not to participate in the voluntary national tests, but certain parents want their children tested. In such cases, parents may elect to have their children tested by appropriately licensed or recognized individuals or organizations. Because all parents who may wish to have their children take the test may not have the resources to pay for private testing, the test and scoring guides could be made available for free, or at a minimal charge, after the period for conducting the testing is completed.

#### Intended Use of the Voluntary National Tests

The intended use of the voluntary national tests is related to the statement of purpose and definition of "voluntary" suggested above. The Governing Board suggests that Congress and the President consider the following as the intended use of the proposed voluntary national tests:

To provide information to parents, students, and authorized educators about the achievement of the individual student in relation to the content and the rigorous performance standards for the National Assessment, as set by the National Assessment Governing Board for 4th grade reading and 8th grade mathematics.

#### Rationale

The proposed intended use of the voluntary national tests is purposely narrow, and appropriately so. Consistent with the purpose statement, which is *to measure individual student achievement*, the intended use is *to provide information describing the achievement of the individual student*. Upon receiving the results of the test, parents, students and teachers will have an overall measure of the individual student's achievement in 4th grade reading or 8th grade mathematics. As

described in the following section on reporting, they will have information on the performance standard reached by the student and other detailed related information.

With information in hand from the voluntary national tests and other sources about the child and the school program, it is expected that: (1) parents could become more involved with the child's education, (2) students could study hard and learn more, (3) teachers could work more to emphasize important skills and knowledge in the subjects tested without narrowing or limiting their curricula, and (4) parents, students, and teachers could have a means for better communication about the child's achievement.

While such outcomes can be hoped for, their achievement relies on local effort, resources, skill, and persistence. A test and clear performance standards are necessary, but not sufficient conditions for their achievement. No testing program can determine, ensure, or constrain what will be done with the information it provides. However, when the values of a society at large are focused on a clear goal widely recognized as important, with consistent methods for monitoring progress toward that goal, the likelihood that local effort, resources, skill and persistence will voluntarily be brought to bear on the achievement of that goal is increased.

The Governing Board does not assume that uses of data from voluntary national tests beyond the intended use described above are necessarily inappropriate or should be prohibited to states, districts, and private schools. Any such additional use of voluntary national test data would be done at the discretion of the participating state, district, or private school authorities, who would be responsible for following appropriate technical standards and validation procedures.

However, the voluntary national test are not tied to a preferred curriculum, teaching method or approach. The voluntary national tests are based on the content of the National Assessment of Education Progress. The content of each NAEP test is developed by the Governing Board through a National consensus process involving hundreds of educators, curriculum specialists, school administrators, parents, and members of the public. The content of NAEP is designed to assess what students know and can do, not how they are taught.

The voluntary national tests also are not designed to diagnose specific learning problems or English language proficiency. Tests for such diagnostic purposes are specifically tailored. For

example, a test of English language proficiency may involve speaking and listening as well as reading. A test to diagnose specific learning problems may include motor coordination and perception, but may or may not include mathematics skills. Tests for the general population, such as the voluntary national tests, are inappropriate for these diagnostic purposes.

The voluntary national tests are not intended to be used as the sole criterion in making "high stakes" decisions (e.g., placement or promotion) about individual students. As the National Academy of Sciences/National Research Council (NAS/NRC) stated in its report "High Stakes: Testing for Tracking, Promotion, and Graduation":

Scores from large-scale assessments should never be the only sources of information used to make a promotion or retention decision \* \* \* Test scores should always be used in combination with other sources of information about student achievement \* \* \* Students who fail should have the opportunity to retake any test used in making promotion decisions; this implies that tests used in making promotion decisions should have alternate forms. (p. 12-11).

The NAS/NRC report also recommends against the use of the voluntary national test in any high stakes decision for individual students under any circumstances, whether in association with other sources of information or not. This recommendation is in contrast to the Governing Board's suggestion above that any use of the voluntary national test beyond the stated intended use must follow technical standards and be validated by the participating state, district, or private school authorities. The Governing Board recommends that such uses and their validation be left to the professional discretion of participating states, districts and schools.

#### **Reporting the Results of the Voluntary National Tests**

Consistent with the purpose and intended use of the voluntary national tests, the National Assessment Governing Board suggests that results of the voluntary national tests be provided separately for each student. Parents, students, and authorized educators (those with direct responsibility for the education of the student) should receive the test results report for the student. Test results for the student should be reported according to the performance standards for the National Assessment of Educational Progress (NAEP). These are the NAEP achievement levels: Basic,

Proficient, and Advanced.<sup>1</sup> All test questions, student answers, and an answer key should be returned with the test results; it will be clear which questions were answered correctly and which were not. The achievement levels should be explained and illustrated in light of the questions on the test. Also, based on the nationally representative sample of students who participated in the national tryout of the test the year before, the percent of students nationally at each achievement level should be provided with the report.

There should be no compilations of student results provided automatically by the program. The program should not provide results for the nation as a whole or by state, district, school, or classroom, since the purpose and use of the testing program are directed at individual student level results.

However, it is virtually certain that compilations of student results will be desired and demanded by at least some of the state and district participants and possibly by private school participants as well. These participants should be permitted to obtain and compile the data at their own cost, but they will bear the full responsibility for using the data in appropriate ways and for validating the uses they make of the data.

The Governing Board would develop and provide guidelines and criteria for use by states, districts, and schools for compiling and reporting the data from the voluntary national tests. The guidelines and criteria would explicitly require full and clear disclosure about exclusions and/or absences from testing, so that results and comparisons would be accurately portrayed. Access to the test data by external researchers would be made strictly at the discretion of the participating state, district, or private school, as it would with any other testing program, without prejudice because of federal support for the voluntary national test program.

#### **Other Issues**

There are several issues which the Governing Board would be remiss not to raise, although they are outside the requirements for this report set by Congress and no attempt is made to resolve them here.

