National Assessment Governing Board

Ad Hoc Committee on NAEP Background Information

May 16, 2013
2:00 p.m.-4:00 p.m.

AGENDA

2:00 – 2:05 pm	Welcome and Introductions *Terry Holliday, Committee Chair*	Attachment A
2:05 – 2:25 pm	Student Questionnaires: NAEP and International Assessments—PISA, TIMSS, and PIRLS a. Core Data on Students, Families, and Homes b. Subject-Specific Questionnaires in Reading, Mathematics, and Science *James Deaton and Daniel McGrath, NCES*	Attachment B
2:25 – 2:45 pm	NAEP Data Explorer: Further Improvements *Rebecca Moran, ETS*	Attachment C
2:45 – 3:00 pm	Plans for Focused NAEP Reports *Grady Wilburn, NCES*	Attachment D
3:00 – 3:15 pm	Exploratory Analyses of NAEP Data: (a) Science *Alan Friedman, Governing Board Member* (b) Education Indicators *Larry Feinberg, NAGB Staff*	Attachment E
3:15 – 3:55 pm	Review of NAEP Background Information Framework *Larry Feinberg, NAGB Staff*	Attachment F
3:55 – 4:00 pm	Next Steps *Committee Members*	
Information Item	Data Collection and Reporting on Time and Learning in International Assessments	Attachment G

Attachment A

National Assessment Governing Board

Ad Hoc Committee on NAEP Background Information

BACKGROUND AND PURPOSE

As part of the resolution on NAEP background questions, adopted by the Governing Board in August 2012, an Ad Hoc Committee on NAEP Background Information is to be established for one year. Its purposes are as follows:

1. Monitor implementation of the Policy Statement on NAEP Background Questions and the Use of Contextual Data in NAEP Reporting, adopted August 4, 2012.

2. Review the *NAEP Background Information Framework*, adopted August 1, 2003. Recommend revisions, additions, or replacement, as deemed necessary or desirable.

3. Recommend a permanent arrangement for Board consideration of background questions and the reporting of contextual data in NAEP. This work is now divided between the Reporting and Dissemination and Assessment Development committees of the Board.

SPECIFIC ISSUES

As listed in the resolution, these include:

- Making greater use of contextual data in NAEP Report Cards and focused reports.
- Using background data to describe patterns and trends, including the educational experiences of different student groups.
- Detailed frameworks to support the selection of non-cognitive topics and questions, including their connection to student achievement.
- Clusters of questions on topics of continuing interest, such as technology and out-of-school learning, to be used regularly or rotated across cycles.
- Modules on issues of current policy interest.
- Elimination of duplicative, low-priority, or unproductive topics and questions.
- Use of questions from international assessments, such as TIMSS and PISA.
- Improved measures of socio-economic status (SES), including exploration of SES index.
- Spiral sampling and rotation of background questions in different years.
- Increasing the maximum time for students to answer background questions.
- Additional focused reports with the appointment of advisory committees in some cases.
- Exploratory analyses of existing data that may form the basis for subsequent reports.
- Consistency of wording to preserve trends.
- Further improvements in the NAEP Data Explorer.

COMMITTEE COMPOSITION AND TIMELINE

The Ad Hoc Committee will include six or seven Board members with a variety of perspectives and membership in different standing committees of the Board. The Committee will convene during each quarterly meeting of the Governing Board, and is expected to make its final report in August 2013.

COMMITTEE MEMBERSHIP

Terry Holliday, Chair

Doris Hicks

Andrew Ho

Brent Houston

Dale Nowlin

Joseph O'Keefe, S.J.

Susan Pimentel

Leticia Van de Putte

Adopted Unanimously – 8/4/12

Policy Statement on NAEP Background Questions and the Use of Contextual Data in NAEP Reporting

INTRODUCTION

By statute, the purpose of the National Assessment of Educational Progress is to provide a "fair and accurate" measure of student achievement and achievement trends. Academic or cognitive questions are its primary focus; the American public is its primary audience. However, in addition to reporting on what American students know and can do, NAEP has collected data for more than 40 years that provide a context for reporting and interpreting achievement results. According to the statute, such factors, both in and out of school, must be "directly related to the appraisal of academic achievement."

In each assessment NAEP administers background questionnaires for students, their teachers, and schools. The questionnaires deal with educational experiences and other factors, such as teacher training or out-of-school learning activities, that are related to academic achievement. Data on several hundred background or noncognitive variables are available on the Internet through the NAEP Data Explorer. However, for more than a decade, little use has been made of this information in NAEP reports. The data have received minimal attention and had little impact despite the considerable efforts expended in developing and approving questionnaires and collecting and tabulating responses.

In October 2011 the National Assessment Governing Board convened an expert panel to recommend how to make better use of existing NAEP background questions and to propose an analytic agenda for additional topics and questions that would be useful in developing education policy and of value to the public. The panel report, entitled, *NAEP Background Questions: An Underused National Resource*, was presented to the Board in March 2012 by Marshall Smith, former U.S. Under Secretary of Education, who chaired the six-member panel.

Many of the panel recommendations build on the *Background Information Framework for the National Assessment of Educational Progress*, adopted by the Governing Board after it received final authority from Congress over non-cognitive items on the assessment. The framework was adopted in 2003, but has not been fully implemented.

The following policies are based on recommendations by the expert panel. The Board has also taken into consideration a wide range of public comment and the analysis provided by the National Center for Education Statistics.

It is important to understand that the National Assessment is not designed to show cause-and-effect relationships. Its data should not be used to "prove" what schools should do. But, as the *Background Information Framework* declares, NAEP's "descriptions of the educational circumstances of students…, considered in light of research from other sources, may provide important information for public discussion and policy action." The Board believes the National Assessment should improve upon its efforts to collect contextual information and present it clearly to the public, which will add to NAEP's value to the nation.

POLICY PRINCIPLES

1. NAEP reporting should be enriched by greater use of contextual data derived from background or non-cognitive questions asked of students, teachers, and schools. Such data will be used both in regular Report Cards and in special focused reports.

2. Reporting of background data will describe patterns and trends, including the educational experiences of different groups of students. Care should be taken not to suggest causation.

3. Detailed frameworks will be published with the theoretical rationale and research evidence that support the selection of topics and questions in background questionnaires and their connection to student achievement. Such frameworks should be updated for each assessment cycle and provide the basis for new topics and questions.

4. An ad hoc committee of the Board will be established for one year to monitor implementation of this resolution, review the *NAEP Background Information Framework*, and recommend a permanent arrangement for Board consideration of background questions and the reporting of contextual data in NAEP.

IMPLEMENTATION GUIDELINES

For Questions and Questionnaires

1. Clusters of questions will be developed on important topics of continuing interest, such as student motivation and control over the environment, use of technology, and out-of-school learning, which could be used regularly or rotated across assessment cycles.

2. Modules will be prepared for special one-time studies to provide descriptive information on issues of current policy interest.

3. A thorough review will be conducted to eliminate duplicative or low-priority questions. Unproductive topics and questions will be dropped.

4. NAEP will include background questions from international assessments, such as PISA and TIMSS, to obtain direct comparisons of states and TUDA districts to educational practices in other countries.

5. Because of the value of preserving trends, consistent wording of questions should be maintained on topics of continuing interest. Changes in wording must be justified. However, as practices and circumstances change, new questions will be introduced in a timely manner to gather data on topics of current interest.

6. The development and use of improved measures of socio-economic status (SES) will be accelerated, including further exploration of an SES index for NAEP reporting.

For Data Collection

7. The maximum time for students to answer the background questionnaire will be increased from 10 to 15 minutes on new computer-based assessments. Consideration should be given to a similar increase in paper-and-pencil assessments.

8. Whenever feasible, assessment samples should be divided (spiral sampling) and background questions rotated in different years in order to cover more topics without increasing respondent burden. These practices will be initiated in the assessments of reading and mathematics, which are conducted frequently, and considered for other subject areas if the frequency of testing permits.

For Reporting

9. Special focused reports with data through the 2013 assessment will be issued on the following topics: private schools, charter schools, gender gaps, and black male students. Reports shall include significant contextual information as well as cognitive results. Advisory committees, composed of a range of knowledgeable persons, may be appointed to provide input on reporting issues.

10. Exploratory analyses will be carried out to determine if existing background questions may form the basis for additional focused reports. Such reports may be issued by the Governing Board as well as by the National Center for Education Statistics.

11. The NAEP Data Explorer should be further improved to make data more accessible to general, non-specialist users. Tables and very simple-to-construct charts will be prepared to present data on important topics of wide public interest. Additional means of disseminating information, using new technology such as simple apps that would allow parents, teachers, and others to access background and achievement data, will be explored.

Comparison of NAEP Student Survey Questions to International Student Survey Questions

The purpose of this document is to compare the NAEP student survey questions to those survey questions found on three international assessments: the Progress in International Reading Literacy Study (PIRLS), the Trends in International Mathematics and Science Study (TIMSS), and the Program for International Student Assessment (PISA).

NAEP and the international assessments all have survey questions that are not subject-specific. These are often referred to as "core" questions. The first section in this document will offer a comparison of NAEP core questions to international assessment core questions. Subsequent sections will compare subject-specific questions for NAEP to the international assessments for the following: mathematics, reading, and science. Finally, there will be a discussion of similarities, differences, and implications for NAEP.

Before beginning this comparison, it is helpful to consider the student respondent groups for NAEP compared to the international assessments. NAEP assesses students in grades 4, 8, and 12. PIRLS assesses students in grade 4, TIMSS in grades 4 and 8, and PISA does not target a specific grade, but rather 15-year old students. Students in this age group would typically be in ninth or tenth grade, thus representing respondents between NAEP's grade 8 and 12 assessments.

The NAEP core, mathematics, and reading questions used for comparison purposes in this document were all included in the 2013 operational assessment. The NAEP science questions used for comparisons will be pilot tested in 2014. All PIRLS and TIMSS questions used for comparisons were administered in 2011, while all PISA questions used for comparisons were administered in 2012. The 2012 subject-specific questionnaires for PISA focus on mathematics, which was the subject assessed in-depth in 2012. We have not included comparison to the PISA science questions in 2006 or PISA reading questions in 2009[1] since the 2012 mathematics questions provide the most accurate picture of the types of PISA subject-specific questions that will be used in the future.

1. Student core question comparison

The number of core questions[2] asked of student respondents is substantially greater for PIRLS, TIMSS, and PISA than for NAEP. For example, the NAEP core has 11 questions for grade 4 students, 13 questions for grade 8 students, and 15 questions for grade 12 students. The TIMSS and PIRLS[3] core has 39 to 40 questions for grade 4 students, depending on a potential skip pattern. The TIMSS core has 40 to 42 questions for grade 8 students depending on their answer

[1] Science was assessed in-depth in 2006 and reading was assessed in-depth in 2009.
[2] The number of questions has been calculated based on the number of responses needed. For example, TIMSS asks students: *When were you born?* Students respond by first reporting the month and then the year. This calculation would consider this as two questions since two responses are needed.
[3] The TIMSS and PIRLS core uses the same questions.

to a few specific questions. The PISA core has 115 questions, 65 of which are given to all respondents and 50 of which are spiraled so they are given to two-thirds of respondents.

Table 1 shows the number of questions by topic on the NAEP core compared to the international surveys. The majority of NAEP core questions have comparable questions for TIMSS, PIRLS, and PISA. An important consideration when viewing the table is that NAEP, unlike the international surveys, does not use a matrix question format in any of its core questions. PIRLS, TIMSS, and PISA all have at least some matrix questions on their core questionnaire. Within this table, each sub-item in a matrix question would be counted as one question.

