National Assessment Governing Board

Ad Hoc Committee on NAEP Background Information

February 28, 2013
2:00 p.m.-4:00 p.m.

AGENDA

2:00 – 2:10 pm	Welcome and Introductions *Terry Holliday, Committee Chair*	Attachment A
2:10 – 2:25 pm	Measuring and Reporting on Socio-Economic Status: Follow-up and Response to NCES Expert Panel Paper *Mary Crovo, NAGB Staff* *William Ward, NCES*	Attachment B
2:25 – 3:00 pm	Collection and Reporting of Background Data in International Assessments—PISA, TIMSS, and PIRLS a. Socio-Economic Status b. School Resources and Practices c. Student Behaviors and Out-of-School Activities *Daniel McGrath, NCES*	Attachment C
3:00 – 3:15 pm	Background Questions on Technology and Engineering Literacy (TEL) Assessment *Mary Crovo, NAGB Staff* *William Ward, NCES*	Attachment D
3:15 – 3:30 pm	Exploratory Analyses of NAEP Data *Alan Ginsburg, Consultant*	Attachment E
3:30 – 3:55 pm	Review of NAEP Background Information Framework (Adopted 2003)—Initial Discussion *Larry Feinberg, NAGB Staff* *James Deaton, NCES*	Attachment F
3:55 – 4:00 pm	Next Steps *Committee Members*	

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

Brent Houston

Dale Nowlin

Joseph O'Keefe, S.J.

Susan Pimentel

Leticia Van de Putte

Adopted unanimously – 08/4/12

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.

IMPROVING THE MEASUREMENT OF SOCIOECONOMIC STATUS (SES) FOR THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS: A THEORETICAL FOUNDATION

An Update on Stakeholder Feedback

Submitted to the National Assessment Governing Board
by the National Center for Education Statistics

NCES is actively reaching out to key industry stakeholders (e.g., National Forum on Education Statistics, American Educational Research Association) to solicit their feedback regarding the white paper titled *Improving the Measurement of Socioeconomic Status (SES) for the National Assessment of Education Progress: A Theoretical Foundation*.

Recently, this white paper was presented to the National Forum on Education Statistics' Alternative SES Measure Working Group. The Forum's purpose is to improve the quality, comparability, and utility of elementary and secondary education data by:

- developing and recommending strategies for implementing national education standards useful to education decision making;
- improving the collection and dissemination of education data; and
- addressing emerging issues affecting education data policy.

The Forum's membership consists of representatives of state and local education agencies, offices of the U.S. Department of Education, Regional Educational Laboratories (RELs), and other federal agencies and national associations.

Feedback from the Forum's Alternative SES Measure Working Group regarding this white paper will be presented at the meeting.

Council of Chief State School Officers (CCSSO)-
National Assessment Governing Board Policy Task Force

Views on Socio-Economic Status in NAEP

January 8, 2013

This document lists issues and suggestions from the Task Force's discussion at the January 8, 2013 in-person meeting.

Task Force Discussion: White Paper on Socio-Economic Status in NAEP Reporting

The Governing Board received a report (www.nagb.org/publications/expert-panelnaep-bq-report.pdf) from an expert panel at the March 2012 Governing Board meeting, and established the Ad Hoc Committee on NAEP Background Information later in 2012. One of the issues the Ad Hoc Committee is considering is how to improve the measurement of socioeconomic status (SES). NCES has convened a group of experts to provide focused guidance on this topic for the NAEP context and has also asked the Task Force for feedback. The following provides feedback from the Task Force.

The purpose of SES data collection

- The language in the report states that SES "could include" certain variables, but does not provide a definitive meaning of SES. At some point, it will be helpful to be more definitive as to what SES means. In order to do that, we need to consider the question of the purpose of SES, which the paper does begin to address. However, the purpose is not yet clear.

- Consider: what is the decision to improve learning and teaching that we are trying to achieve by changing the definition of SES or by collecting various data related to SES? This should inform which data are collected.

Identifying appropriate elements to include in SES

- It is important to develop a definition for SES that is strong yet also one for which we are able to collect valid and reliable data. We could develop a great definition but it will not be helpful for this work unless we are able to collect appropriate data in an accurate and cost-effective manner.

- It is important not to confound the individual (or home) notion of socioeconomic status and community socioeconomic status. If we put these two components into the same variable, we lose the ability to disentangle where policy can impact improving schooling and where there are factors upon which we are not able to have a policy impact.
    - Keeping individual and community elements of SES separate helps to support understanding of what should be done to maximize the potential of these students given specific factors.

- Rather than confounding the measures, we want a better measure of what resources are available in a student's home.