<sup>1</sup> N.B. In making the determination that the achievement levels will be the basis for reporting voluntary national test results, the Governing Board is aware that Congress has asked for its response to the assertion that the process for setting the levels is "flawed." The Governing Board is submitting simultaneously, under separate cover, a report describing its response to this assertion and its plan for investigating alternative standard-setting methods.

### Implementation

By law, the Governing Board has exclusive authority for test development. The Governing Board has been meticulous in staying within the law's boundaries. The Governing Board has focused its efforts on developing test questions and on associated activities. Appropriately, the Governing Board has not taken up implementation issues such as

- The process by which states, districts and schools commit to participate, to what entity the commitment is made, and in what form and of what nature the commitment should be
- How information about the test program and the opportunity to participate will be made available to parents, teachers, and students
- Whether and how quality control monitoring of testing should occur
- How printing of test booklets, scoring of student responses, and reporting of test results would be handled
- Whether the testing program should be controlled by a federal agency or private commercial interests
- Whether all or part of the costs for the test program should be paid by the federal government

### Linking the Voluntary National Tests to NAEP

Underlying the concept of the proposed Voluntary National Tests is the desire to measure and report student achievement based on the content and rigorous performance standards of NAEP. Indeed, the directive from Congress to the Governing Board is to link the VNT to NAEP "to the maximum extent possible." Accomplishing this linkage presents a significant challenge—one which affects the design of the VNT as well as the manner in which data are calculated and reported. Two tests can be linked to the degree that they have common characteristics, including types of questions, range of content, test administration procedures, etc. Thus, the first task facing the Governing Board is to forge a close relationship between the two tests as the VNT is being created.

Linking two tests also depends upon the particular statistical approach that can be used. Unless a strong statistical procedure can be used legitimately, the VNT results cannot be reported directly on NAEP scales. This would necessarily mean that the VNT may have to be reported without direct reference to NAEP.

Solutions to the challenge of linking will evolve as (and if) work on the VNT

continues. The Governing Board intends to develop options to create a good linkage between the VNT and NAEP. If the linkage cannot be established, alternative reporting strategies for the VNT will be prepared. These alternatives would, of course, be based on NAEP content and performance standards to the maximum extent possible.

These questions of implementation and linking do not need to be settled immediately. They will, however, need to be considered and must be settled in a timely manner if Congress and the President decide that the voluntary national test program should go forward.

### Summary

This report presents the Governing Board's response to the congressional assignment to determine the purpose and intended use of the proposed voluntary national tests, including the definition of the term "voluntary" and a description of the achievement levels and other means for reporting results. The Governing Board has prepared this report over an eight month period that included extensive deliberation, expert advice, four regional public hearings and two successive periods of public comment (the first to develop the draft report, the second to review the draft report).

Although the legislation requiring the report calls for a "determination," the Governing Board views this report as advisory. Any final determination on these matters would be made in legislation enacted by Congress and signed by the President.

In submitting this report, the Governing Board is neither advocating for or against a voluntary national test. Rather, the Governing Board interprets and assignment from Congress to be a present a sound and logical case about the potential purpose and use of the voluntary national tests.

### Recommendation

The Governing Board is submitting this report in June, three months before the required due date of September 30, 1999. This is to assist the Congress and the President in deliberations toward a timely decision on the future of the voluntary national tests.

The Governing Board recommends that a decision be made before September 30. The schedule for the voluntary national test, if the decision is made to proceed, calls for a pilot test in March 2000 of test questions developed by the Governing Board. In order for the pilot test to be properly carried out in March 2000, a decision is needed before

September 30, 1999. This will permit the test development contractor to proceed in an orderly and efficient manner to carry out activities that are essential to the pilot test, such as determining the sample of participating schools and arranging for the printing of booklets of test questions.

A decision to proceed that comes too late will set the schedule for the pilot test back one year, to March 2001. This is because pilot testing must occur in the same month that testing is to occur, which is March. If authorization to proceed does not come before September 30, it may not be possible to carry out all of the necessary steps that lead up to the pilot test in time for it to occur in March 2000.

If, on the other hand, the decision is made not to proceed, a decision prior to September 30 will allow for an orderly and cost-effective termination of the test development contract.

It is important to note the purpose of pilot testing. The purpose of pilot testing is to determine the quality of each individual test question. There are no individual student scores reported. In pilot testing, individual questions are evaluated singly. There are no overall test scores calculated, even though a student in the pilot test will respond to many test questions. The only data collected are statistics that relate to the specific test question, such as the percent of students who answered the question correctly. From the analysis of student responses on the individual test questions, three decisions are possible: drop the test question, keep the test question as is, or keep the test question with changes. Only from the set of test questions that remain after pilot testing will test booklets be constructed, which then will be tried out in field-testing. The *field test* stage, unlike the pilot test, is designed to simulate the plans for actual testing. If the decision is made to proceed, a field test would be conducted in March 2001.

The optimal outcome would be to have a timely final decision on whether or not there shall be voluntary national tests. Another possible outcome would be to have agreement to proceed with the pilot test of questions, while continuing to deliberate on the prospects for the voluntary national test program itself. If the pilot test proceeds, the test questions could be considered for use in the National Assessment of Educational Progress, should the ultimate outcome be the continuing prohibition of voluntary national tests.

### National Assessment of Educational Progress: Design 2000–2010

What should the Nation's Report Card on student achievement look like during the next decade? How can it most effectively help the public understand the academic readiness of our youth at grades 4, 8, and 12—key transition points in American education? Ultimately, how can the National Assessment of Educational Progress (NAEP) best be used as an indicator of national and state educational preparedness for the challenges facing our society?