The NAEP core has four topics that do not have a comparable equivalent found on the PIRLS, TIMSS, and PISA core: reading, post-high school activities (grade 12 only), high school program (grade 12 only), and zip code.

Three topics covered in all international core questionnaires that do not have a comparable NAEP core question are gender, birthdate, and birthplace. However, NAEP does collect gender and birthdate information through school records. NAEP also administered a birthplace question to students in a pilot test in 2012. This question was not included in the 2013 revised operational assessment core questionnaire due to minimal variation across response options as well as political sensitivity issues related to administering this question in Puerto Rico.

Table 1: Number of NAEP Core Questions Compared to PIRLS, TIMSS, and PISA by Topic

	NAEP	PIRLS / TIMSS (Grade 4)	TIMSS (Grade 8)	PISA (15-year-old)[4]
Gender	*School records collected*	1	1	1
Race/ethnicity	2	2	2	2
Grade	-	-	-	1
Birthdate	*School records collected*	2	2	1
Birthplace (parents and self)	-	3	3 or 4	3
Books in home	1	1	1	1
Computer	1	3	3	-
Items in home[5]	1	8	8	22
Reading	1	-	-	-
Family involvement in school	1	4	4	-
Student teacher relations	-	-	-	5

[4] The PISA questions that are given to two-thirds of respondents (50 total) all fall into the following categories: student teacher relations, student attitudes about school, and problem solving. The remaining PISA questions are given to all respondents.

[5] There are questions that ask students to fill in one or more response options. All of the questionnaires, NAEP included, use this format to ask students their race, presenting the following options: White, Black or African American, Asian, American Indian or Alaska Native, and Native Hawaiian or other Pacific Islander. However, NAEP also uses this for the items in home question. All the international questionnaires use a yes/no structure format for items in home. This yes/no format aids analysis, but also substantially increases burden. For example, the PISA questionnaire has two matrix questions regarding items in the home. The first matrix question asks about 17 different items, whereas the second asks students about five different items. Each of these requires a yes/no response, meaning students would ultimately provide 22 answers for both matrix questions.

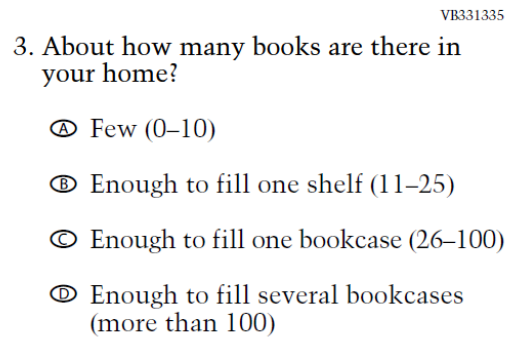
	NAEP	PIRLS / TIMSS (Grade 4)	TIMSS (Grade 8)	PISA (15-year-old)[6]
Student attitudes about school	-	3	3	23
Problem-solving	-	-	-	22
Negative incidents at school	-	6	6	-
Outside of school activities	-	4	-	-
School absenteeism	1	1	1	3
Earlier education	-	-	-	3
Repeated a grade	-	-	1	4
Parental education	2 *(grades 8 and 12 only)*	-	2	10
Parental occupation	-	-	-	6
Language in home	1	1or 2	1 or 2	1
People living in home	1	-	-	6
Educational expectations	-	-	1	1
Post-high school activities	1 *(grade 12 only)*	-	-	-
High school program	1 *(grade 12 only)*	-	-	-
Zip code	1	-	-	-
Total	**15**	**39 - 40**	**40 - 42**	**115**

Books in the home: One example of similarities and difference across surveys

The vast majority of questions on the NAEP student core can be mapped to comparable questions on PIRLS, TIMSS, and PISA. However, while the topic may be similar, or even exactly the same, the question itself may be somewhat different – in terms of phrasing and response options.

For example, the NAEP, PIRLS, TIMSS, and PISA questionnaire all have a question regarding books in the home. However, this question is asked in four different ways. This includes differences in the item stem, guidance as to what to count as a book, and varying response options. We will first show the four different versions of this question and then outline the differences between them.

NAEP version of books in the home:



VB331335

3. About how many books are there in your home?

Ⓐ Few (0–10)

Ⓑ Enough to fill one shelf (11–25)

Ⓒ Enough to fill one bookcase (26–100)

Ⓓ Enough to fill several bookcases (more than 100)

[6] The PISA questions that are given to two-thirds of respondents (50 total) all fall into the following categories: student teacher relations, student attitudes about school, and problem solving. The remaining PISA questions are given to all respondents.

PIRLS and TIMSS (Grade 4) version of books in the home:

5

About how many books are there in your home? (Do not count magazines, newspapers, or your school books.)

*Fill in **one** oval only.*

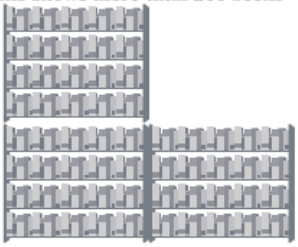
None or very few (0–10 books)---- ① This shows 10 books

Enough to fill one shelf (11–25 books)---- ② This shows 25 books

Enough to fill one bookcase (26–100 books)---- ③ This shows 100 books

Enough to fill two bookcases (101–200 books)---- ④ This shows 200 books

Enough to fill three or more bookcases (more than 200)---- ⑤ This shows more than 200 books



TIMSS (Grade 8) version of books in the home:

8

About how many books are there in your home? (Do not count magazines, newspapers, or your school books.)

Fill in one oval only.

None or very few (0–10 books) --①

Enough to fill one shelf (11–25 books) --②

Enough to fill one bookcase (26–100 books) --③

Enough to fill two bookcases (101–200 books) --④

Enough to fill three or more bookcases (more than 200) --⑤

PISA version of books in the home:

Q30 How many books are there in your home?

There are usually about 15 books per foot of shelving. Do not include magazines, newspapers, or your schoolbooks.

(Please darken only one circle.)

0-10 books	O 1
11-25 books	O 2
26-100 books	O 3
101-200 books	O 4
201-500 books	O 5
More than 500 books	O 6

Item stem comparison

The NAEP version asks students: *About how many books are there in your home?* The PIRLS and TIMSS versions have the same item stem, but with an added parenthetical: *Do not count magazines, newspapers, or your schools books*. In contrast to the other surveys, the PISA version omits *About* in the item stem. It also includes two sentences of guidelines underneath the item stem the first of which is unique to PISA: *There are ususally about 15 books per foot of shelving*. The second guidline is very similar to the paranthetical found in the PIRLS and TIMSS version. The PISA phrasing is "Do not include" whereas as the PIRLS and TIMSS phrasing is "Do not count." PISA has "schoolbooks" as one word, while PIRLS and TIMSS has it as two words [i.e., "school books"].

Response option comparison

The NAEP version presents students with four possible response options ranging from 0-10 to more than 100. The PIRLS and TIMSS (Grade 4) version presents students with five possible options ranging from 0-10 to more than 200. Response options A through C are the same as the the NAEP version in terms of numeric ranges: 0-10, 11-25, and 26-100. However, the description of this numeric range differs. For instance, on the NAEP questionnaire it is "Few (0 – 10)" whereas on the PIRLS and TIMSS questionnaire it is "None or very few (0 – 10 books)." Most notably, the PIRLS and TIMSS (Grade 4) version provides a visual for students to help them conceptualize what different numbers of books would look like. The TIMSS grade 8 version is the same as the grade 4 version, except without the visual. The PISA version contains the most response options (six total) as it further dissagregates those at the upper end of the spectrum with the greatest being "More than 500 books."

Structural differences

The structure of PISA questions is at times radically different than those on NAEP, PIRLS, or TIMSS. For example, PISA has some questions that introduce a scenario and ask students several questions based on this scenario. Here is an example of this type of question[7]:

Q41 *Suppose that you have been sending text messages from your cell phone for several weeks. Today, however, you can't send text messages. You want to try to solve the problem.*

What would you do? For each suggestion, darken the option that best applies to you.

(Please darken only one circle in each row.)

	I would definitely do this	*I would probably do this*	*I would probably not do this*	*I would definitely not do this*
a) I press every button possible to find out what is wrong.	○ 1	○ 2	○ 3	○ 4
b) I think about what might have caused the problem and what I can do to solve it.	○ 1	○ 2	○ 3	○ 4
c) I read the manual.	○ 1	○ 2	○ 3	○ 4
d) I ask a friend for help.	○ 1	○ 2	○ 3	○ 4

NCES is open to exploring new ways to structure questions. We recognize moving NAEP assessments and questionnaires from a paper-pencil format to a technology-based platform will expand opportunities. We also want to be cognizant of the cognitive load placed on students. The scenario type of questions on PISA may more appropriate for grade 8 and 12 NAEP respondents than for grade 4 respondents. Moreover, introducing different types of questions does increase cognitive demands on respondents.

2. Student mathematics question comparison

Mathematics questions for NAEP can be compared to those on TIMSS and PISA. The length of the NAEP and TIMSS mathematics questionnaire is relatively similar. The NAEP mathematics questionnaire has 19 questions for grade 4, 34 questions for grade 8, and either 28 or 56 questions for grade 12. There are 28 questions if students are *not* currently taking a mathematics course and 56 if they are currently taking a mathematics course. The TIMSS mathematics questionnaire has 18 questions for grade 4 and 27 questions for grade 8.

[7] This PISA core question was classified under the topic "problem-solving."

In comparison to NAEP and TIMSS, PISA has many more mathematics questions. PISA has 167 mathematics questions with each question given to two-thirds of respondents. On average, a student completing the PISA questionnaire would receive approximately 110 mathematics-specific questions.

Table 2 shows the number of questions by topic on the NAEP mathematics section of the questionnaire compared to TIMSS and PISA. As seen in the table, the majority of NAEP mathematics questions do not have comparable TIMSS or PISA questions. NAEP's mathematics questions are grouped into five categories: coursework, computer use, calculator use, after-school / tutoring programs, and student attitudes about subject. The category "student attitudes about subject" is the only one shown in the table that is found on NAEP, TIMSS, and PISA. While the categories in the table were constructed as mutually-exclusive, this does not seem to have resulted in any significant misrepresentation of overall content coverage. For example, PISA has two questions that mention computers. One question asks students: *How often do you do the following things at school and outside of school?* One of the eight sub-items to this question is: *I program computers*[8]. Another question asks students: *Thinking about all school subjects: on average, how many hours do you spend each week on the following?* One of the six sub-items to this question was: *Practice content from school lessons by working on a computer (e.g., learn vocabulary with training software)*[9]. In both cases, it is clear that the focus of the questions is not computer use specifically related to mathematics class.

Table 2: Number of NAEP Mathematics Questions Compared to TIMSS and PISA by topic

	NAEP (Grade 4)	NAEP (Grade 8)	NAEP (Grade 12)	TIMSS (Grade 4)	TIMSS (Grade 8)	PISA
Coursework	-	2	21	-	-	-
Computer use	6	12	0 or 9	-	-	-
Calculator use	2	8	1 or 14	-	-	-
After-school / tutoring programs	1	1	0 or 1	-	-	-
Student attitudes about subject	7	7	3 or 7	13	20	45
Approach to subject	-	-	-	-	-	4
Engagement with assessment	3	3	3	-	-	-
Experience with teacher	-	1	0 or 1	5	5	31
Experience with class	-	-	-	-	-	9
Perception of teachers regarding classroom management	-	-	-	-	-	3
Homework	-	-	-	-	2	
Perception of teachers regarding homework	-	-	-	-	-	3
Mathematics activities (at school / outside of school)	-	-	-	-	-	8
Mathematics compared to other subjects	-	-	-	-	-	15
Experience with types of problems / concepts	-	-	-	-	-	33
Ability to do specific tasks	-	-	-	-	-	8

[8] This PISA question was classified under the topic "Mathematics activities (at school / outside of school)."
[9] This PISA question was classified under the topic "Overall time spent on school."