Other considerations in measuring SES

- *International context.* What composite can we develop about which we could say these are the things that really have an influence on student performance? Some factors that are not included in the paper (running water, electricity, moped, etc.) are more focused on access and how the factors will help students be ready to attend schools. In many countries, free and reduced lunch is a given so this variable is not a useful consideration. The meaning of the factors we are considering is very context dependent. As we consider what to include in the U.S., we should focus on the factors that prepare students or help them to be better prepared (access to a computer, for example). There is a significant shift in previous variables collected compared to factors that are relevant in the digital age.

- *Poverty density in a community.* There is significant research about the role of the density of poverty in a community (as opposed to an individual family's status alone). The approach would be very different in terms of what can be changed depending on whether it is a high poverty density community or a situation where one family is very poor living in a community that is not very impoverished. The high-density poverty measure has been used as one way to define high-risk students in some states.

- *Data from self-reporting.* We also have to remain cognizant of the difficulty of self-reporting, both at the young (4^{th} graders) and older (12^{th} graders) ages (though for different reasons).

National Assessment Governing Board
Council of Chief State School Officers
Policy Task Force

As part of the Board's continuing outreach efforts, the Governing Board contracted with the Council of Chief State School Officers (CCSSO) in September 2007 to form a Task Force charged with providing state feedback and recommendations to the Board on NAEP policy areas and projects. The Task Force consists of 12 high-level state education agency staff members who were chosen based on expertise and interest in assessment, and geographic representation of the nation. Task Force members include:

- 1 chief state school officer
- 5 deputy superintendents
- 3 associate superintendents of accountability and assessment
- 3 public information officers

Policy Task Force Members (2012-2013)

Patricia Wright, Task Force Chair
Superintendent of Public Instruction
Virginia Department of Education
Richmond, VA

Deborah Sigman, Task Force Vice Chair
Deputy Superintendent
California Department of Education
Sacramento, CA

David V. Abbott
Deputy Commissioner/General Counsel
Rhode Island Department of Elementary and Secondary Education
Providence, RI

Liza Cordeiro
Executive Director, Office of Communications
West Virginia Department of Education
Charleston, WV

Tom Foster
Director of Assessment
Kansas State Department of Education
Topeka, KS

Pete Goldschmidt
Director of Assessment
New Mexico Public Education Department
Santa Fe, NM

Dan Hupp
Director of Student Assessment
Maine Department of Education
Augusta, ME

Susie Morrison
Deputy Superintendent/Chief of Staff
Illinois State Board of Education
Springfield, IL

Nate Olson
Communications Manager
Washington Office of Superintendent of Public Instruction
Olympia, WA

Joel Thornton
Chief of Staff
Georgia Department of Education
Atlanta, GA

Joyce Zurkowski
Director of Student Assessment
Colorado Department of Education
Denver, CO

IMPROVING THE MEASUREMENT OF SOCIOECONOMIC STATUS FOR THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS: A THEORETICAL FOUNDATION

Contents

Expert Panel

Charles D. Cowan
Partner and Chief Executive Officer
Analytic Focus, LLC

Robert M. Hauser
Vilas Research Professor Emeritus
Department of Sociology, University of Wisconsin-Madison

Executive Director
Division of Behavioral and Social Sciences and Education
National Research Council

Robert A. Kominski
Assistant Chief for Social Characteristics
Social, Economic, and Housing Statistics Division, U.S. Census Bureau

Henry M. Levin
Professor of Economics and Education
Teachers College, Columbia University

Samuel R. Lucas
Professor
Department of Sociology, University of California-Berkeley

Stephen L. Morgan
Jan Rock Zubrow '77 Professor in the Social Sciences
Department of Sociology, Cornell University

Margaret Beale Spencer
Marshall Field IV Professor of Urban Education
Social Sciences Division, Comparative Human Development
University of Chicago

Chris Chapman
Acting Associate Commissioner
Early Childhood, International, and Crosscutting Studies Division
National Center for Education Statistics (Ex officio)

EXECUTIVE SUMMARY

At the request of the National Assessment Governing Board (NAGB, 2003), the National Center for Education Statistics (NCES) convened a panel of experts to provide recommendations concerning socioeconomic status (SES) as a construct, with the understanding that their recommendations might ultimately lead to a new measure of SES that could be used for the National Assessment of Educational Progress (NAEP). The current, most prominent NAEP measure of student SES, National School Lunch Program (NSLP) eligibility, has become less valid over time. The panel's main focus was on the theoretical foundations of SES.

Objectives

Several objectives guided the panel's recommendations:

- provide a definition of SES,
- identify components of SES,
- review data collection and measurement approaches,
- create an SES composite, and
- consider implications of a new measure of SES.