The purpose of this report to Congress and the President is to describe the recommendations of the National Assessment Governing Board for answering these questions. The report will provide a summary of the Governing Board's policy to redesign the National Assessment, describe the status of implementation of the redesign policy, and address the implications for reauthorization of the National Assessment of Educational Progress.

#### Background

In 1996, prompted by increasing demand for more and more frequent information about the status and progress of student achievement in the United States, the National Assessment Governing Board, an independent, bipartisan citizen's group created by Congress to set policy for the National Assessment, charted a course for NAEP through the year 2010. The policy to redesign the National Assessment followed two years of study, expert advice, deliberation by the Governing Board, and public comment.

In 1997, the National Center for Education Statistics (NCES) developed a plan to implement the redesign policy. The plan has two phases. The first phase covers assessments in the year 1999–2003. In 1998, NCES awarded new contracts for NAEP covering this period. During this first phase, the Governing Board's annual schedule of assessments will be carried out (see Table 1), National Assessment student achievement data will be released more quickly, National Assessment reports will be redesigned for the general public, and research will be conducted to foster a streamlined design for the National Assessment. The second phase of National Assessment redesign, covering assessments for the years 2004–2007, will continue the earlier improvements and begin to implement the innovations aimed at streamlining the design of NAEP.

Even as redesign implementation begins under the new contracts, the

Governing Board continues to weigh new evidence that may bear on the shape of the NAEP redesign policy. For example, following the adoption of the redesign policy in 1996, there have been evaluation reports issued on the National Assessment, reviews by other experts, and papers prepared for the November 1998 Ten-Year Anniversary Conference sponsored by the Governing Board. The views expressed raise issues or concerns that bear on six areas of the redesign policy. The Governing Board decided to examine once again these six areas of the redesign policy to determine whether any modifications to the policy are in order. These six policy areas were reviewed in detail in a forum conducted by the Governing Board on April 15 with technical experts, consumers of NAEP data, representatives from the National Center for Education Statistics and the NAEP contractors. The results of the April 15 forum are incorporated in this report.

#### *National Assessment Redesign: A Summary and Status Report*

##### Introduction: The Redesign Principles

Over its thirty-year history, the National Assessment has earned respect and credibility. The National Assessment is widely recognized for the comprehensiveness of its tests, the quality of its technical design, the accuracy of its reports, and innovation in its execution. The data produced by the National Assessment are unique. No other program provides regular reports on the status and progress of student achievement for our nation as a whole and that are comparable state-by-state.

Although its original purpose was to measure and report on the status of student achievement and on change over time, recognition of the quality and integrity of the National Assessment led to a multitude of demands and expectations beyond reporting on achievement. Meeting those expectations was done with good intentions and seemed right for the situation at the time. However, some additions that the National Assessment performs less effectively were "tacked on" to the original design.

The National Assessment was being asked to do too many things, some even beyond its reach to do well, and was attempting to serve too many audiences. For example, in contrast to the 1970's in which a single 120 page report on mathematics was deemed sufficient, the 1992 NAEP mathematics reports numbered seven and totaled about 1,800 pages.

The result of attempting to respond to demands beyond NAEP's central

purpose was to overburden NAEP's design, drive up costs and reduce the number of subjects that could be tested. For example, the National Assessment tested two or three subjects each year during the 1970's, its first decade, but only every other year after the 1980's. Another indicator that NAEP had too many distractions was that results could be released as many as two to three years after testing. This simply was not acceptable, particularly with the advent of state-level assessments in the 1990's.

The Governing Board's solution was to focus NAEP on what it does best: measure and report on the status of student achievement and change over time. Focusing NAEP on what it does best would permit NAEP's design to be simplified and also would mean putting limits on demands that are outside NAEP's central purpose. Another part of focusing NAEP is to define the audience for reports. The Governing Board has determined that the NAEP program should not attempt to serve multiple audiences directly. The audience for reports should be the general public.

Specialized needs for NAEP data should be accommodated by making the NAEP data easily accessible for analysis by others—educators, researchers, policymakers, and the media, among others. In order to make data more understandable and useful to the general public, the Governing Board has determined that achievement levels, or performance standards, should be the primary means for reporting NAEP results.

Thus, five principles undergird the Governing Board's policy for the redesign of the National Assessment:

- Conduct assessments annually, following a dependable schedule
- Focus NAEP on what it does best
- Define the audience for NAEP reports

- Report results using performance standards

- Simplify NAEP's technical design

Details on these and other aspects of the redesign policy follow.

#### *Annual Schedule*

A centerpiece of the National Assessment redesign is a dependable annual schedule of assessments through the year 2010 (Table 1). In the past decade, the focus on education reform, new and revised state assessments, and the national education goals have led to demand for National Assessment testing more frequently than the biennial schedule of the 1980's and most of the 1990's. The schedule for the period 1996 through 2010 was adopted in March 1997 and revised in November 1998. It provides for annual assessments

at the national level and state-level assessments in even-numbered years. The long-term trend assessments in reading, writing, mathematics, and science continue on a once per four-year cycle beginning in 1999.

At the national level, grades assessed will be 4, 8 and 12. Subjects covered will be reading, writing, mathematics, science, geography, U.S. history, world history, civics, economics, foreign

language, and the arts. These are the subjects listed in the current national education goals. Reading, writing, mathematics and science will be assessed once every four years. Other subjects will be assessed less frequently, but there will generally be two assessments in a subject over a ten-year period.

Testing at the state level will occur in even-numbered years, with reading and

writing in grades 4 and 8 alternating with mathematics and science in grades 4 and 8. Student achievement results in these subjects and grades at the state level will be reported on a once per four-year basis.

Many of the other redesign policies, described below, are aimed at making the annual schedule affordable through cost-saving efficiencies.

TABLE 1.—SCHEDULE FOR THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

[The following schedule was adopted by the National Assessment Governing Board on March 8, 1997 and revised in November 1998. Assessments shown as scheduled for 1996, 1997, and 1998 were approved previously by the Board.]