	NAEP (Grade 4)	NAEP (Grade 8)	NAEP (Grade 12)	TIMSS (Grade 4)	TIMSS (Grade 8)	PISA
Overall time spent on school	-	-	-	-	-	7
Class size	-	-	-	-	-	1
Total	**19**	**34**	**28 or 56**	**18**	**27**	**167**

The overall scope of topics covered on the PISA questionnaire is substantially greater than both NAEP and PISA. This may be attributed, in large part, to the length of the questionnaire. For example, there are approximately nine times as many questions on the PISA questionnaire than the grade 4 NAEP questionnaire, five times more than the grade 8 NAEP questionnaire, and three times[10] more than the grade 12 NAEP questionnaire. NCES recognizes that the mathematics questionnaire could expand its scope and maintain current burden constraints by reducing the number of questions asked on computer and calculator use[11].

Table 2 shows 12 different topics in which PISA has at least one question – a topical range much greater than NAEP or TIMSS. However, nearly two-thirds of the 167 PISA questions could be grouped into one of three categories: student attitudes about subject, experience with teacher, and experience with types of problems / concepts. NCES has exercised significant caution regarding topics that might be categorized as assessing attitudes or beliefs due to legislation which stipulates that NAEP "only collect information that is directly related to the appraisal of academic achievement."

3. Student reading question comparison

Reading questions for NAEP can be compared to those on PIRLS. The number of PIRLS questions (37 total) is greater than for NAEP grade 4 (18 total), NAEP grade 8 (28 total), and NAEP grade 12 (36 total). Table 3 shows the number of questions by topic on the NAEP reading section of the questionnaire compared to PIRLS.

Over 70 percent of the PIRLS reading-specific questions ask students about their attitudes related to the subject. While the NAEP reading-specific questions address the topic, the majority asks students about class or school activities related to reading, which is not emphasized as much on PIRLS. Over 20 percent of the PIRLS questions ask students about reading outside of school. This topic is also covered in the NAEP questionnaires, albeit with one question that asks students how often they read for fun on their own time. While PIRLS asks students more reading-specific questions, the range of topics appears slightly greater on the NAEP reading questionnaire. The NAEP grade 12 reading questionnaire addresses coursework and computer use. All the NAEP reading questionnaires address after-school / tutoring programs and reading discussions with friends and family, whereas PIRLS has a question on library use.

Table 3: Number of NAEP Reading Questions Compared to PIRLS by topic

	NAEP (Grade 4)	NAEP (Grade 8)	NAEP (Grade 12)	PIRLS
Coursework	-	-	3	-
Computer use	-	-	7	-
After-school / tutoring programs	1	1	1	-

[10] For those grade 12 NAEP students who are currently taking a mathematics class.
[11] This was also recommended within the Expert Panel paper on NAEP Background Questionnaires.

	NAEP (Grade 4)	NAEP (Grade 8)	NAEP (Grade 12)	PIRLS
Student attitudes about subject	1	2	3	26
Reading outside school	1	1	1	8
Reading discussions with friends/family	1	1	1	-
Class or school activities related to reading	11	20	17	2
Engagement with assessment	3	3	3	-
Library use	-	-	-	1
Total	**18**	**28**	**36**	**37**

4. Student science question comparison

Science questions for NAEP can be compared to those on TIMSS. The number of science-specific questions is greater on NAEP than on TIMSS. Comparable to the NAEP mathematics questionnaire for grade 12, the NAEP science questionnaire for grade 12 asks students if they are currently taking a science course. Those students who report "No" are instructed to skip to a later question and would be asked to respond to 29 questions, whereas students who report "Yes" would be given a total of 56 questions.

Table 4 shows the number of questions by topic on the NAEP science section of the questionnaire compared to TIMSS.

Over 70 percent of the TIMSS science-specific questions at grade 4 and 75 percent at grade 8 ask students about their attitudes related to the subject. NAEP also covers this construct, albeit with less emphasis. NAEP science questions generally concentrate around class or school activities related to science. Moreover, NAEP science questions cover more topics. These include: coursework (grade 12 only), computer use, science activities outside school, and library use.

Table 4: Number of NAEP Science Questions Compared to TIMSS by topic

	NAEP (Grade 4)	NAEP (Grade 8)	NAEP (Grade 12)	TIMSS (Grade 4)	TIMSS (Grade 8)
Coursework	-	-	16		-
Computer use	1	1	0 or 1		-
Student attitudes about subject	4	4	5	13	21
Experience with teacher	2	1	0 or 1	5	5
Science activities outside school	2	2	2	-	-
Class or school activities related to science	19	31	4 or 27	-	-
Library use	1	1	0 or 1	-	-
Engagement with assessment	3	3	3	-	-
Homework	-	-	-	-	2
Total	**32**	**43**	**29 or 56**	**18**	**28**

5. Conclusion: Similarities, differences, and implications for NAEP

The similarities are greater between NAEP and international survey content coverage for core questions than for subject-specific questions. NAEP subject-specific questionnaires have aimed for breadth of coverage, whereas the international survey questionnaires aim more for depth of coverage. The Policy Statement on NAEP Background Questions and the Use of Contextual Data in Reporting that was adopted by the Board in August 2012 stated, "Clusters of questions

will be developed on important topics of continuing interest, such as student motivation and control over the environment, use of technology, and out-of school learning, which could be used regularly or rotated across assessment cycles." This recommendation favors depth of coverage over breadth. As part of our recent review of the Background Information Framework, NCES suggested the current framework be updated to include a discussion of the benefit of exploring questionnaire indexes on important topics rather than relying on stand-alone items only. The advantage of scales over single items is that item-wording effects are reduced and measurement quality is increased. This is not to suggest that NAEP should immediately delete many of its current questions and replace them with those in international surveys.

NAEP has strived to create continuity, when possible, across its questionnaires: whether that be for three different student groups (grade 4, 8, and 12), respondents (student, teacher, and school administrator), or subjects. One of the advantages of PIRLS and PISA is that only *one* student group is assessed (grade 4 students for PIRLS and 15-year old students for PISA). NAEP frequently assesses three different student groups. Consequently, with NAEP student questionnaire development there are two competing and valid desires: 1) develop questions that are appropriate and accessible for a specific grade and, 2) create uniformity across grades so questionnaire results from grades 4, 8, and 12 can be compared with each other when possible. The former desire has led NAEP to avoid using matrix questions for grade 4 students.

Plans for Reporting Contextual Variables in the NAEP Data Explorer

The NAEP Data Explorer (NDE) serves as the backbone of the NAEP Web delivery infrastructure. In addition to being an interface for users to obtain statistical results, it represents (1) a comprehensive database of historical NAEP data, (2) a computational engine that produces statistics and conducts tests of statistical significance for use in the NDE itself and most other web reporting applications, and (3) a graphics engine to visualize results in accessible ways. Capabilities in the NDE include the ability to obtain statistical results in the form of student group distributions (percentages) and achievement results (average scale score, achievement-level percentages, percentile estimates) for a single variable (e.g., Race/Ethnicity), two-way interactions (e.g., Race/Ethnicity by School Resources), and three-way interactions.

Users can produce tables of statistics for a single or multiple years for one or more jurisdictions, produce graphical summaries, perform significance testing, and extract tables, graphs, and significant testing tables (or USA maps) to desktop programs such as Word and Excel. In any given month, the NDE services approximately 100,000 requests for NAEP information. That number typically doubles around the time of a report release.

The NDE database includes all contextual/survey questions obtained from school records, asked of students, teachers, and school administrators, other reporting variables (e.g., school type, region), and in the case of technology-based assessments student use of tools during the assessment (e.g. use of "text to speech", thesaurus). The sheer volume of information in NDE may render it challenging for users to obtain desired information. For this reason new tool, tentatively called NAEP Simple, is being discussed. This new avenue for searching NAEP results would allow users to make straightforward, plain language queries. For example, users might type "reading at home" or "class size" and they would be presented with a series of links similar to what they might find using any Web search engine. NAEP Simple would allow user to obtain multiple subjects (e.g. reading and math) and cross-grade results.

second potential enhancement to the NDE is a tool called NDE Quick. This tool would generate multiple sets of results based o a single input query. Users could request numerous statistics and variables across subjects/grades/years through single request from simplified menu-driven interface.

NDE and the NAEP website have been successful for many years. However it is recognized that bringing in additional organizations with specialties in web delivery of content and computational optimization would be beneficial. As part of the recent NAEP contract awards, the NAEP Alliance will partner with three new organizations:

- **IBM** will provide database architecture and processing support to enhance the database performance of the NDE — which, in turn, will make possible the development of faster, more user-friendly tools.

- **Forum One** will help devise and lead the Internet design and communications strategy for revamping the overall approach to Web reporting — incorporating more intuitive and appealing data visualization techniques to clearly communicate NAEP results.

- **Levine & Associates** will lead the creation and implementation of designs that support our overall communications strategy for print and digital reports.

Other planned enhancements to the NAEP Data Explorer will be presented at the committee meeting.

NOTE TO Ad Hoc Committee on Background Information on Exploratory Analyses of NAEP Data

The Board's consultant, Alan Ginsburg, former Director of Policy and Program Studies at the U.S. Department of Education, has completed an exploratory analysis of background information obtained in the special NAEP 2011 assessment of 8th grade science. The study examines student attitudes and other factors, such as teacher preparation, science equipment, and instructional methods, which are related to science achievement.

Ginsburg worked on the study with Board member Alan Friedman, former director of the New York Hall of Science who now is a consultant on science museums and informal science education.

At this meeting Friedman will discuss the main findings of the report. In addition, Larry Feinberg, of the Governing Board staff, will update the Committee on a new project on using NAEP data for key education indicators that was introduced at the meeting in late February 2013. The work will be undertaken by Ginsburg and Marshall S. Smith, former Dean of the Stanford University Graduate School of Education, who chaired the Board's Expert Panel on Strengthening NAEP Background Questions, which presented its report in February 2012.

The science report and materials on the indicators project will be provided before the meeting.

Background Information

Framework

for the

National Assessment of

Educational Progress

National Assessment Governing Board
Adopted August 1, 2003
POSSIBLE CHANGES May 2013

National Assessment Governing Board

Darvin M. Winick
Chair

John H. Stevens
Chair, Ad Hoc Committee on Background Questions

Charles E. Smith
Executive Director

Lawrence Feinberg
Project Officer

Background Framework Project

Mary Lyn Bourque, Writer/Editor
Mid-Atlantic Psychometric Services

David Grissmer
RAND

Consultants

Paul E. Barton
Ina V.S. Mullis
Michael Podgursky
Mark Reckase
Herbert J. Walberg

Table of Contents

Authorised User 4/26/13 10:19 AM
Formatted: Font:12 pt

Preface

by the National Assessment Governing Board

The National Assessment of Educational Progress (NAEP) has been established by law to monitor the academic achievement of American students. In addition to its academic assessments, NAEP has collected information from hundreds of non-cognitive or background questions about students, their educational experiences in class and at home, their teachers, and their schools. Some of these questions provide data for NAEP's reporting categories, but far more have been used to give context to NAEP results or to track factors associated with academic achievement. Some have been used by scholars in social science research.