Defining SES

The panel developed the following consensus definition of SES:

> *SES can be defined broadly as one's access to financial, social, cultural, and human capital resources. Traditionally a student's SES has included, as components, parental educational attainment, parental occupational status, and household or family income, with appropriate adjustment for household or family composition. An expanded SES measure could include measures of additional household, neighborhood, and school resources.*

Components and Correlates of SES

The panel concluded that the components of a core student SES measure were the "big 3" variables (family income, parental educational attainment, and parental occupational status), but also suggested that home neighborhood and school SES could be used to construct an expanded measure of SES. Identifying such variables and including them in an expanded SES composite could help improve the explanatory power of SES in accounting for NAEP scores. In addition, some psychological process variables (e.g., coping mechanisms, emotional control, or perceptions of the environment) and some subjective measures (i.e., how one views one's

SES), might be understood as useful contextual and potentially explanatory variables that could help interpret student NAEP scores. Although psychological process and subjective factors were not included as components of a core or expanded SES as developed by the panel, it is important that research be conducted to evaluate the effects of these factors on achievement.

Approaches to Measuring SES Components

In addition to current measures of family income, additional variables, such as housing tenure (rent or own), number of moves in the past year, presence of a household member needing healthcare assistance, and others, could be studied for potential use as indirect measures of family income. Parental educational attainment is currently measured through the NAEP questionnaire, but only for 8th- and 12th-graders, and parental occupational status – one of the big three variables – is not collected in the Student Questionnaire, nor is it available through school records. Cognitive laboratory studies should be conducted on various question types for collecting student reports on parental occupation.

There are currently no direct measures of neighborhood components of a possible expanded SES measure, although NAEP student questionnaire items and information from school records could be aggregated to serve as neighborhood measures. American Community Survey data could be used to provide much of the information not available through NAEP questionnaires and school records. The upcoming Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011) (NCES, 2012a), which tests 4th-graders in 2014, represents an ideal opportunity to inform SES measurement.

SES Composite

There are reporting and interpretation advantages and disadvantages for treating SES as a single measured variable, as several single measured variables, or as a composite of several measured variables. The advantages of a composite variable over the use of single variables outweigh the disadvantages. There are a variety of schemes by which SES components could be combined into a composite measure. A challenge in developing an SES composite is determining whether weights should vary depending on factors such as location or grade level. A review of the existing literature and data quality should be conducted before proposing a recommendation on a component weighting scheme.

Implications

A new SES measure will affect NAEP reporting, including whether and how to characterize SES levels, whether a bridge study must be conducted to link new and old measures of SES, and how a new SES measure will affect NAEP's conditioning model. The research,

framework, and findings associated with the development of a new SES measure could benefit other programs that measure SES, both within and outside NCES. For example, states are continually seeking better measures of SES. In addition, the development of a new SES measure is likely to incur both anticipated and unanticipated side effects, including the requirement to coordinate with other federal programs within and outside NCES, and consequences such as attention given to equity and educational resource distribution.

Key Recommendations

The panel made four key recommendations to improve measurement and reporting of SES:

1. Family income and other indicators of home possessions and resources, parental educational attainment, and parental occupational status should be considered components of a core SES measure, and should be the subject of immediate focus for NAEP reporting.

2. Neighborhood and school SES could be used to construct an expanded SES measure, and measures of these variables could contribute to an expanded SES.

3. Composite measures have many advantages, such as being a single summary useful for reporting, greater reliability, and representing the full range of SES factors. In addition, treating SES as a composite measure does not preclude reporting on relationships between individual SES components and achievement. Therefore, attempts should be made to develop an SES composite measure.

4. The validity of NSLP eligibility has been decreasing due to jurisdiction-wide eligibility and other factors, and that trend is likely to continue. Furthermore, there is concern over the quality of student reports, particularly regarding parental educational attainment (for 4th-graders) and occupational status (for all grades). Due to these data quality issues, along with burden considerations, attempts should be made to explore the possibility of linking to Census data on SES components.

1. SETTING THE STAGE

The National Assessment of Educational Progress (NAEP) is a congressionally authorized project of the National Center for Education Statistics (NCES) within the Institute of Education Sciences of the U.S. Department of Education. The Commissioner of NCES is responsible for carrying out the NAEP project, while the National Assessment Governing Board (Governing Board) oversees and sets policy for NAEP. NAEP measures student progress over time in a variety of subject areas, including reading, writing, mathematics, science, and U.S. history. NAEP does not report individual student scores; rather, the assessment is designed to produce public-domain data about student achievement at the group level. Because NAEP results are meant to inform educators, policymakers, and the general public about the performance of students at the 4th, 8th, and 12th grade levels, reports include overall results as well as scores for student subgroups that are of interest to the target audiences, such as gender, race/ethnicity, and socioeconomic status (SES).