Year	National	State
1996	Mathematics Science Long-term trend* (reading, writing, mathematics, science)	Mathematics (4, 8). Science (8).
1997	Arts (8).	
1998	Reading Writing Civics..	Reading (4, 8). Writing (8).
1999	Long-term trend*.	
2000	Mathematics Science Reading (4).	Mathematics (4, 8). Science (4, 8).
2001	U.S. History. Geography.	
2002	Reading Writing Civics.	Reading (4, 8). Writing (4, 8).
2003	<b>FOREIGN LANGUAGE (12).</b> Long-term trend*.	
2004	<b>MATHEMATICS</b> Science	<b>MATHEMATICS (4, 8).</b> Science (4, 8).
2005	<b>WORLD HISTORY (12).</b> <b>ECONOMICS (12).</b>	
2006	<b>READING</b> Writing	<b>READING (4, 8).</b> Writing (4, 8).
2007	<b>ARTS.</b> Long-term trend*.	
2008	Mathematics <b>SCIENCE</b>	Mathematics (4, 8). <b>SCIENCE (4, 8).</b>
2009	<b>U.S. HISTORY.</b> <b>GEOGRAPHY.</b>	
2010	Reading <b>WRITING</b>	Reading (4, 8). <b>WRITING (4, 8).</b>

**Note:** Grades 4, 8, and 12 will be tested unless otherwise indicated. Comprehensive assessments are indicated **BOLD ALL CAPS**; standard assessments are indicated in upper and lower case.

\* Long-term trend assessments are conducted in reading, writing, mathematics and science. These assessments provide trend data as far back as 1970 and use tests developed by the National Assessment at that time.

Status of Implementation

The work in the new NAEP contracts covers the schedule as adopted by the Governing Board for the years 1999–2003. The long-term trend assessments in reading, writing, mathematics, and science will be conducted in 1999 and 2003. In 2000, mathematics and science assessments will be conducted in grades 4 and 8 at the state level and at grades 4, 8, and 12 at the national level. In addition, a reading assessment at grade 4 at the national level will be conducted. In 2001, geography and U.S. history assessments will be conducted

at grades 4, 8, and 12 at the national level. In 2002, reading and writing assessments will be conducted at the state level in grades 4 and 8 and at the national level in grades 4, 8, and 12. In 2003, assessments will be conducted at the national level in civics in grades 4, 8, and 12 and in foreign language at grade 12.

Define the Audience for NAEP Reports

The expanded demands and expectations noted above reflected the many varied audiences that NAEP was attempting to serve. Trying to serve too

many audiences has meant that no audience is optimally served by the National Assessment. The NAEP redesign policy makes the distinction between the *audience* for reports prepared by the NAEP program and the *users* of NAEP data. The audience for NAEP reports is the American public. The primary users of NAEP data are national and state policymakers, educators, and researchers.

This distinction in the policy between the audience for reports and users of data is important. It is intended to address the needs of various groups and



individuals interested in NAEP results, while providing an appropriate division of labor between them and the federal government.

National Assessment reports released by the U.S. Department of Education should be objective, providing the facts about the status and progress of student achievement. Providing objective information about student achievement is an appropriate federal role. Since the public is the primary audience, NAEP reports should be understandable, jargon free, easy to use, widely disseminated, and timely.

On the other hand, the redesign policy suggests that interpreting NAEP data (e.g., developing hypotheses about achievement from relationships between test scores and background questions) is a role that falls primarily to those outside the Department of Education—the states that participate in NAEP, policymakers, curriculum specialists, researchers, and the media, to name a few. For the NAEP program itself to address the myriad of interests and questions of these diverse groups seems both impractical and inappropriate. However, the federal government should encourage and provide funds for a wide range of individuals and organizations with varied interests and perspectives to analyze NAEP data and use the results to improve education. This is the point of the redesign policy. Thus, the redesign policy provides that National Assessment data are to be made available in easily accessible forms to support the efforts of states and others to analyze the data, interpret results to the public, and improve education performance.

#### Status of Implementation

The National Center for Education Statistics is placing a high priority on “highlight” reports and national report cards for each subject, which are aimed at the general public. NAEP data will be accessible through a new Internet web site, customized for particular data users. Priorities for NAEP secondary analysis grants were revised to encourage wider use of NAEP data by national and state policy makers, educators, and researchers and to focus the analyses on interpretive and education improvement purposes. Also, NCES is continuing to develop and provide training on software for analyzing NAEP data.

#### Report Results Using Performance Standards

In 1988, Congress created the Governing Board and authorized it to set performance standards—called achievement levels—for reporting

National Assessment results. Under the redesign policy, achievement levels are to be used as the primary (although not exclusive) means for reporting National Assessment results. The achievement levels describe “how good is good enough” on the various tests that make up the National Assessment. Previously, the National Assessment reported average scores on a 500-point scale. There was no way of knowing whether a particular score represented strong or weak performance and whether the amount of change from previous years’ assessments should give cause for concern or celebration. The National Assessment now also reports the percentage of students who are performing at or above “Basic,” “Proficient,” and “Advanced” levels of achievement.

The achievement levels have been the subject of several independent evaluations, some controversy, and conflicting recommendations. Recommendations have been carefully considered and some have been used to improve the standard-setting procedures. While the current procedures are among the most comprehensive used in education, the Governing Board remains committed to making continual improvements.

#### Status of implementation

The Governing board will continue to set achievement levels for reporting NAEP results. These achievement levels are to be used on a developmental basis until a determination is made that the levels are reasonable, valid, and informative to the public. At that point, the developmental designation will be removed.