Concerns have been raised about the selection of background variables, the quality of the information obtained, and the validity of inferences drawn from it. There is also concern about the burden that collecting background information places on respondents and on the NAEP program. After the National Assessment Governing Board was granted final authority over the background questions in early 2002, it adopted a policy to focus NAEP background data on the primary purpose of the National Assessment—to provide sound, timely information on the *academic* achievement of American students. The Board also initiated a process to prepare a general framework to guide the collection and reporting of background data.

It is important to understand the National Assessment is not designed to prove cause-and-effect relationships; it cannot prescribe what should be done. But its descriptions of the educational circumstances of students at various achievement levels—considered in light of research from other sources—may provide important information for public discussion and policy action. Used with other research, the contextual data collected by NAEP may give insights into how achievement can be improved as well report to the public on how school personnel and resources related to achievement are distributed.

This framework will define the purpose and scope of NAEP's system of collecting background information, including background questionnaires and other sources of non-cognitive data. It will establish criteria for reporting background information as part of the National Assessment. The approach it suggests provides for asking various groups of questions to various samples of students at various times.

The framework reflects the following key principles:

- The selection of background topics and questions shall be designed to fulfill all legal requirements for the National Assessment and to carry out decisions regarding what NAEP will report and how to report it.

- Background information shall provide a context for reporting and interpreting achievement results and, as the statute provides, must be "directly related to the appraisal of academic achievement and to the fair and accurate presentation of such information."

- The collection of background data shall be designed to obtain information that is objective, valid, reliable, and of consistently high quality.

- The system of background data collection shall be efficient and designed to minimize the burden on respondents and on the NAEP program. As much data as possible should be obtained from school records and other reliable data sources.

- These principles shall apply both to the collection of general background information and to subject-specific background questions. The frameworks for the latter must be focused and prioritized, indicating a core set of variables for regular reporting and a more comprehensive set to be collected and reported less frequently.

- The priority order for background information is as follows: (1) reporting categories, as required by law; (2) contextual factors with a well-established relationship to achievement; and (3) subject-specific information.

There is one other consideration—the new role of the National Assessment in the No Child Left Behind Act of 2001. Under this law, all states receiving federal Title I aid are required to participate every two years in NAEP's state-level samples of reading and mathematics in grades 4 and 8. The results will provide an independent yardstick to compare trends on NAEP with performance on each state's own set of required exams.

Because No Child Left Behind places particular emphasis on closing the persistent performance gaps between various student groups, NAEP must be able to report on changes in achievement for all groups specified by law. Through its background questions, the National Assessment might also provide useful information about the students left behind and those who are ahead of them, including the sorts of schools that high-achieving and low-achieving students attend, the courses they take, the patterns of how they are taught, and the qualifications of their teachers. Over time, such descriptive information will allow NAEP to track changes in contextual and instructional factors related to student achievement and in the distribution of important educational resources.

In sum, the purpose of this Background Information Framework is to focus the collection and reporting of background data by the National Assessment and to establish clear priorities and limits. We hope to make it possible that with far fewer non-cognitive questions than it has had in the recent past, NAEP will serve the purposes of law and

provide the American public and decision makers with useful information. We are committed to improving the quality of data collected and the reporting of results.

Executive Summary

The National Assessment of Educational Progress (NAEP) is a federally authorized survey of student achievement at grades 4, 8, and 12 in various subject areas, such as mathematics, reading, writing, science, U.S. history, the arts, and foreign languages. The No Child Left Behind Act of 2001 (P.L. 107-110) requires the assessment to collect data on specified student groups, including race/ethnicity, gender, socio-economic status, disability, and limited English proficiency. It requires fair and accurate presentation of achievement data and permits the collection of background or descriptive information that is related to academic achievement and aids in fair reporting of results. The intent of the law is to provide representative-sample data on student achievement for the nation, the states, and subpopulations of students and to monitor progress over time.

The National Assessment Governing Board (NAGB) sets policy for NAEP and determines the content framework for each assessment. As a result of the No Child Left Behind Act, the Board is responsible for selecting and approving all of NAEP's non-cognitive or background questions, as well as the cognitive items over which it has had final authority since 1988. This Background Information Framework will guide the development and selection of non-cognitive topics and questions, starting with the NAEP 2006 assessment. It will fulfill the purposes of law and implement Board policy.

When NAEP began in 1969-70, its background information was limited to gender, race/ethnicity, and literacy materials at home. During the 1980s the array of non-cognitive questions expanded greatly, both to provide more contextual information and in an effort—never fully realized—to use the assessment for educational research.

This background data framework will refocus the collection of non-cognitive variables on NAEP's primary mission: providing a fair and accurate measure of student achievement and on achievement trends over time. Thus, the framework is a guide for gathering important information that will assist in reporting and understanding NAEP results. NAEP may contribute to research into improving education policy and practice, Its role in this respect is limited, but, used with other research, the contextual data collected by NAEP may give insights into how achievement can be improved as well report to the public on how school personnel and resources related to achievement are distributed.

Since by law NAEP may only collect information that is "directly related to the appraisal of academic achievement," it must concentrate on non-cognitive variables that are known from other research to have such a relationship. The law also specifically prohibits NAEP from asking about personal or family beliefs and attitudes. These points are emphasized in the Governing Board Policy Statement on the Collection and

Authorised User 4/25/13 2:36 PM
Deleted: but

Authorised User 4/25/13 2:36 PM
Deleted: i

Authorised User 4/25/13 2:38 PM
Deleted: and the framework is not a comprehensive list of possible factors to explore.

Reporting of Background Data by the National Assessment (adopted on May 18, 2002). That policy is incorporated into this framework. It is attached in the appendix.

PRIORITIES

The following priorities for collecting and reporting non-cognitive information should be followed in planning background questionnaires, the frequency with which questions are asked, and the samples from which data are collected.

(1) ***Student reporting categories*** that are required by law must be collected as a regular component of all NAEP assessments. These include race, ethnicity, gender, socio-economic status, disability, and limited English proficiency. A core of SES information should be collected in every assessment, such as type of community and poverty status. An expanded set of SES variables may be included periodically or administered to limited samples. Efforts should be made to develop a composite measure or index of SES.

(2) ***Other factors that provide a context for results*** should be sampled periodically, or on a rotating basis, over several NAEP cycles, although a limited set may be asked in every assessment. Contextual factors may include courses taken, student mobility, school safety and discipline, teacher-related factors such as demographics and experience, other factors related to students and schools, and educationally-relevant variables outside school. Modules should be prepared for special studies to provide descriptive information on issues of current policy interest. In all cases, non-cognitive variables must be clearly related to academic achievement or to the fair presentation of achievement results.

(3) ***Subject-specific background information*** should be gathered at the same time that achievement in a subject is assessed. This may include relevant course content and requirements, teacher preparation, and other factors related to student achievement. Questions will not be designed to determine effective practices, but to show patterns and trends of factors of interest, based on previous research. Like the contextual information, most of these variables should be sampled periodically, or on a rotating basis, over several administrations of the subject exam, although a limited core set may be repeated every time the assessment is given.

Authorised User 4/25/13 2:41 PM
Deleted: Although many non-cognitive variables may be of interest, they must be limited to meet the needs of NAEP reporting.

Authorised User 4/25/13 2:41 PM
Deleted: they

SELECTION CRITERIA

Key criteria for selecting non-cognitive topics and questions are as follows:

- ***Does the current or proposed non-cognitive variable relate to the primary purpose of NAEP and how?*** The primary purpose of NAEP is to report on the academic achievement of students to the American public. It is not to report on the causes of that achievement. Other surveys with longitudinal data are far

better suited to examining causality. NAEP's choice of which non-cognitive variables to measure should be guided by how and to what extent the variables selected will support NAEP's primary mission.

- ***Do the current or proposed non-cognitive variables meet professional standards for reliability and validity?*** The NAEP legislation requires that the assessment "use widely accepted professional testing standards (P.L. 107-110, Sec. 411 (b) (5)." This requirement applies equally to non-cognitive and academic variables.

- ***How stable is the non-cognitive variable from period to period?*** If a variable shows little change from year to year, it should be reviewed to determine whether it should be deleted or used on a periodic basis rather than in every assessment.

- ***If new questions are added, have others been deleted in order to limit the burden and expense of NAEP's background questionnaires?*** There will always be pressure to collect more information. Mechanisms must be developed to make sure the burden of background questionnaires does not expand over time.

- ***Does a question address specific behavior rather than conclusions?*** Even for such questions, however, caution is advisable because self-reports are often unreliable.

- ***Will the topic or question meet the test of broad public acceptability and not be viewed as intrusive or prying?*** NAEP's non-cognitive questions are not kept secure, and all of them are to be posted on the Internet. Possible objections should be considered in deciding whether or not a question will be asked.

- ***Does the topic or question deal with a factor in which trends over time are important?***

- ***Will the information obtained be of value in understanding academic performance and taking steps to improve it?*** This is a fundamental issue to be addressed in evaluating all background questions proposed for NAEP.

Because of the value of preserving trends, consistent wording of questions should be maintained on topics of continuing interest. Changes in wording must be justified. However, as practices and circumstances change, new questions will be introduced in a timely manner to gather data on topics of current interest. NAEP shall include background questions from international assessments, such as PISA (Program for International Student Assessment) and TIMSS (Trends in International Mathematics and Science Study), to obtain direct comparisons of states and TUDA districts to educational practices in other countries.

Authorised User 4/26/13 10:59 AM
Deleted:

DATA COLLECTION

Whenever possible, NAEP should use information from school records and other reliable data collections in order to improve the validity of the information collected and limit the background questionnaires in NAEP itself. In exploring the utility of different data sources, the following criteria should be considered: (1) reliability, (2) universality, (3) currency, (4) respondent burden, (5) logistics, (6) efficiency and cost-effectiveness, and (7) the impact on timeliness of NAEP reporting.

Of the student reporting categories in Priority 1, information on gender, race/ethnicity, disability status, and limited English proficiency shall be collected in a uniform manner in all NAEP samples. NAEP is also required to collect information on socio-economic status. This will continue to be done in all samples, although there may be some variation in the number of factors on which data are obtained with a uniform core and more extensive data gathering in some cases.

Because socio-economic status cannot be measured simply or directly, NAEP has used "proxy" variables, such as eligibility for free or reduced-price lunch (a measure of poverty), parent education, and the number of reading materials in the home. The framework provides that NAEP explore development of a composite index for SES derived from information collected from students and schools. To the extent that the index can be sharpened by additional data from readily available sources, such as zip codes and the census, this option should also be considered. Occasionally and in limited samples, more extensive SES questions may be asked. Although NAEP may never be able to produce a full composite of SES, based on family income, education, and occupation, efforts should be accelerated to develop and use improved measures of socio-economic status, including an SES index.

Authorised User 4/26/13 10:26 AM
Deleted: the proxy variables

Authorised User 4/26/13 10:26 AM
Deleted: currently

Authorised User 4/26/13 11:04 AM
Deleted: should be made to find an approximation that is more informative than the current set of proxy variables.

For the past two decades, NAEP has collected information on a lengthy list of student, teacher, school, and beyond-school factors that may provide a context for achievement results and are of interest to policymakers, researchers, and the public. Yet, NAEP's design as a cross-sectional survey places serious limitations on the inferences that can properly be drawn from this information. We propose a careful review of the contextual factors in NAEP to focus on the most important variables related to public policy. All such information must be clearly related to student achievement, as shown by other research. Different questions should be cycled in and out of the assessment periodically, and the use of data from non-NAEP sources should increase. Information should be collected at meaningful intervals in ways that may show significant patterns and change over time.