In response to a call by the Governing Board (NAGB, 2003) to improve the measurement and reporting of SES and its relationship to academic achievement in the context of NAEP, NCES convened a panel of experts in the fields of economics, education, statistics, human development, and sociology with substantive expertise in the effects of poverty and disadvantage on student achievement as well as methodological expertise in the measurement of socioeconomic standing. The panel was asked to provide recommendations concerning SES as a construct with the understanding that those recommendations might ultimately lead to a new SES measure that could be used in programs such as NAEP. The guidance was to focus on the theoretical aspects of SES measurement, not on operational aspects. Specifically, the panel was tasked with considering issues surrounding SES, including the creation of a composite measure of SES, how a new SES variable could be used in a reporting context, and how its derivation could be explained to both technical and general audiences. The panel met three times between 2010 and 2012.

This report reflects the discussions and recommendations of the panel and provides the context and background for those discussions. The report was prepared with key NCES stakeholders in mind, including the general public and education policymakers at both the state and national levels. The panel discussed the reporting of SES in NAEP and other large-scale assessments, such as the Program for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS), to learn from those assessments and also to seek to inform them.

2. BACKGROUND

There is a long history of SES being reported to correlate with educational achievement (Cuff, 1934; Holley, 1916; Lynd & Lynd, 1929). The Equality of Educational Opportunity Commission Report (Coleman et al., 1966) played a major role in bringing this finding to prominence in policy circles. Since then, measures of SES have been routinely included in educational research studies as background variables. Researchers and policy makers are interested in SES as a contextual variable to study educational equity and fairness issues, as a covariate with achievement to examine the effects of other variables such as class size or school governance policies, and as a matching variable to ensure the equivalence of treatment and control groups in educational intervention studies. NAEP treats SES as one of five background reporting variables (see law excerpt, below) and scores are reported separately for different SES subgroups. NAEP is mandated to report scores by SES by the No Child Left Behind Act of 2001 (P.L. 107-110, 2002), which acknowledges the importance of SES in educational achievement:

> *The Commissioner, in carrying out the measurement and reporting described in paragraph (1), shall— ''(G) include information on special groups, including, whenever feasible, information collected, cross tabulated, compared, and reported by race, ethnicity,* ***socioeconomic status****, gender, disability and limited English proficiency* (Sec. 411. National Assessment of Educational Progress, Paragraph (b) Purpose; State Assessments; Subparagraph (2) Measurement and Reporting). (115 STAT. 1898)

However, the mandate does not provide specific guidelines on how SES is to be measured, nor even on how it is defined. Current NAEP practice is to measure SES through a set of proxy variables, most notably eligibility for the National School Lunch Program (NSLP; 2008), but also through school Title 1 status, parental educational attainment, and reading materials in the home. For reporting purposes, all of these are treated as individual variables, rather than as a composite SES variable.

It is instructive to review how socioeconomic status is treated in NAEP score reporting. Typically no mention is made of SES per se, but NAEP scores are reported by different variables that might be interpreted as SES measures. In the recent 2009 NAEP Science report (NCES, 2011), for example, for 4th- and 8th-grade students, NAEP scores were reported by eligibility for NSLP in three categories (*not eligible, eligible for a reduced-price lunch,* and *eligible for a free lunch*). NSLP eligibility was reported to be "an indicator of low income" (p. 60). The *Technical Notes* section of the report states that scores were not reported by NSLP eligibility for 12th-grade students "[b]ecause students' eligibility for free or reduced-price school lunch may be underreported at grade 12." (p. 60) (See discussion in the *Measuring SES* section, below.) These data are obtainable from the NAEP Data Explorer, however. For 8th-

grade and 12th-grade students, NAEP scores were reported by parental educational attainment, which is widely regarded as a central component of SES.

A broad and widely accepted definition of SES in the scientific literature emphasizes its role in reflecting access to resources. Furthermore, students' SES is traditionally defined as a combination of family income, parental occupational status, and parental educational attainment. Although the proxy variables currently used in NAEP reflect these factors to some extent, questions have been raised about the quality of the data, the narrowness of the measure, and the lack of a composite SES measure.

Consequently, there have been calls to explore alternative SES measures. Among the suggestions have been to create a composite measure rather than relying on single proxy measures (Barton, 2003), and to use data linked from other sources, such as the U.S. Census Bureau's American Community Survey, which provides data on income, parental occupation, and parental educational attainment (Hauser & Andrew, 2007). The problems identified with current NAEP practice in measuring SES, along with conceptual and empirical developments in understanding SES, suggest that the time is right to consider alternatives in developing a new SES measure for NAEP.

Measuring SES

The history of SES measurement and the identification of possible explanatory correlates show that SES is defined as a broad construct, ideally measured with several diverse indicators. However, there are some advantages to using NSLP eligibility as an operational SES measure for NAEP. First, NSLP eligibility is available through school records for almost every student in the U.S., making data collection inexpensive and minimizing problems with missing data. In addition, NSLP status is a three-level categorical variable, which is convenient for reporting purposes and easily understood by a variety of audiences. Finally, NSLP eligibility status is also tied to federal definitions of poverty, which means that maintenance or updating is handled automatically through updating of federal poverty guidelines.