The Governing Board views standard setting as a judgmental, not a scientific, process. However, the process must be conducted in a manner that is technically sound and defensible. The Governing Board is preparing a report required by Congress to respond to the assertion that the process for setting the achievement levels is “flawed.” This report will include a detailed plan for reviewing the criticisms and compliments found in the evaluation reports that studied the achievement levels. The plan also will address alternatives to the current level-setting procedures.

#### *Simplify the Technical Design for the National Assessment*

The current design of the National Assessment is very complex. The redesign policy requires that the research and testing companies that compete for the contract to conduct the National Assessment must identify

options to simplify the design of the National Assessment. Examples of NAEP’s complexity include: (1) National and state results are based on completely separate samples. (2) No student takes the complete set of test questions in a subject and as many as twenty-six different test booklets are used within a grade; thus scores on NAEP are calculated using very sophisticated statistical procedures. (3) Students, teachers, and principals complete separate background questionnaires, which may be submitted at different times, complicating their use in calculating assessment results. (4) The data for every background question collected must be compiled before any report can be produced, regardless of whether the data from the background question will be included in a report, lengthening the time from data collection to reporting.

#### Status of Implementation

This is a “work in progress.” Options for combining the national and state samples are being developed by the contractors in collaboration with NCES and the Governing Board. Similarly, options to reduce the size of the state sample are being considered. An option to increase the precision of the state results will be implemented in the year 2000 mathematics and science state assessments. Progress also has been made in shortening the time between data collection and reporting by eliminating the requirement to link certain background questionnaires to student achievement data. Plans for a short-form of the National Assessment, using a single test booklet, are being implemented, with a pilot possibly as early as the year 2000. The purpose of the short-form trial is to enable faster initial reporting of results and, possibly, for states to have access to NAEP assessment results in years in which NAEP assessments are not scheduled in particular subjects. Plans also are in the development stage for improving the quality, relevance, and efficiency of background questionnaires.

#### *Measure Student Achievement at Grades 4, 8, and 12*

The primary purpose of the National Assessment is to measure student achievement at grades 4, 8, and 12 in academic subjects at the state and national level and for subgroups, showing trends over time in the percent of students at or above each achievement level. The subjects to assess are those listed in the national educational goals—reading, writing, mathematics, science, U.S. history, geography, world history, civics,

economics, the arts, and foreign language. Grades 4, 8 and 12 are considered to be important transition points in American education. Reporting by grade is generally thought to be relevant for policy than the reporting by age which was used at NAEP's inception and in long-term trend reporting.

Although grade 12 performance is important as an "exit" measure from the K-12 system, here are problems with grade 12 results. The problems are that student and school participation rates and student motivation at grade 12 are now. The Governing Board has considered whether the change NAEP to another grade at the high school level, examining both anecdotal and empirical evidence. Anecdotal evidence about the low motivation of high school students taking low stakes tests in the spring of their senior year raises serious questions about whether NAEP should test a grade 12. However, the empirical evidence in NAEP does not indicate that switching to grade 11 would result in higher motivation on the part of students or greater accuracy in the results. In fact, there is some evidence that twelfth graders taking NAEP may try harder in some cases than eleventh graders. The redesign policy asks the companies that compete for the NAEP contract to find ways to increase school and student participation rates and student motivation. Until they increase, National Assessment reports should include clear caveats about interpreting grade 12 results.

#### Status of Implementation

Because the empirical evidence does not warrant a change at this time, NAEP should continue to test at grade 12. New NAEP contracts have been awarded for the conduct of assessments through the year 2003. The contracts are designed to measure student achievement at grades 4, 8, and 12; report state, national, and subgroup results; report trends over time; and use performance standards for reporting results. Caveats for interpreting grade 12 results have been added to reports. However, more attention needs to be placed on improving grade 12 participation rates and student motivation. Toward this end, NCES is planning a series of studies, including NAEP transcript studies, to examine the relationship between student achievement and motivation.

#### What NAEP Is Not Designed To Do

The NAEP redesign policy attempts to focus NAEP on what it does best. What the National Assessment does best is measure student achievement. Focusing

NAEP on what it does best comes with a related idea—recognizing and limiting what NAEP is not designed to do.

Although the National Assessment is well designed for measuring student achievement and trends over time, it is not a good source of data for drawing conclusions about or providing explanations for the level of performance that is reported. It also is not a measure of personal values, a national curriculum, an appropriate means for improving instruction in individual classrooms, or a basis for evaluating specific pedagogical approaches.

The National Assessment is what is known as a "cross-secondary survey," an effective and cost-efficient means for gathering data on student achievement. A cross-sectional survey gathers data at one point in time. In the case of NAEP, data are gathered on national and state-representative samples of students at a particular time during the school year. The sample is large enough to permit reasonably accurate estimates of subgroup performance (e.g., by sex, race, and ethnicity). Change over time can be measured by administering the same survey again in later years, under the same testing conditions, with samples of students that are similar to the ones tested earlier. Comparisons can be made within and cross the subgroups and for the whole sample.

However, a cross-sectional survey cannot provide answers about what causes the level of performance that is reported. Measuring the causes of achievement would involve an experimental design, with specific research questions to answer, pre- and post-testing of students, and comparisons of results between groups of students receiving a particular educational approach with those that are not. While some may view such research as a worthwhile part of NAEP, the need for pre- and post-testing alone would double the costs of NAEP testing. Because pre- and post-testing would require additional administrative burden on schools and more time away from instruction for students, it could severely hamper school and student participation rates in NAEP, especially with NAEP's annual assessment schedule. Too few schools and students in the sample, in turn, would jeopardize NAEP's ability to provide national and state-representative student achievement results.

The best that can be done regarding explanation or interpretation of results is to report on background variables that may be associated with achievement. However, in many cases, the data from background questions collected by

NAEP are inconclusive or counter to what one would expect. Even where the associations are stronger, the data are not adequate for supporting conclusions that explain why achievement is at the level reported. Clearly, the use of NAEP background data to explain or interpret achievement results should be done with caution.