The collection of subject-specific background information should be focused, limited, and prioritized as part of the subject-matter frameworks adopted by the Board. For subjects tested regularly at two-year or four-year intervals, there should be a small core set of background items administered to the full sample each time a subject is assessed. An additional, more comprehensive set of questions should be administered periodically or to smaller subsamples.

Authorised User 4/26/13 10:30 AM
Deleted: each subject

Clusters of questions should be developed so that indexes may be prepared on important topics rather than relying on individual items alone.

Detailed frameworks will be published with the theoretical rationale and research evidence that support the selection of topics and questions in background questionnaires and their connection to student achievement. Such frameworks should be updated for each assessment cycle and provide the basis for new topics and questions.

In constructing questionnaires it is important to place strict limits on the respondent burden they impose. As much data as possible should be obtained from school records and other reliable data sources. The average individual response time to answer background questionnaires for each assessment, as calculated in accordance with Office of Management and Budget (OMB) procedures, shall be limited as follows: 10minutes for each student on paper-and-pencil tests, 15 minutes per student on computer-based assessments, 20 minutes for each teacher, and 30 minutes for each school.

REPORTING

NAEP reporting should include contextual variables and subject-specific background information to enrich and give perspective to results. Consistent with space and operational limitations, descriptive information should be part of NAEP Report Cards and summary and highlights reports. The reports should present information on patterns and trends of non-cognitive variables known to have a relationship to academic achievement and may contain disaggregated data on school conditions and practices for various groups of students. Data on courses taken before NAEP assessments (either from transcripts or questionnaires) is of great public interest and can be related to academic results.

In addition, special reports should be prepared that focus on particular topics of public interest and importance. These reports should feature significant contextual information as well as cognitive results..

All background questions and data collected by NAEP should be posted on the Internet so the public may be able to consider them in discussing results. Complete data files should be made available to researchers for further analysis. In all cases, NAEP reports must not state conclusions as to cause and effect relationships.

RESEARCH

As a cross-sectional survey without longitudinal data, the National Assessment is able to document school conditions and practices. It can report on achievement results. But it cannot properly be used to establish direct cause-and-effect relationships. Still,

Authorised User 4/25/13 2:58 PM
Formatted: No bullets or numbering

Authorised User 4/25/13 2:53 PM
Deleted: NCES will prepare for Board review and approval a plan indicating the frequency, sample size, and schedule of rotation for all background variables and questions on which information is to be collected by NAEP. This should include both questionnaires and alternate data sources to obtain core reporting data, subject-specific information, and data on achievement-related contextual variables from a variety of NAEP samples—national only, national and state, and a subset of the national sample. The plan should indicate the frequency and schedule of rotation for each of the questions proposed. It should also indicate any questions needed for quality control purposes. The recommendations should be prepared with input from researchers and state policy analysts, as appropriate, and updated on a regular basis.

Authorised User 4/25/13 2:46 PM
Deleted: ten

Authorised User 4/25/13 2:46 PM
Deleted:

Authorised User 4/25/13 3:01 PM
Deleted: supplemental

Authorised User 4/25/13 3:01 PM
Deleted: may

Authorised User 4/25/13 3:02 PM
Deleted: aspects of the background data collected

Authorised User 4/25/13 3:05 PM
Moved down [1]: In all cases, NAEP reports published by the National Center for Education Statistics must not state conclusions as to cause and effect relationships and avoid simplistic presentations that imply best practice.

Authorised User 4/25/13 3:05 PM
Moved (insertion) [1]

Authorised User 4/25/13 3:06 PM
Deleted: published by the National Center for Education Statistics

Authorised User 4/25/13 3:07 PM
Deleted: and avoid simplistic presentations that imply best practice.

over the past three decades, NAEP has been part of two important research endeavors—exploring changes in the black-white test score gap since 1970 and seeking to establish the impact of state-level reforms during the 1990s. By monitoring achievement well, NAEP has provided sound data for researchers to use. NAEP results have been critical in identifying research hypotheses. Its contextual variables have added valuable information. Its large data sets have been combined with other information to tease out meaning and policy implications, though NAEP's own reports have properly steered clear of these activities.

The Governing Board believes that by doing its main task of monitoring educational achievement well NAEP can make a valuable contribution to education research. Researchers should be involved, under the auspices of NCES, in developing NAEP background questionnaires, validity studies, and other data collection efforts to carry out the provisions of this framework.

The primary purpose of NAEP is to provide fair and accurate information on student achievement. Its primary audience is the American public. The Governing Board believes that in serving its purpose and audience well, NAEP can contribute to educational research. It welcomes the interest and efforts of researchers.

Authorised User 4/25/13 5:20 PM
Deleted:

Authorised User 4/25/13 3:08 PM
Deleted: The NCES program of secondary analysis grants for researchers to analyze NAEP data should continue.

Authorised User 4/25/13 3:08 PM
Deleted: Educational researchers

Chapter One: Introduction

The National Assessment of Educational Progress is the only continuous long-term measure of student achievement in the United States in elementary and secondary schools. Its primary purpose is to report to the American public on academic achievement and its change over time.

Nature and Purpose of NAEP

The NAEP survey consists of two major components: academic assessments that measure the achievement of students on a broad range of content, and non-cognitive survey questions that collect descriptive information from students, teachers, and school administrators about demographic characteristics and the educational process. Since 1969 NAEP has measured achievement in most areas of the school curriculum, including mathematics, reading, writing, science, U.S. history, world geography, civics, economics, foreign language, computer science, and the arts. The content of NAEP assessments is determined through a framework development process that articulates the content parameters for each area and recommends subject-specific non-cognitive areas for data collection and reporting.

NAEP's purpose is to report to the public on the status of academic achievement in America. The assessment does not report results for individual students, but only for groups of test-takers having large, representative samples, e.g., students from rural schools, from various ethnic groups, or from participating states, and, on a trial basis, large urban school districts. It must be able to provide data for fair and accurate comparisons between the states and subgroups on which it reports. The background data play a crucial role in ensuring the fair comparisons—over time and between student groups—that are at the heart of NAEP's mission and value.

Nature and Purpose of Background Data

The most recent NAEP reauthorization (P.L. 107-110) gives the National Assessment Governing Board "final authority" to approve "all cognitive and non-cognitive assessment items." This framework deals with the non-cognitive side of the Board's responsibility, including the items that identify students in NAEP's required reporting categories and the other information that provides a context for results and tracks factors associated with academic achievement.

The term "non-cognitive," as used in the law, seems more inclusive than the phrase "background questions" by which the collection of non-academic information has been termed by NAEP in the past. However, non-cognitive is also less readily understandable than background information, and so the two terms will be used interchangeably in this document. Both will refer to all of the information beyond the academic assessment that NAEP uses to make its academic results more meaningful to the public.

When NAEP began, the collection of non-cognitive data was limited to the demographic categories of gender and race/ethnicity, and to two measures of home environment or socio-economic status—level of parents' education and literacy materials in the home. In addition, an index was constructed, based on data from the U.S. Census and a brief school questionnaire, to report achievement results for schools in three types of communities—disadvantaged urban, advantaged urban, and rural.

During the 1980s the use of non-cognitive questions was greatly expanded to accommodate several functions within NAEP (Reckase, 2002). First, they were used to define a more extensive array of subgroups of the student population for reporting purposes. For example, NAEP results are now reported by gender, race/ethnicity, parents' highest level of education, type of school, participation in Title I, and eligibility for free/reduced-price lunch

A second reason for collecting non-cognitive information is to inform educational policy by describing the contexts for learning, sometimes called opportunity-to-learn (Mullis, 2002). Broadly, this involves the content specified in the curriculum, whether and how that content actually is taught, students' propensity to learn, as well as home and school factors that can enhance learning.

Authorised User 4/25/13 5:28 PM
Deleted: opportunities
Authorised User 4/25/13 5:28 PM
Deleted:

In conjunction with the descriptions of students, background information about educational settings and experiences can reveal striking differences in how important aspects of education and educational resources are distributed among different groups. For example, do disadvantaged minority students have less access to science laboratory equipment than more advantaged groups? Do girls take less rigorous mathematics courses than boys? The data on course taking has been used widely to discuss the patterns and trends in mathematics achievement. Having this information as part of NAEP has added to the public impact of assessment results.

A third function of the non-cognitive questions has been to support research into factors that may be related to student achievement. The background questions serving this function have sought information not only on curriculum, teaching methods, and discipline in the school, but also on educational activities at home. For example, *The 1998 NAEP Reading Report Card* (Donahue, Voelkl, Campbell, & Mazzeo, 1999) reports on television viewing, daily reading habits, classroom reading and writing assignments, and discussion of schoolwork at home. While secondary researchers have used NAEP to investigate relationships to student achievement, the basic design of the assessment as a cross-sectional survey without longitudinal data limits its usefulness. Research has been

most productive when NAEP is combined with other data sources and in descriptive studies that track changes over time.

Non-cognitive data are also necessary to support certain technical functions of NAEP. For example, some non-cognitive information is used to evaluate the potential for bias resulting from non-participation. That is, did the students absent or refusing to participate in the assessment differ in such significant ways from those who did take part that results were changed? Non-cognitive variables also play an important role in NAEP's sampling and weighting procedures, and sometimes in checking the validity of results. Many of these variables are taken from other data sources, such as the Common Core of Data (CCD), but some come from the administration roster collected from schools prior to testing, the records kept by test administrators, and student questionnaires.

Finally, NAEP non-cognitive questions have been used in the technical process for preparing estimates of student proficiency distributions on the cognitive component of the assessment. But their role in this process is limited to facilitating data analysis. Only the student responses to cognitive questions are used to determine achievement results. Background variables are used to define the groups for which cognitive data are reported.

Once test results for a group are determined, the NAEP analytic process makes use of background data available to prepare a second data set—identical in its group scores to the first—that can be handled by much simpler computer programs to prepare other analyses and reports. However, only the background factors to be reported on are needed for this analytical work, called conditioning. The precision of NAEP results is not reduced if background items not used for reporting are eliminated.

This background information framework will focus the collection of non-cognitive information on NAEP's primary mission: providing, as the law stipulates, "a fair and accurate measurement of student academic achievement and reporting trends in such achievement" over time. Thus, the framework is a guide for gathering important information that will assist in reporting and understanding NAEP results.

Development of NAEP Background Information Framework

In the Policy Statement on Redesigning the National Assessment of Educational Progress (adopted in August 1996), the Governing Board sought to improve the validity of background information on NAEP, increase the efficiency with which it is collected, and reduce the number of background questions in the assessment itself. The statement was based on the report of a Design/Feasibility Team (Forsyth et al, 1996), headed by Robert Forsyth, which recommended a design that would rotate the collection of non-cognitive data into distinct modules administered over several assessment cycles. NAGB endorsed implementing that recommendation through a system of *comprehensive* and *standard* NAEP assessments that would be administered on a cyclical basis (NAGB, 1996).

Standard assessments would ask a short, essential core of background questions associated with a content area. Periodically, a *comprehensive* assessment would employ a much fuller complement of such questions to probe that area more extensively. Although some efforts have been made to reduce the background questionnaires and streamline data collection, the full impact of the NAGB policy has not yet been realized.

In early 2002, the No Child Left Behind Act transferred final authority over the non-cognitive questions from the National Center for Education Statistics to the National Assessment Governing Board. The Board adopted a new policy governing the development and selection of non-cognitive questions in May 2002, and initiated a process to prepare a general framework for non-cognitive data (NAGB, 2002). This framework would define the scope of NAEP background questionnaires, the priorities for collecting non-cognitive information, and the criteria for reporting non-cognitive data in NAEP. (See Appendix for full text of the policy.)