On the other hand, there are problems with using NSLP eligibility as the main measure of SES in NAEP reporting. These problems can be summarized as follows:

1. *NSLP eligibility measures only one SES component, family income* (adjusted for household composition). NSLP eligibility does not reflect parental educational attainment or occupational status.
2. *Due to the process of eligibility certification, NSLP eligibility may not be the most reliable measure of family income* (Harwell & LeBeau, 2010). Approximately 20 percent of students either are not eligible but are deemed eligible or are eligible but are not recognized

as such (Food and Nutrition Service, 1990; Harwell & LeBeau, 2010; Hauser, 1994). The problem of eligible students failing to apply (whether due to social stigma or some other cause) increases with grade level, and is particularly prevalent for 12th-graders (Office of Research, Nutrition, and Analysis, 1994). Failure to apply when eligible is also thought to correlate with immigration status and to be more prevalent among students who speak English as a second language.

3. *Because there are only three levels of NSLP eligibility, there are large within-category SES differences, particularly in the non-eligible category.* Furthermore, the categories contain uneven shares of the distribution; there is approximately an 8:1 ratio of students in the free vs. reduced-price lunch categories.
4. *School-level and jurisdiction-level eligibility threatens the validity of NSLP eligibility as a measure of an individual student's family income*. All students in a school with greater than 80 percent eligibility are categorized as NSLP eligible, regardless of their family income. Likewise, all students in some jurisdictions, such as Puerto Rico, and many of the urban districts are declared eligible regardless of family income levels.[1]

The remainder of this paper is organized into seven additional chapters. *Objectives* reviews project goals, which are to articulate a definition of SES; identify SES components; address data collection issues, which should lead to a new SES composite; and consider implications of a new SES measure. *Defining SES* emphasizes a broad definition of SES as a student's access to resources and reviews its expected relationship to achievement. *Components and Correlates* presents the idea of SES as a composite of "the big 3" variables—family income, parental educational attainment, and parental occupational status—and additional variables, most notably neighborhood and school SES. Also included here is a discussion of variables that could be considered as either components or correlates of SES, including subjective (perceived) SES, cultural capital, and other factors. The chapter also reviews variables that correlate with SES and variables (moderators) that interact with SES in its relationship with achievement. *Approaches to Measuring SES Components* reviews how the proposed components of SES can be measured. *SES Composite* reviews various ways in which SES components can be weighted and combined, and discusses issues with missing data. *Implications* focuses on the consequences of a new measure of SES for reporting, including anticipated and unanticipated side effects, and discusses possible uses of the new SES measure

[1] Regarding this last point, the most significant problem with the NSLP eligibility measure for the future is the introduction of Community Eligibility (Provision 4) through The Healthy, Hunger-Free Kids Act of 2010 (PL 111-296). Community Eligibility means that schools will no longer be required to keep individual student eligibility information once they have determined a baseline percentage of eligible students, which may result in missing or inaccurate individual student eligibility information. This change in eligibility certification is expected to be phased in, but would nevertheless affect the validity of the NSLP eligibility measure.

by other units, departments, and agencies. Finally, *Discussion* provides a general summary of the paper and concludes with key recommendations.

3. OBJECTIVES

The panel was to provide recommendations for a new SES measure for NAEP that would continue to meet the requirements of reporting SES and also improve the measurement and reporting of SES through the collection of higher quality data.

The primary purpose of proposing a new measure of SES was to meet the requirements of the No Child Left Behind Act of 2001 (P.L. 107-110, 2002) in the best feasible way. The law mandates the reporting of scores by SES in acknowledgement of the importance of SES in educational achievement. The law was not specific about how SES was to be measured. The primary objective was to have a panel provide expert guidance and interpretation on how the law's mandates could be realized.

Specifically, the panel was charged with the following study objectives:

Provide a Definition of SES

Although it is possible to develop and use a measure of socioeconomic status without a clear definition of what it is—for example, basing it on measures that have been used or are currently used in different projects—there are many advantages to communicating a clear definition of SES.

Identify Components of SES

Historically, SES has been defined as some combination of family income, parental educational attainment, and parental occupational status. Other variables have also been considered as part of SES, including various school factors, community or neighborhood factors, and subjective measures of socioeconomic status, such as where individuals see themselves on a status ladder. An objective for this study was to identify which of the various components should be included as part of SES for NAEP reporting.