#### Status of Implementation

Under the new NAEP contracts, the collection of background information will be more focused. The plan is to collect a well-defined core of background information. For example, the well-defined core of background information will include the data that are required for every assessment—e.g., data on sex, race, ethnicity, whether the students are in public or private schools, etc. In addition, each assessment will have a set of background questions designed specifically for the subject being assessed, with each set being determined by policy. Therefore, the background questions for the mathematics assessment will vary from those for the science or reading assessments.

The intent is not only to be more purposeful about what is collected, but more strategic about how it is collected as well. For example, in the past, information on TV watching by students was collected regularly as a part of every assessment. In the same year, the same background questions could be asked of the students in each separate national sample. Clearly, whether two or more subjects are being assessed in a particular year, it may not be necessary to ask identical questions across all of the assessments. Similarly, it may not be necessary to ask certain questions every year. In addition, the background questions themselves will be pilot tested to reduce the possibility of misinterpretation.

#### Reporting NAEP Results

The redesign policy provides the National Assessment results should be released with the goal of reporting results six to nine months after testing. Reports should be written for the American public as the primary audience and should be understandable, free of jargon, easy to use and widely disseminated. National Assessment reports should be high technical quality, with no erosion of reliability, validity, or accuracy.

The amount of detail in reporting should be varied. *Comprehensive reports* would be prepared to provide an in-depth look at a subject the first time it is assessed using a newly adopted test

framework, testing many students and collecting background information. Although scale scores also will be used, achievement levels shall continue to be the primary method for reporting NAEP results. Test questions, scoring guides, and samples of students work that illustrate the achievement levels—Basic, Proficient, and Advanced—will receive prominence in reports. Data also would be reported by sex, race/ethnicity, socio-economic status, and for public and private schools; other reporting categories also are possible. *Standard reports* would be more modest, providing overall results in the same subject in subsequent years using achievement levels and average scores. Data could be reported by sex, race/ethnicity, socio-economic status, and for public and private schools, but would not be broken down further. The amount of background data collected and reported would be somewhat limited in comparison to a comprehensive report. *Special, focused assessments* on timely topics also would be conducted, exploring a particular question or issue and possibly limited to one or two grades.

#### Status of Implementation

The new NAEP contracts provide for faster release of data, standards-based reporting, reports that are targeted to the general public, and three different kinds of reports: "comprehensive," "standard," and "focused." The 1998 national reading results were released in 11 months of testing; the state results in 12 months. Although still short of the Board's goal of reporting results in 6 to 9 months following testing, progress is being made.

#### Simplify Trend Reporting

The NAEP redesign policy requires the development of a carefully planned transition to enable "the main National Assessment" to become the primary way to measure trends in reading, writing, mathematics and science. This is because there are now two NAEP testing programs for reading, writing, mathematics and science. The two programs use different tests, draw different samples of students (i.e., one based on age—9, 13 and 17-year-olds, the other based on grade—4, 8 and 12), and report results in two different ways. Not surprisingly, the two different programs can yield different results, which complicates the presentation and explanation of NAEP results. In addition, this redundancy boosts costs, potentially limiting assessments in other subjects.

The first program, referred to as the "long-term trend assessments,"

monitors change in student performance using tests developed during the 1960's and 1970's. The sample of students is based on age (i.e., 9, 13, and 17-year-olds) for reading, mathematics, and science and on grade for writing (i.e., grades 4, 8 and 11). The age-based samples include students from two or more grades. For example, the 9-year-old sample has 3rd, 4th, and 5th grade students. Long-term trend assessment results are reported displaying changes over time in average scores. The second program, referred to as "main NAEP," uses tests developed more recently, reports results by grade, and employs performance standards for reporting whether achievement is good enough. As an example of the potential for confusion in maintaining two separate programs, in 1996 the long-term trend assessment program declared mathematics results flat since 1990, while main NAEP reported significant gains.

Some argue against the policy to make main NAEP the primary means for monitoring trends. They feel that being able to compare student achievement in the 1990's to achievement in the 1970's and 1980's is too important to eliminate. Others argue that the long-term trend assessments are not relevant for policy makers. This is because these assessments primarily use a sample based on the students' grade, the content of the tests is simpler, there is no standards-based reporting, and the results at times conflict with main NAEP.

#### Status of Implementation

This is a "work in progress." The National Center for Education statistics is just beginning to develop options for making the transition from long-term trend to main NAEP as the primary means for monitoring trends in achievement. Identifying options that are practical, affordable, and technically feasible will take time. The Governing Board has scheduled long-term trend assessments to be conducted in 1999, 2003, and 2007. This will afford adequate time to evaluate the viability of the options that may be proposed and at the same time maintain the long-term trend line. The immediate effect is to change the schedule for this part of the testing program from once every two years to once every four years.

#### Keep NAEP Assessment Frameworks Stable

The NAEP redesign policy states that assessment frameworks shall remain stable for at least ten years. The purpose is three-fold: to provide for measuring

trends in student achievement, to allow for change to frameworks when the case for change is compelling, and to manage costs.

By law, National Assessment frameworks are developed by the Governing Board through a national consensus process involving hundreds of teachers, curriculum experts, state and local testing specialists, administrators, and members of the public. The assessment frameworks describe how an assessment will be constructed, provide for the subject area content to be covered, determine what will be reported, and influence the cost of an assessment.

Both current practice and important developments in each subject area are considered: How much algebra should be in the 8th grade mathematics assessment? Should there be both multiple choice and constructed response items and if so, what is the appropriate mix? How much of what is measured should students know and be able to do? The frameworks receive wide public review before adoption by the Governing Board.