The Board created an Ad Hoc Committee on Background Questions and conducted an all-day workshop on the NAEP non-cognitive questions on September 24, 2002. Six consultants prepared and presented papers at the meeting that was attended by Board members, academic researchers, representatives of the national teacher organizations and other education groups, and NAEP contractors and staff. The six consultants are identified on the title page as contributors to this document.

In the months after the workshop, a draft framework was prepared. It was refined at several meetings of the Ad Hoc Committee, posted for public comment on the Internet, and was the subject of a public forum in Washington, D.C., on May 1, 2003. Altogether, oral comment and written testimony were received from 22 persons and organizations, many with differing perspectives and views. The Ad Hoc Committee and the Board carefully considered these comments, and the draft framework was revised at a Committee meeting on June 25. The Committee heard additional comment and made final revisions on July 31. The background information framework was reviewed by the full Governing Board several times during the course of its development. The Board adopted it unanimously on August 1, 2003.

While this framework is not a consensus document, it does encompass the thinking of a wide range of researchers, policy analysts, and users of NAEP data. It is the product of discussion and deliberation by the Governing Board, and incorporates Board decisions on the nature and focus of the background information to be included in NAEP. The framework will become operative in the 2006 National Assessment.

Requirements of NAEP Statute

The No Child Left Behind Act of 2001 (P.L. 107-110) requires NAEP to collect information on gender, race/ethnicity, socio-economic status, disability, and limited English proficiency. It must report test data on these groups, whenever feasible, that is cross-tabulated, compared, and reported according to the categories required.

The law also requires NAEP to collect only information that is directly related to academic achievement and to the presentation of such information in a fair and accurate manner. This means that NAEP needs to concentrate on variables that are known to be related to achievement rather than on theoretical constructs. The statute requires the Governing Board to ensure that all NAEP questions are "free from racial, cultural, gender, or regional bias"—a provision from previous law. But it adds new language that questions must be "secular, neutral, and non-ideological" and must not "evaluate or assess personal or family beliefs and attitudes."

In their report on the bill, the House-Senate conference committee that negotiated its final form says the law "does not preclude the use of non-intrusive, non-cognitive questions, approved by the National Assessment Governing Board, whose direct relationship to academic achievement has been demonstrated and is being studied as part of [NAEP] for the purposes of improving such achievement." The report language is not binding, but is intended to guide implementation of the law. ***This framework emphasizes that the legal prohibitions must be followed in preparing background questions and collecting any other non-cognitive data for NAEP***.

In addition, the law makes it clear that NAEP may not disclose any personally identifiable information or maintain any system of records that contains such data. These restrictions are not new. They have dictated careful procedures in the past, which must be continued.

Purpose and Rationale of Background Information Framework

The purpose of the framework on background information is similar to that of NAEP's content area frameworks: to guide the development of the assessment. The content frameworks have described the topics to be tested by NAEP and provided an outline of the assessment for each subject area. Purposefully, the frameworks attempt to be independent of a particular pedagogy. They do not specify what educational resources or processes should be used, but rather describe important achievement results. They provide states, schools, policymakers, and the public with a logical outline of the approach used in constructing the assessment.

The framework for NAEP background data will specify the parameters of the assessment from a reporting perspective. The background information that NAEP uses in its reports helps to give context and meaning to the cognitive results. It must be collected in a systematic way from the NAEP testing samples either through questionnaires or from other reliable sources, such as school records and other federal surveys. Collecting descriptive information from a variety of sources can improve the quality of the data obtained and increase efficiency while reducing the burden on respondents.

The Governing Board adopted a Policy Statement on the Collection of Reporting of Background Data on May 18, 2002 (NAGB, 2002). The statement is incorporated into this framework and attached in the Appendix.

A further statement, entitled Policy Statement on NAEP Background Questions and the Use of Contextual Data in NAEP Reporting, was adopted by the Board on August 4, 2012. It has been used in revising the framework text and has been added to the Appendix.

Chapter Two: Priorities and Criteria For Collecting and Reporting Non-cognitive Data on NAEP

This chapter presents priorities for collecting and reporting non-cognitive information on NAEP. It also includes the criteria for selecting particular topics and questions, and for determining the frequency with which various data elements are reported. A final section presents criteria for identifying and selecting background data sources.

Priorities for Non-Cognitive Information

The following priorities for collecting and reporting non-cognitive information are based on legal requirements, the purposes of NAEP, and the strengths and limitations of the assessment. They should be followed in planning background questionnaires, the frequency with which questions are asked, and the samples from which data are collected.

(1) ***Student reporting categories that are required by law must be collected as a regular component of all NAEP assessments.*** These include race, ethnicity, gender, socio-economic status, disability, and limited English proficiency. A core of SES information should be collected in every assessment, such as type of community and poverty status. An expanded set of SES variables may be included periodically or administered to limited samples. Efforts should be made to develop a composite measure or index of SES.

(2) ***Other factors that provide a context for results should be sampled periodically, or on a rotating basis over several NAEP cycles, although a limited set may be asked in every assessment.*** Contextual factors may include courses taken and course requirements, student mobility, school safety and discipline, teacher-related factors such as teacher demographics, preparation, credentials, and experience, and other factors related to students, schools, and educationally-relevant variables beyond the school. Modules should be prepared for special studies to provide descriptive information on issues of current policy interest. In all cases, non-cognitive variables must be clearly related to academic achievement or to the fair presentation of achievement results.

Authorised User 4/26/13 10:48 AM
Deleted: ,

Authorised User 4/26/13 10:46 AM
Deleted: Although these types of non-cognitive variables are of interest, they must be limited so that they meet the needs of NAEP reporting

Authorised User 4/26/13 10:46 AM
Deleted: they

(3) ***Subject-specific information may be gathered at the same time that academic achievement in a particular area is assessed***. This may include relevant course content and requirements, teacher preparation, and other factors related to achievement in the subject assessed. Questions will not be designed to determine effective practices, but to show the patterns and trends of factors of interest, based on previous research. Like other contextual information, most of these variables should be sampled periodically, or on a rotating basis, over several administrations of the subject exam, although a limited core set may be repeated every time the assessment is given.

With regard to the points above, Walberg (2002) makes a suggestion that might be a workable solution to consider. Just as students in the NAEP samples do not respond to all the questions, say, in reading, but only to a portion of those for any one grade-level, so too, the non-cognitive questions could be rotated through different (smaller) NAEP samples. These non-cognitive "testlets" could be rotated through the NAEP samples by class or school, with students receiving different, expanded "testlets" in addition to a core set of background questions.

Criteria for Selecting Non-cognitive Topics and Questions

The Advisory Council on Education Statistics (ACES), a technical panel that used to advise the National Center for Education Statistics, spent a considerable amount of effort on the issue of NAEP non-cognitive questions. Its guidelines, adopted in May 1997, include a set of key questions that should be utilized in selecting topics and questions for NAEP background data collection. The questions with commentary are summarized below:

- ***Does the current or proposed non-cognitive variable relate to the primary purpose of NAEP and how?*** The primary purpose of NAEP is to report on the academic achievement of students to the American public. It is not to report on the causes of that achievement. Other surveys with longitudinal data are far better suited to examining causality. NAEP's choice of which non-cognitive variables to measure should be guided by how and to what extent the variables selected will support NAEP's primary mission.

- ***Do the current or proposed non-cognitive variables meet professional standards for reliability and validity?*** The NAEP legislation requires that the assessment "use widely accepted professional testing standards (P.L.107-110, Sec. 411 (b) (5)." This requirement applies equally to non-cognitive and academic variables. It is already known that some non-cognitive variables in NAEP have weak reliability (e.g., data from 4th graders on their parents' highest level of education and the self-reports of teachers on classroom practice). If more reliable sources of such data cannot be found, these variables should be deleted from the assessment.

- ***How stable is the non-cognitive variable from period to period?*** If a variable shows little change from year to year, it should be reviewed to determine whether it should be deleted or used on a periodic basis rather than in every assessment.

- ***Is the proposed or current non-cognitive variable of timely interest?*** The educational environment changes from time to time, and consequently public interest in particular variables will change as well. It would serve NAEP well to review the set of non-cognitive variables periodically with this criterion in mind, deleting those that do not meet the test of timeliness and substituting others of current interest.

- ***If new questions are added, have others been deleted in order to limit the burden and expense of NAEP's background questionnaires?*** There will always be pressure to collect more information. Mechanisms must be developed to make sure the burden of background questionnaires does not expand over time.

- ***Does a question address specific behavior rather than conclusions?*** For example, a question that asks teachers whether they adhere to national standards in mathematics or another subject is conclusionary and hard to interpret, since many teachers are apt to say yes, regardless of what they do. It would be better to ask about specific behaviors, such as homework assignments or computer use. Caution is advisable in this area too because self-reports are often unreliable.

The Board believes three other important criteria must also be considered:

- ***Will the topic or question meet the test of broad public acceptability and not be viewed as intrusive or prying?*** NAEP's non-cognitive questions are not kept secure and must readily be available to anyone requesting a copy. Under Board policy, all questions asked are to be posted on the Internet. Possible objections should be considered in deciding whether or not to ask them.

- ***Does the topic or question deal with a factor in which trends over time are of importance?*** If trends are deemed important and the factor is related to achievement, the topic or question should be included periodically on a four-year or eight-year cycle, rather than being part of the background questionnaire each year. For example, measuring television watching in every NAEP assessment is not necessary. But it can be valuable to measure TV-watching every four or eight years to find out whether or not it is increasing.

- ***Will the information obtained be of value in understanding academic performance and taking steps to improve it?*** This is a fundamental issue to be addressed in evaluating all background questions proposed for NAEP.

Because of the value of preserving trends, consistent wording of questions should be maintained on topics of continuing interest. Changes in wording must be justified.However, as practices and circumstances change, new questions will be introduced in a timely manner to gather data on topics of current interest. NAEP shall include background questions from international assessments, such as PISA (Program for International Student Assessment) and TIMSS (Trends in International Mathematics and Science Study), to obtain direct comparisons of states and TUDA districts to educational practices in other countries.

Authorised User 4/26/13 11:13 AM
Formatted: Indent: Left: -0.06", First line: 0.56"

Criteria for Selecting Data Sources

NAEP has collected non-cognitive information from students, teachers, and schools, using NAEP background questionnaires. There are also administration rosters, completed by test administrators at the school level in advance of testing to determine characteristics of the testing samples. The Common Core of Data (CCD) is used to identify characteristics of schools (e.g., Title I funding), and schools also complete a questionnaire on special needs students (e.g., students with disabilities and limited English proficiency).

However, the collection of non-cognitive data may be shifted among these sources or to new sources in order to improve reliability, increase efficiency, or reduce burden. State management information systems and data collected for school report cards, have become increasingly useful for NAEP. ***Whenever possible, NAEP should use information from school records and other reliable data collections about students and schools in order to improve the validity of the information collected and limit the background questionnaires in NAEP itself.***

Authorised User 4/26/13 11:14 AM
Deleted: as required by the No Child Left Behind Act, may

Authorised User 4/26/13 11:15 AM
Deleted: very

In exploring the utility of different data sources, the following criteria should be considered:

- ***Validity*** – Is the data obtained from the new source a valid indicator of what it purports to measure?
- ***Reliability*** – Is the data from the new source at least as reliable and consistent as that from the source previously used?
- ***Universality*** – Can the required data be collected by this method for all (or almost all) of the students and schools participating in NAEP and will it support valid comparisons over time?
- ***Currency*** – Will data obtained from a new data source be current enough to relate clearly to the assessment being conducted? If data from the census or some other source is several years old it may not accurately describe school or neighborhood conditions at the time of testing.