Review Data Collection and Measurement Approaches

Some SES measures, such as eligibility for the National School Lunch Program, have been obtained from school records. Others, such as parental educational attainment, have been obtained from the student questionnaire. A methodology that has been discussed, and experimented with, involves obtaining geographically aggregated Census data (aggregated over ZIP code tabulation areas, Census blocks, or Census tracts) to impute student family data,

such as family income, household status (e.g., single vs. dual head of household), and parental occupations. Census data can be obtained either from the United States Census 2000 long form (to analyze previously-collected data for research purposes only), or from the ongoing American Community Survey 5-year estimates (which could be analyzed for both research purposes and operational use). An objective of this study was to review data collection and measurement pertinent to these various approaches.

Create an SES Composite (or Justify Use of Multiple Single Variables)

An initial panel objective was to consider alternatives and make recommendations on how an SES composite could be formed. However, the scope was widened to include the possibility of using multiple single variables to measure SES rather than a composite. Thus an objective for the study was to consider the pros and cons of an SES composite vs. multiple single-variable measures of SES. The charge was also to consider various issues in forming a composite, such as how to weight the components of a composite, and whether to vary or keep weights constant across grades, whether to adjust weights (such as income) for locality, how to deal with the issue of missing data, and so forth.

Consider Implications of a New Measure of SES

A new measure of SES would have implications for the reporting of NAEP scores. For example, a new measure of SES might show greater achievement differences between low and high SES groups, compared to free lunch vs. non-subsidized lunch groups. A sudden change in how SES was defined might therefore disrupt trends in the relationship between SES and NAEP achievement scores, which would create significant challenges to interpreting SES estimates over time. Eligibility for a free or reduced-price lunch is a variable with three categories, which is convenient for reporting. A new measure of SES could be a continuous variable, in which case a decision would have to be made about whether to transform it into a categorical variable, or treat it in some other fashion. An objective for this study was to consider these and other implications of a new measure of SES.

4. DEFINING SES

SES is measured by different variables in different studies (e.g., Sirin, 2005), which makes it difficult to appreciate exactly what it is, or what researchers or policy makers mean by SES. However, studies on the relationship between SES and educational achievement cover more than nine decades of research (e.g., Bryant, Glazer, Hansen, & Kursch, 1974; Coleman et al., 1966; Cowan & Sellman, 2008; Harwell & LeBeau, 2010; Holley, 1916; Lynd & Lynd, 1929; White, 1982). It is useful to consider this history in developing a definition of SES.

SES emerged as a concept because of observations that students of parents with low educational attainment, working in low-status jobs, or with low income performed more poorly in school and on tests that reflected school achievement. One of the earliest SES conceptualizations was Taussig's (1920) classification, which was based solely on father's occupational status, classified into seven categories. In a later study by Cuff (1934), the Sims (1927) Score Card was employed as a measure of SES. The Sims Score Card contained a survey with 23 questions about home possessions (books, telephones), rooms in the home, extracurricular and cultural activities, parents' educational attainment, and father's occupation. The Chapin (1933) scale was a rating scale based on the idea that socioeconomic status reflected cultural and material possessions (furniture, accessories), income, and participation in community activities, and which were reflected in and therefore could be measured by home possessions in and the condition of one's living room.

The development of instruments such as the Sims Score Card and the Chapin scale led to increased measurement sophistication of SES. An example was Sewell's (1940) classic study of the measurement of SES in farm families, one of the earliest sociological applications of factor analysis. Ganzeboom, De Graaf, and Treiman (1992) developed a model-based approach in which they proposed that occupational status mediated the relationship between education and income. They then computed occupational status accordingly. The Ganzeboom et al (1992) measure is currently used in The Program for International Student Assessment (PISA) to measure occupational status. Hauser and Warren (1997) similarly took into account educational levels in measuring occupational status.

In the present day, large-scale international assessments routinely include measures of SES. PISA, for example, includes items administered to fifteen-year-old students that form an SES composite called the PISA index of economic, social, and cultural status (ESCS) (OECD, 2010a; see pp. 131). The ESCS is a weighted composite (based on a principal component analysis) of three variables:

- occupational status of the parent with the higher occupational status (based on the Ganzeboom et al. [1992] model, described above),
- educational attainment of the parent with the higher educational attainment, and
- an index of home possessions.

The index of home possessions is itself a composite of three variables (derived from 16 survey questions related to home possessions) and a categorical measure of total number of books in the home:

- wealth (room of their own, Internet link, dishwasher, DVD player, and 3 country-specific measures),

- cultural possessions (classic literature, books of poetry, classic art), and
- home educational resources (desk and quiet place to study, a computer available for school work, educational software, books to help with school work, technical reference books, and a dictionary)
- number of books in the home (four categories: 0-10; 11-100; 101-500; over 500).