#### Status of Implementation

The Governing Board is solely responsible for developing and approving assessment frameworks and has been adhering to its policy of keeping the frameworks stable. With a decision to be made this year about whether to conduct a national consensus process for the 2004 mathematics assessment, the Governing Board is beginning to examine criteria for determining when a new framework is necessary. An important factor will be the impact of changing the framework on the measurement of trends in student achievement.

#### Use International Comparisons

The NAEP redesign policy states that National Assessment frameworks, test specifications, achievement levels, and data interpretations shall take into account, where feasible, curricula, standards, and student performance in other nations, and promote studies to "link" the National Assessment with international assessments.

The National Assessment is, and should be, an assessment of student achievement in the United States. It should be focused on subjects and content deemed important for the U.S. through the national consensus process used to develop NAEP frameworks. However, decisions on content, achievement levels, and interpretation of NAEP results, where feasible, should be informed, in part, by the expectations for education set by other industrialized

countries, and comparative test results. Although there are technical hurdles to overcome, consideration of such information can be useful in determining "how good is good enough" in an assessment for U.S. students.

#### Status of Implementation

The National Center for Education Statistics conducted a linking study of the 1996 NAEP science and mathematics assessments with the 1995 Third International Mathematics and Science Study (TIMSS). The Government Board used information from this linking study in setting the achievement levels for the 1996 science assessment. NCES will be conducting TIMSS again in the spring of 1999 and thirteen states have agreed to participate to collect state-representative TIMSS data. NCES will be applying a methodology for relating TIMSS to NAEP and will be evaluating the strength of the relationship.

#### *Use Innovations in Measurement and Reporting*

The NAEP redesign policy states that the National Assessment shall assess, and, where warranted, implement advances related to technology and the measurement and reporting of student achievement. In addition, the competition for NAEP contracts for assessments beginning around the year 2000 shall include a plan for conducting testing by computer in at least one subject and grade and for using technology to improve test administration, measurement, and reporting.

#### Status of Implementation

The newly awarded NAEP contracts include plans for a short-form test (described above) in 4th grade mathematics in the year 2000 and for the development of a computer-based assessment.

#### *Help States and Others Link To NAEP and Use NAEP Data To Improve Education Performance*

The NAEP redesign states that the National Assessment shall assist states, districts and others, who want to do so at their own cost, to link their test results to the National Assessment. The policy also provides that NAEP shall be designed to permit access and use by others of NAEP data and materials. These include frameworks, specifications, scoring guides, results, questions, achievement levels, and background data. In addition, the policy provides that steps be taken to protect

the integrity of the NAEP program and the privacy of individual test takers.

#### Status of Implementation

The State of Maryland and the State of North Carolina have collaborated with Governing Board on studies to examine the content of their respective state mathematics test in light of the content of NAEP. The National Center for Education Statistics has a special grants program that provides funds to analyze NAEP data. The NCES has amended priorities for this grants program to encourage applications from states (and others) to conduct analyses that will be practical benefit in interpreting NAEP results and in improving education performance. The National Academy of Sciences report "Uncommon Measures," describes the many technical difficulties involved in linking state results to NAEP. The NCES is planning a major conference with the states to provide a forum for discussing and addressing these difficulties. In addition NCES is planning to conduct studies on various linking methodologies to provide insight on how the linking of NAEP and state assessments may best be done.

#### *National Assessment Redesign: Implications for Reauthorization*

The Governing Board's redesign policy is directed at the operation of the National Assessment program. It does not address governance of the National Assessment. While there are a number of areas in the current NAEP legislation for which change should be considered, the NAEP redesign policy can, with two exceptions, be implemented within the current NAEP legislation.

The first exception has to do with the subjects to assess. Current law ties the subjects covered by NAEP to reading, and the other subjects listed in the national education goals. The Governing Board agrees that these subjects should be assessed by the National Assessment and, accordingly, has adopted the schedule displayed in Table 1 above. However, the national education goals are about to expire. The Governing Board recommends that, with respect to subjects to assess, the reauthorization of the National Assessment should be consistent with the schedule of assessments adopted by the Governing Board.

The second issue has to do with long-term trend assessments. Current law requires that assessments using age-based samples be conducted at least once every two years. Since the only assessments using age-based samples are the reading, science and mathematics long-term trend

assessments, this provision is interpreted as requiring long-term trend assessments once every two years. In accordance with the schedule of assessments, the Governing Board recommends that the NAEP legislation be modified so that the frequency of the long-term trend assessments is changed to at least once every four years.

#### *Conclusion*

The National Assessment in the next century will provide student achievement results at the national level each year. State-level data will be provided every other year. Student achievement in reading, writing, mathematics and science will, appropriately, receive the most attention, with testing once every four years, but not to the exclusion of other important subjects. By continuing to report results using achievement levels and improving the process by which achievement levels are set, the National Assessment will help advance standards-based assessment and reporting in the United States. With a focus on its core purpose—measuring and reporting on the status of student achievement and change over time—the National Assessment design can be made more streamlined, more effective, and more efficient. With a clear sense of its primary audience—The general public—National Assessment reports will have more impact.

With a predictable schedule of assessments and reporting of National Assessment results, the public at regular intervals will discuss and debate education quality, states can plan ahead for their participation, and educators will have an external standard against which to compare their own efforts.

Additional Information: Written comments must be received by June 9, 1999 at the following address: Mark D. Musick, Chairman (Attention: Ray Fields), National Assessment Governing Board, 800 North Capitol Street NW, Suite 825, Washington, DC 20002-4233.

Written comments also may be submitted electronically by sending electronic mail (e-mail) to Ray\_Fields@ED.GOV by June 9, 1999. Comments sent by e-mail must be submitted as an ASCII file avoiding the use of special characters and any form of encryption. Inclusion in the public record cannot be guaranteed for written statements, whether sent by mail or electronically, received after June 9, 1999.