- ***Respondent Burden*** – Will the new source(s) reduce the burden on students, teachers, and schools in filling out NAEP questionnaires? Will the total amount of respondent burden be decreased?

- ***Logistics*** – Will the alternative source(s) be logistically possible, or will there be more logistical problems than with the previous data source? Logistics includes such considerations as cost, time, administrative personnel resources, and steps needed to ensure accurate coding and data analysis.

- ***Efficiency and cost-effectiveness*** – How efficient will the new data source be in comparison to the previous one? For example, it may be more efficient to collect data from a state management information system about the state's schools, teachers, or students, rather than obtaining it from the test samples directly, but efficiency and cost-effectiveness should be determined before a change is made.

- ***Timeliness of NAEP reporting*** – How will a change in data sources affect the speed with which NAEP can be reported? Some changes will speed operations, but those that slow down NAEP reporting are not desirable.

Chapter Three: Topics and Types of Background Data

This chapter will cover the non-cognitive topics that are required for reporting under the No Child Left Behind Act of 2001 (P.L. 107-110), as well as those that should be considered for inclusion in NAEP on a cyclical basis. It discusses socioeconomic status (SES), contextual factors of interest to public policy, and subject-specific variables.

Demographic Reporting Categories

The demographic variables collected by NAEP come from two sources. Information is obtained from school records on gender, age, race/ethnicity, and two elements of socio-economic status (SES)— participation in Title I and eligibility for free or reduced-price lunch, which is based on family income. The shcool records are also used to indicate whether a student is classified as disabled or limited English proficient. In addition, data on race/ethnicity is alsocollected on the NAEP student questionnaire, and students are asked to report on the highest level of each parent's education and on several aspects of home enviroment, including number of books, internet access, and whether they have their own bedroom.

Authorised User 4/26/13 11:37 AM
Deleted: currently

Authorised User 4/26/13 11:30 AM
Deleted: are

Authorised User 4/26/13 11:43 AM
Deleted:

Authorised User 4/26/13 11:37 AM
Deleted: In addition, information is obtained on disability status and on students who are classified as

Authorised User 4/26/13 11:27 AM
Deleted: All of this information is collected on an administration roster, completed from school records in advance of testing

Authorised User 4/26/13 11:46 AM
Deleted:

Authorised User 4/26/13 11:40 AM
Deleted:

A more extensive questionnaire is completed by school staff on each student selected for NAEP who is classified as either disabled or limited English proficient (LEP). For students with disabilities (SD), the questionnaire collects data on the specific disability and its severity, the student's Individualized Education Plan (IEP), type of curriculum, whether the student participates in standardized testing (with our without accommodations), and the accommodations allowed on state and district standardized tests in presentation, response, setting, and timing. For LEP students, the questionnaire covers native language, number of years of academic instruction in English, percent of instruction in English and/or native language, and the testing accommodations provided under district or state policy. In the future, NAEP might also identify students who recently exited from LEP programs and track their achievement.

NAEP is required to collect information on all of these categories (except age), but has some discretion in determining definitions and aggregating responses. These data will continue to be collected in a uniform manner in every NAEP assessment, although, for socio-economic status, as explained in the section below, there may be some variation, with a uniform core and more extensive data-gathering in some cases.

Socio-economic Status (SES)

Under current law, NAEP is required to collect information on socio-economic status. SES also is clearly a factor that has been shown to be related to academic achievement in many research studies, beginning with the Equality of Educational Opportunity Commission Report (Coleman et al., 1966). The research community's consensus over the past four decades has been to deal with the influence of SES on other achievement-related variables by holding SES constant while examining the other effects, for example, adjusting for SES while looking at effects of class size or teacher training. NAEP does not adjust for SES, but it does report on the relationship between student achievement and SES proxy variables like parents' education or Title I participation.

NAEP has not been able to measure SES directly, using its present set of questions and data sources, i.e., the student, teacher, and school questionnaires. The assessment has used "proxy variables" for SES, including students' eligibility for the National School Lunch program, participation in Title I, parents' education, and the number of reading materials in the home (newspapers, magazines, books, etc.)—information on the latter two factors being reported by students in the assessment samples. In addition, NAEP uses census data to classify schools into different types of location, based on Census Bureau definitions, such as central city, suburban/large town, and rural/small town. The questions on newspapers and magazines were dropped in the mid-2000s as circulation dwindled, and were replaced by an item on internet access.

Strictly speaking, these are individual proxy variables and are not combined into a composite variable. However, both the questions on parent education and home environment have been coded in a pseudo-composite manner. For example, the parent education related to the student is the higher of either the mother's or father's education level. On the four home environment questions used until the mid-2000s student responses were coded differently for a "yes" answer to two questions or fewer, "yes" to three questions, and "yes" to four questions, as well as omitted responses (Allen, Carlson, & Zelenak, 1999).

Authorised User 4/26/13 11:52 AM
Deleted: are

Authorised User 4/26/13 11:56 AM
Deleted: ,

Authorised User 4/26/13 11:56 AM
Deleted: are

At the lower grade levels, students' reports of their parents' education are questionable at best, while the National School Lunch program sorts students only into three categories (Yes, No, and Unknown) and Title I into two categories (Yes or No). For many years, NAEP used a reporting category of disadvantaged urban schools, constructed from information provided by school principals. This was discontinued in the mid-1990s because the category lacked a consistent definition from year to year and between different state samples. There also were serious doubts about the reliability of the information on which it was based. The data on eligibility for the National School Lunch Program have also become increasingly problematic because of expansion of the program and administrative changes allowing whole-school or whole-district eligibility in high-poverty areas. In short, there has been considerable concern over many years about the quality of the SES measures in NAEP, both for reporting to the public and for analysis by researchers.

Barton (2002) suggests two alternative approaches for improvement: (1) a composite index for SES, or (2) a parent questionnaire. A composite index is viable using the same information that is currently collected in NAEP, or perhaps augmented with a few targeted questions or census data, possibly the zip code of student home addresses. ***The necessary analytical work should be initiated through small research studies using extant NAEP data sets in order to check systematically the validity of a composite index as a better measure of SES in NAEP samples. The results could vary by grade level, in which case, adjustments might be needed in the way the data are collected, augmented, and/or confirmed. NAEP may never be able to produce a full composite of income, education, and occupation, but efforts should be accelerated to develop and use improved measures of socio-economic status, including an SES index.***

In November 2012, an expert panel convened by the National Center for Education Statistics recommended prompt development of an SES composite measure.

Authorised User 4/26/13 12:13 PM
Deleted: *made to find an approximation that is more reliable than the current set of individual proxy variables.*

Authorised User 4/26/13 12:20 PM
Formatted: Font:Not Bold, Not Italic

The argument in favor of this approach is that it advances the goals of the current law without impacting data collection in unforeseen ways. Barton suggests that such an index would enable NAEP to report results in terms of SES quartiles (much the same way that the National Educational Longitudinal Survey, NELS, does). Further, it would allow the assessment to report cross-tabulations on distributions of students in the NAEP achievement level categories by SES. A good measure of SES would improve the monitoring of achievement gaps among various racial/ethnic groups, although sample sizes may not be large enough within all ethnic groups or types of schools. Finally, a composite SES index may be beneficial to states and districts in the Trial District Assessment (TUDA), enabling NAEP to compare the performance of groups of students with the same socio-economic status, which is a factor of high public and policy interest.

The argument against such an approach is that SES would continue to be measured indirectly, i.e., by using proxy variables, albeit through a composite index. There would also be disagreements about precisely which variables to include in the index and how to weight different factors. For example, Armor (D. J. Armor, personal communication, December 18, 2002) has suggested that two variables deleted from the NAEP student questionnaire in 2000 be reinstated, namely, the number of siblings in the home and family status (student lives with both parents, mother or father, neither). These variables were dropped because of concerns about intrusiveness, but they may be of considerable importance in constructing an SES index. The item on number of parents in the home was restored in 2013.The Board will have to weigh the considerations involved, and may decide there is value in using them periodically or in limited samples.

Authorised User 4/26/13 12:14 PM
Deleted: recently

A parent questionnaire has been proposed as a more reliable means of collecting SES data than relying on student reports, school records, or census data. Other National Center for Education Statistics surveys, for example, NELS and the Early Childhood Longitudinal Study, have employed parent questionnaires that ask direct questions regarding occupation and income.

However, the National Assessment of Educational Progress involves far more students than any of these research surveys. Accordingly, a parent questionnaire on

NAEP would entail far more respondent burden and might arouse more controversy, making it more difficult to accomplish the primary mission of the assessment to measure student achievement. A parent questionnaire has been considered by NAGB in the past, but rejected as too burdensome and intrusive. Because these considerations are still persuasive, particularly as the scope of NAEP has expanded, no work should be undertaken on developing a parent questionnaire.

In sum, because of its importance and the requirements of law, information on socio-economic status must be collected in all NAEP samples, although there may be some variation in the number of factors on which data are obtained. Efforts should be made to develop a composite measure or index of SES based on school records and the student questionnaire. To the extent that an index can be sharpened by additional information from readily available sources, such as zip codes and/or census data, this option should be considered as well.

A core of SES information should be collected in every assessment, such as type of community (e.g., central city, rural, etc.), poverty status (e.g., eligibility for free or reduced-price lunch and Title I participation), reading materials in the home, and level of parent educationSteps must be taken to ensure that such data are reliable. Additional SES variables may also be included such as number of siblings and parents at home, possession of computers, and parent occupation. Periodically, an expanded set may be administered.

Public Policy Contextual Factors

For the past two decades NAEP has collected information on student, teacher, school, and beyond-school factors that are of interest to policymakers and the public. For students, some of these factors have included course-taking patterns, TV-watching, homework, and use of computers. For teachers, the contextual factors have included educational background, credentials, years of experience, and participation in professional organizations, to name a few.

The lists of factors have been long. They have become burdensome both to respondents and to the efficient scoring, analysis, and reporting of the NAEP survey. The way they have been reported—through simple one-way tabulations—has encouraged unwarranted conclusions about cause-and-effect relationships.

We propose a careful review of the contextual factors on which information is collected by NAEP to focus on the most important variables related to public policy. All such information must be clearly related to student achievement, as shown by other research. Modules should be prepared for special studies to provide descriptive information on issues of current policy interest. Data should be collected at meaningful intervals in ways that may show significant patterns and change over time.

Two documents are helpful in surveying the research base and presenting alternatives for NAGB to consider. The first is *Monitoring School Quality: An*

Authorised User 4/26/13 12:42 PM
Deleted: *Research*

Authorised User 4/26/13 12:42 PM
Deleted: *conducted*

Authorised User 4/26/13 12:42 PM
Deleted: *into creating*

Authorised User 4/26/13 12:54 PM
Moved (insertion) [2]

Authorised User 4/26/13 12:55 PM
Deleted:

Authorised User 4/26/13 12:45 PM
Deleted: *n—though steps*

Authorised User 4/26/13 12:48 PM
Deleted: *An expanded set of*

Authorised User 4/26/13 12:45 PM
Deleted: *be included*

Authorised User 4/26/13 12:46 PM
Deleted: *periodically and*

Authorised User 4/26/13 12:46 PM
Deleted: *administered to limited samples, including*

Authorised User 4/26/13 12:56 PM
Deleted: *NAEP should explore the use of an SES index derived from proxy variables currently in either the administration roster or student questionnaire. To the extent that an index can be sharpened by additional information from readily available sources, such as zip codes and/or census data, this option should be considered as we* ... [1]

Authorised User 4/26/13 12:54 PM
Moved up [2]: *To the extent that an index can be sharpened by additional information from readily available sources, such as zip codes and/or census data, this option should be considered as well.*

Authorised User 4/26/13 1:13 PM
Deleted: *Data should be obtained periodically, on a rotating basis, over several NAEP cycles, although a limited set of factors may be included in every assessment*

Authorised User 4/26/13 1:14 PM
Deleted: *Information*

Indicators Report (Mayer, Mullens, & Moore, 2001), prepared by Mathematica Policy Research, Inc. for NCES. This report presents a research synthesis, indicating factors for which there is a research base showing a strong relationship to academic achievement. The synthesis, involving a review panel as well as statistical analyses, identifies the following as factors related to student results: the academic skills of teachers, teacher assignments (such as out-of-field teaching), course content, student discipline and school safety, class size, and focus on academic achievement. Other sources of information are available on all of these factors, but only through NAEP can they be related to the achievement of broad groups of students over time.