Numerous studies over the years have attempted to provide an explanation for why SES correlates with academic outcomes. The Wisconsin Model developed by William H. Sewell and colleagues (Sewell, Haller, & Portes, 1969), based on the Blau and Duncan model (1967), was one of the first attempts to account for educational and occupational attainment by proposing a recursive model including personal aspirations, the influence of peers, educational achievement, parents' SES, and cognitive ability. Along these lines, SES is related to the kind of school and the kind of classroom a student attends (Reynolds & Walberg, 1992), with schools differing characteristically in the kind of instruction offered, materials provided, teacher experience, and access to teachers (Wenglinsky, 1998), as well as the kind of relationship that exists between school staff members and parents (Watkins, 1997).

It may not be family income or poverty per se that drives the relationship between SES and achievement, and life success (Mayer 1997). Spaeth (1976) suggested that SES might indicate the complexity of a child's cognitive environment and that exposure to cognitively challenging home environments prepares students better for the challenges of school. Levin & Belfield (2002) suggested several "pathways" or home environment variables through which SES might affect student achievement. These include the learning environment, language and literacy, parent-child interactions, and daily routine. Low SES children are less likely to have a "school-like" home and follow a daily routine; they have weaker language interaction with parents, weaker literacy engagement, and more conflicting interactions. Walpole (2003) noted that low SES students also tend to have less access to cultural capital (specialized or insider knowledge not taught in schools) and social capital (contacts in networks that can lead to personal or professional gains; Coleman, 1988), which have been argued to be key components of a student's educational success. Recent research in genetics suggests that SES may limit opportunities for children to pursue and benefit from educational experiences congruent with genetically-influenced intellectual interests (Tucker-Drob & Harden, 2012). There also is research linking family socioeconomic resources, including a consideration of family size and structure, to student test scores (Duncan & Magnuson, 2005).

Together, these studies suggest that SES may broadly be seen as a general variable that indexes resources available to the student, including economic, social, and cultural resources. Furthermore, the "big 3" variables discussed earlier can be thought to capture different aspects of resources available to students.

Recently, the American Psychological Association (2007a) created an Office on Socioeconomic Status and issued a report from a specially commissioned American Psychological Association Task Force on Socioeconomic Status. The commission provided a framework for defining and developing SES measures. They characterized three models for understanding SES and social class-related inequalities, across three domains: education, health, and human welfare. One model, reflecting most of the SES literature as reviewed here, was what they referred to as the traditional materialist model. Another model emphasized social gradients and individuals' positions relative to others', which motivates the use of subjective SES measures. A third model focused on social capital, but seemed not to have resulted in specific SES measurement approaches.

Several studies have investigated what kinds of variables have been used in studies of educational achievement to measure SES. White (1982) conducted a meta-analysis on studies before 1980, and Sirin (2005) conducted a meta-analysis of studies published between 1990 and 2000. Both studies indicated a medium to strong relationship between SES and achievement with some measures showing stronger relationships than others. Sirin (2005) found that measures could be placed into the SES categories of parental educational attainment (30 studies), parental occupational status (15 studies), family income (14 studies), free or reduced-price lunch (10 studies), neighborhood (6 studies), and home resources (4 studies). Parental educational attainment was also the most commonly used measure in the studies White (1982) reviewed, and parental occupational status and family income were frequently used as SES measures.

Based on both the history of SES and the measures used to assess SES in studies of educational achievement, the primary measurement of SES over the years has been the "big 3" variables: (a) family income, (b) occupational status of heads of household, and (c) educational attainment of heads of household, consistent with what Hauser (1994) pointed out. However, school and neighborhood variables have also been included in SES measurement for some time (Hauser, 1969). For example, Fertig (2003) examined student peer group achievement heterogeneity on student achievement using PISA data. Van Ewijk and Sleegers (2010) conducted a meta-regression analysis of the effects of peer socioeconomic status on student achievement, and showed effects at both the individual and class levels.

Panel Recommendation: A Definition of SES

A consensus definition of SES is as follows:

> *SES can be defined broadly as one's access to financial, social, cultural, and human capital resources. Traditionally a student's SES has included, as components, parental educational attainment, parental occupational status, and household or family income, with appropriate*

adjustment for household or family composition. An expanded SES measure could include measures of additional household, neighborhood, and school resources.

5. COMPONENTS AND CORRELATES OF SES

In thinking about how SES should be defined for NAEP, it is useful to draw a distinction between components and correlates of SES. An SES component is a variable that should be included as part of SES—that is, as part of the measurement of SES. An SES correlate is simply a variable that correlates with SES, but should not be considered part of SES. It is a high priority in future NAEP data collections to include SES components, while collecting data on SES correlates is a lower priority.

This chapter focuses on determining the components and correlates of SES without regard to the practical aspects of measurement. The next chapter focuses on possibilities for measuring the components of SES.

The "Big 3"

Given the history of SES and how it has been defined over the years, and given the common ways it has been measured in research (see chapter 4), the "big 3"—family income, parental educational attainment, and parental occupational status—should be considered components of SES. Home possessions could be used to measure family income, but there are several drawbacks to using possessions as such a measure: possessions are not typically measured in surveys, do not necessarily represent an accurate picture of family income, and vary over the life cycle in uneven ways with respect to income. Still, possession measures are widely used as SES measures in student educational surveys because they are less intrusive than income measures.