*Public Record:* A record of comments received in response to this notice will be available for inspection from 8 a.m. to 4:30 p.m., Monday through Friday, excluding legal holidays, in Suite 825,

800 North Capitol Street, NW,  
Washington, DC, 20002.

Dated: May 17, 1999.

**Roy Truby,**

*Executive Director, National Assessment  
Governing Board.*

[FR Doc. 99-12746 Filed 5-19-99; 8:45 am]

BILLING CODE 4000-01-M

**DEPARTMENT OF ENERGY**

**Federal Energy Regulatory  
Commission**

[Docket No. EC99-50-000]

**PacifiCorp; Notice of Filing**

May 14, 1999.

Take notice that on May 13, 1999, PacifiCorp filed a correction to Richard T. O'Brien's testimony which was filed with the Application in this docket.

Specifically, the name "ScottishPower" should be substituted for the name "PacifiCorp" at the end of the 17th line on page 4 of Mr. O'Brien's testimony.

Any person desiring to be heard or to protest such filing should file a motion to intervene or protest with the Federal Energy Regulatory Commission, 888 First Street, N.E., Washington, D.C. 20426, in accordance with Rules 211 and 214 of the Commission's Rules of Practice and Procedure (18 FR 385.211 and 385.214). All such motions and protests should be filed on or before May 24, 1999. Protests will be considered by the Commission to determine the appropriate action to be taken, but will not serve to make protestants parties to the proceedings. Any person wishing to become a party must file a motion to intervene. Copies of this filing are on file with the Commission and are available for public inspection. This filing may also be viewed on the Internet at <http://www.ferc.fed.us/online/rims.htm> (call 202-208-2222 for assistance).

**Linwood A. Watson, Jr.,**

*Acting Secretary.*

[FR Doc. 99-12682 Filed 5-19-99; 8:45 am]

BILLING CODE 6717-01-M

**DEPARTMENT OF ENERGY**

**Federal Energy Regulatory  
Commission**

[Docket No. GT99-30-000]

**Reliant Energy Gas Transmission  
Company; Notice of Proposed  
Changes in FERC Gas Tariff**

May 14, 1999.

Take notice that on May 7, 1999, Reliant Energy Gas Transmission Company ("REGT"), formerly NorAm Gas Transmission Company, tendered for filing its FERC Gas Tariff, Fifth Revised Volume No. 1 superseding Fourth Revised Volume No. 1, to be effective June 6, 1999.

REGT states that the purpose of this filing is to reflect its name change to Reliant Energy Gas Transmission Company and also to make minor ministerial changes.

Any person desiring to be heard or to protest said filing should file a motion to intervene or a protest with the Federal Energy Regulatory Commission, 888 First Street, NE, Washington, DC 20426, in accordance with Sections 385.214 or 385.211 of the Commission's Rules and Regulations. All such motions or protests must be filed in accordance with Section 154.210 of the Commission's Regulations. Protests will be considered by the Commission in determining the appropriate action to be taken, but will not serve to make protestants parties to the proceedings. Any person wishing to become a party must file a motion to intervene. Copies of this filing are on file with the Commission and are available for public inspection in the Public Reference Room. This filing may be viewed on the web at <http://www.ferc.fed.us/online/rims.htm> (call 202-208-2222 for assistance).

**Linwood A. Watson, Jr.,**

*Acting Secretary.*

[FR Doc. 99-12670 Filed 5-19-99; 8:45 am]

BILLING CODE 6717-01-M

**DEPARTMENT OF ENERGY**

**Federal Energy Regulatory  
Commission**

[Docket No. RP99-303-000]

**Western Gas Interstate Company;  
Notice of Request for Waiver**

May 14, 1999.

Take notice that on May 7, 1999, Western Gas Interstate Company (WGI), 211 North Colorado, Midland, Texas 79701, tendered for filing a petition for waiver of the electronic communication

and Internet transaction requirement of the Commission's Order Nos. 587, et seq. WGI states that it is a small company located in a discrete geographic area with only five customers, and that the cost of compliance is prohibitive. Further, WGI states that all gas on its system is delivered to end users and municipal distribution systems, and that no gas is delivered for further transportation to any interstate or intrastate pipeline. According to WGI, its customers have never released capacity or used its existing electronic bulletin board, and make their nominations directly to other pipelines that control all system interconnects.

WGI states that copies of its filing have been mailed to all of its jurisdictional customers and to affected state regulatory commissions.

Any person desiring to be heard or to protest said filing should file a motion to intervene or a protest with the Federal Energy Regulatory Commission, 888 First Street, NE., Washington, DC 20426, in accordance with Sections 385.214 or 385.211 of the Commission's Rules and Regulations. All such motions or protests must be filed in accordance with Section 154.210 of the Commission's Regulations. Protests will be considered by the Commission in determining the appropriate action to be taken, but will not serve to make protestants parties to the proceedings. Any person wishing to become a party must file a motion to intervene. Copies of this filing are on file with the Commission and are available for public inspection in the Public Reference Room. This filing may be viewed on web at <http://www.ferc.fed.us/online/rims.htm> (call 202-208-2222 for assistance).

**Linwood A. Watson, Jr.,**

*Acting Secretary.*

[FR Doc. 99-12668 Filed 5-19-99; 8:45 am]

BILLING CODE 6717-01-M

**DEPARTMENT OF ENERGY**

**Federal Energy Regulatory  
Commission**

[Docket No. GT99-29-000]

**Williston Basin Interstate Pipeline  
Company; Notice of Filing**

May 14, 1999.

Take notice that on May 7, 1999, Williston Basin Interstate Pipeline Company (Williston Basin), P.O. Box 5601, Bismarck, North Dakota 58506-5601, tendered for filing as part of its FERC Gas Tariff, Second Revised Volume No. 1, the following revised