The second document, *Making Connections* (Greenberg, Stancavage, Farr, & Bohrnstedt, 2001), was prepared for NCES by the American Institutes for Research and presents an elaborate typology of non-cognitive variables that could be measured by NAEP. It is organized into seven broad categories of non-cognitive information related to students, instructional content and practice, teachers, schools, school community factors, beyond school factors, and federal, state, and district policy. The listing goes beyond what NAEP can and should handle, but its discussion is thoughtful and the document is useful for planning.

Subject-Specific Background Data

For each subject assessed by NAEP, additional subject-specific background information has been collected from students, teachers, and schools. These data fall into the broad category of instructional content and practice. Under that umbrella come such topics as the curriculum taught, course offerings, class management and style, ability grouping, and modes of instruction. Subject-specific data collection has expanded enormously over the past two decades, and in recent years has included five to ten minutes of questions for students, about 30 minutes of questions for teachers, and 30 to 45 minutes for school administrators.

These questions should be focused, limited, and prioritized. Future subject-matter frameworks adopted by the Governing Board should spell out clearly what these priorities will be.

Authorised User 4/26/13 1:15 PM
Deleted: Now is the time for t

Authorised User 4/26/13 1:15 PM
Deleted: to

The design for doing this was presented to the Board in the 1996 report of a Design/Feasibility Team of prominent researchers (Forsyth, R. et al, 1996). The group recommended that a core set of non-cognitive questions should be administered to students each time a subject is assessed by NAEP. In addition, a more comprehensive questionnaire would be given whenever a new framework is introduced and repeated every eight to ten years. Similar patterns should be established for the school and teacher questionnaires.

Authorised User 4/26/13 1:18 PM
Deleted: For example, an extensive set of background questions in reading and mathematics (grades 4 and 8) was administered in 2003, the baseline year for the No Child Left Behind legislation. Another complete set should be administered in mathematics in 2005 and in reading in 2007, the years in which revised frameworks are first used, and then should be repeated at an interval of eight years. In the intervening years, only the more limited core modules will be administered.

The NAEP assessments in other subjects, such as writing, science, history, geography, and civics, should have a core set of non-cognitive questions administered to the full sample, with longer, more extensive questionnaires being administered to smaller

sub samples. With states required to participate in NAEP every two years, the total number of students tested has expanded substantially. This makes even more compelling the case for limiting the NAEP background questionnaires and rotating the background questions.

Clusters of questions should be developed so that indexes may be prepared on important topics rather than relying on stand-alone items only.

Authorised User 4/26/13 1:18 PM
Deleted: now

Authorised User 4/26/13 1:16 PM
Deleted: ***NCES should prepare for Board review and approval a plan indicating the frequency, sample size, and schedule of rotation for all background variables and questions on which information is to be collected by NAEP. This should include both questionnaires and alternate data sources to obtain core reporting data, subject-specific information, and data on achievement-related contextual variables from a variety of NAEP samples—national only, national and state, and a subset of the national sample. The plan should indicate the frequency and schedule of rotation for each of the questions proposed. It should also indicate any questions needed for quality control purposes. The recommendations should be prepared with input from researchers and state policy analysts, as appropriate, and updated on a regular basis.***
Data Collection Year for
Data Collection Year for
Reading
2003, 2007, 2013
2005, 2009, 2011
Foreign Language (12)
2004, 2012
... [2]

Chapter Four: Non-cognitive Data Sources and Collection

This chapter discusses the sources of non-cognitive information for NAEP and the reporting categories that the information describes. It includes a NAEP Background Information Matrix, organized by priorities, which summarizes the types of descriptive information NAEP collects, reporting units, and data sources.

NAEP Student, Teacher, and School Samples

The NAEP student samples vary in size and purpose. Their overall total has become very large. Starting in 2003, national NAEP samples are specified at the state and jurisdictional levels, with approximately 3,000 students per subject and grade (4 and 8 only) for each of the 50 states, plus the District of Columbia, and Department of Defense domestic and overseas schools. Puerto Rico (in mathematics only) has a sample of about 3,000. In addition, the ten Trial Urban District Assessment (TUDA) districts have sample sizes of the order of 3,000 to 5,000 each. There also are a nationally-representative sample of charter schools, totaling about 3,000 students, and national private school samples totaling about 12,000 in each grade.

At grade four, therefore, the total NAEP sample approximates 436,000 students. The grade eight sample is about the same at 432,000 (excepting charter schools). The grade 12 sample is for a pilot test and includes only about 6,000 students (Rust, 2002). In most future years the twelfth grade samples are expected to have about 30,000-40,000 students assessed in national samples only for three subjects.

In addition to the nearly one million students tested, about 80,000 teachers of those students complete teacher questionnaires and some 13,000 schools complete school questionnaires. Several thousand school districts also supply data for the assessment. The sampling and weighting procedures in NAEP use data from the CCD files as well as

census data and school-level achievement data from the states for improving NAEP stratification procedures. The NAEP non-cognitive data collection effort is enormous and challenging.

Other Data Sources

The Governing Board is strongly committed to improving the quality of background information while reducing respondent burden and the complexity of data collection and analysis. The self-report questionnaires given to students, teachers, and schools are sometimes burdensome to fill out, labor-intensive to collate and analyze, and subject to concerns about reliability. All questionnaires should be scrutinized to replace as many items as possible with data from centralized records, gathered by test administrators, or, ideally, from computerized data files.

The data available from federal, state, district, and school records should be carefully explored. In recent years much more information has become available in standardized computer formats. Barton (2002) has suggested some specific sources of data collected outside of NAEP that should be considered to improve NAEP reporting. These include the U.S. Census, Quality Education Data, Inc. (QED), and the Common Core of Data (CCD) and School and Staffing Survey (SASS), both compiled by the National Center for Education Statistics.

Authorised User 4/26/13 1:19 PM
Deleted: With implementation of the school report card requirements of the No Child Left Behind law

Authorised User 4/26/13 1:19 PM
Deleted: should be

Authorised User 4/26/13 1:20 PM
Deleted: soon

This approach of utilizing more data from outside specific NAEP data collections has been elaborated on extensively in the most recent evaluation of NAEP by the National Academy of Sciences (Pellegrino, J.W., Jones, L.R., & Mitchell, K.J., 1999). The panel proposed "a coordinated system of indicators for assessing educational progress, housed within NCES and including NAEP and other currently discrete, large-scale data collections (p. 34)." Figure 1 is reprinted from the NAS report to show the extent of these data collections on students, teachers, and schools, and to indicate what might be obtained from these other sources. To use them for NAEP would greatly lessen the burden on the assessment itself. Merged data sets could be made available, some to the general public, and more to researchers in restricted data files.

For many years state-level NAEP reports have included appropriate collateral data that provide a context for interpreting NAEP results; see for example the *NAEP 1996 Mathematics: Report Card for the Nation and the States* (Reese et al., 1997). These state contextual variables have included enrollment in elementary and secondary schools, poverty status of children from 5 to 17 years old, number of children receiving disability services, per-pupil expenditures, pupil-teacher ratios, and average teacher salaries. To the extent that these data are readily available and are helpful in setting a context for interpretation of NAEP results the practice ought to be continued. However, more effort should be made to ensure that such data are up-to-date and easily-accessed as part of NAEP reporting on the Internet.

Authorised User 4/26/13 1:21 PM
Deleted: as

Authorised User 4/26/13 1:21 PM
Deleted: as possible.

Data Collection and Reporting on Time and Learning in International Assessments

All three major international assessments of K-12 students asked students and school principals questions related to learning time in and out of school in their most recent administrations. PIRLS[1] 2011, which assessed reading at grade 4, and TIMSS[2] 2011, which assessed mathematics and science at grades 4 and 8, also asked teachers questions about student learning time. PISA[3] 2012, which assessed 15-year-old students on mathematics, reading, science, and financial literacy and problem solving, included a teacher questionnaire only as an optional component.

Table 1, below, presents a comparison of NAEP and the international assessments on learning time questions for NAEP reading and mathematics for 2013 and the most recent administrations of PIRLS, TIMSS, and PISA. Results are presented separately for the subtopics, Student Absenteeism, Instructional Time, Homework, and Other. The international assessments generally collected more in-depth information than did NAEP.

International assessment reports include country comparisons on hours of instruction (per week or per year), the extent to which student absenteeism is considered a hindrance to learning, hours of homework per week, and some measures of out-of-school learning. Associations between learning time and achievement have been limited, in part because the purposes (enrichment or remedial) of additional instructional or out-of-school learning time are difficult to discern.

[1]PIRLS is the Progress in International Reading Literacy Study, an assessment of reading at grade 4. PIRLS is coordinated internationally by the International Association for the Evaluation of Educational Achievement (IEA).

[2]TIMSS is the Trends in International Mathematics and Science Study, an assessment of mathematics and science at grades 4 and 8. TIMSS is coordinated internationally by the International Association for the Evaluation of Educational Achievement (IEA).

[3]PISA is the Program for the International Student Assessment, an assessment of 15-year-olds in mathematics, reading, and science literacy, as well as occasional other subjects. PISA is coordinated internationally by the Organization for Economic Cooperation and Development (OECD).

Table 1: Comparison of NAEP and the International Assessments on Time and Learning

Topic and respondent	NAEP	PIRLS (Grade 4)	TIMSS (Grs. 4 & 8)	PISA (15-year-old)
Student absenteeism				
Schools	Percent absent on an average day	Extent to which tardiness and absenteeism a problem		Extent to which tardiness, absenteeism, skipping classes a problem
Teachers	--	--	--	N/A
Students	--	Frequency of absences		Frequency of absences, tardiness, class-skipping
Instructional time				
Schools	--	Minutes of instruction students receive per day, days per week, days per year		--
Teachers	Minutes students receive instruction in a given subject in a typical week	Minutes students receive instruction in a given subject in a typical week; how much reading time in other subjects	Minutes students receive instruction in a given subject in a typical week	N/A
Students	Are you taking more than one mathematics course (grade 8)	--	--	Minutes per class, classes per week in mathematics, reading, science
Homework				
Schools	--	--	--	--
Teachers	Minutes of mathematics homework per day	How often homework assigned, how many minutes per day		N/A
Students	Pages of homework per day	How often homework assigned, how many minutes per day		Hours per week spent on homework
Other				
Schools	--	--	--	--
Teachers	--	--	--	--
Students	How often they read for fun and how often they receive tutoring or other help outside of regular instruction	How often they read for fun, how often they participate in various other reading activities, and if they are studying something outside of school	If they are studying something outside of school	Participation in math club and competitions, other academic clubs, lessons or classes outside school hours, purposes of lessons or classes, hours per week studying w/ tutors, commercial companies, parents, and software, participation in other out-of-school learning