There are additional factors that could be considered components of SES. Specifically, insofar as SES is defined as access to financial, social, and human capital resources, particularly as these factors relate to schooling, they could also be considered components of SES.

Neighborhood SES

The argument for including neighborhood SES information in an expanded measure of student SES is that not all financial, social, and human capital resources available to the individual student come from the family. Some resources come from the neighborhood or community in which the student resides. The resources shape the home environment, broadly conceived, and have been shown to be associated with school achievement.

Traditional indicators of neighborhood SES include the percentages of families below the poverty line, unemployed adults in a neighborhood, and the adults in the neighborhood with a low education level (e.g., percentage without a high school credential). Additional indicators could include the percentage of single parent homes and the percentage of homes where English is not spoken well. In addition, there are social and physical resources associated with neighborhoods, both negative (e.g., the presence of abandoned buildings and roads and walkways in poor condition), and positive (e.g., the availability of parks, recreational areas, and public libraries), that could also be considered part of a neighborhood SES construct. There also are family and household characteristics of a neighborhood, such as aggregated family possessions (e.g., number of rooms in residents' homes, books in their homes, and backyard facilities) which may indicate social and cultural status of a neighborhood.

Empirically, it is not necessarily the case that neighborhood SES data adds information not already available from individual level data. For example, some previous analyses of 8th-graders found that neighborhood data from Census added very little to the relationship between student and parent SES reports (Rivas & Hauser, 2008). However, there is a conceptual distinction between individual family and neighborhood measures of SES, and neighborhood SES should be considered an additional SES component.

There are advantages in including neighborhood SES as part of an individual-level measure of student SES. For one, neighborhood SES can be critical to understanding how student psychological processes (see description in the *Psychological Process Variables* section, below) interact with the context "in real time," and these processes may be influenced by, for example, the creation or expansion of libraries and parks or the diminishment of features such as abandoned buildings and unsafe walkways. Additionally, Census variables that might not be linked at the individual level could be used at the neighborhood level.

Defining what is meant by "neighborhood" (e.g., ZIP code, tract, block group) is difficult, however, and should be considered an operational decision to be decided later. There is also an important distinction to be made between school neighborhood and living neighborhood, as the neighborhood where students live may not have the same characteristics as the neighborhood of the school the student attends, even if they are located in the same ZIP code. For these reasons and others, the prospects for creating components of neighborhood SES that are specific enough to increase the prediction of individual-level NAEP tests scores are uncertain. The odds are sufficiently high that additional work is warranted.

School SES

Many students attend school in the neighborhood in which they live, but some students attend schools outside of their neighborhood due to school choice initiatives and other factors. School

choice is a major movement that may lead to more disconnect between neighborhood SES and the SES composition of the schools that students attend. Therefore, both school and neighborhood SES information could be included as distinct components in an expanded measure of SES. School SES can be defined as the aggregate of the individual students' SES. Currently, school SES is commonly measured by Title 1 status and percentage of students eligible for NSLP.

There are other characteristics of schools (e.g., school safety, physical surroundings) that are relevant for student achievement. However, they should not be considered direct components of an expanded SES measure.

Psychological Process Variables

Research has shown that students at different SES levels have varied levels of exposure to experience with events such as frequent moving or having contact with law enforcement in different ways. Research has also suggested that that low SES is associated with significant risk exposure and low protection factors, and these are likely to influence achievement. Student perceptions of parental involvement and parental monitoring may affect NAEP outcomes. In addition, certain neighborhoods may lead students to adopt coping mechanisms that may not function well in a school environment, or inhibit the development of noncognitive skills such as emotional control. While these are important variables for understanding how students make sense of their environments, psychological process variables, such as coping mechanisms, perceptions, and emotional control, are variables best understood as consequences or correlates of SES rather than as necessary components of SES.

Subjective SES

Research on subjective SES suggests that how one thinks of one's status subjectively can be as important as objective SES measures in relating to outcomes. For example, subjective SES has been shown to predict physical and mental health outcomes after controlling for objective SES (Demakakos, Nazroo, Breeze, & Marmot, 2008). That is, believing you are high status might compensate for lower objective status.

Measurement of subjective SES has relied extensively on the SES ladder technique (e.g., Demakakos, et al., 2008), in which respondents are shown a picture of a ten-rung ladder designed to reflect SES and asked to indicate where they think they (or their family) would stand on the ladder. Other methods for measuring subjective SES include a simple "get along" measure, asking whether the student or student's family has enough money to get along, which has the advantage of being a relative measure that is adjusted over time. Gallup has

administered a “get along” question for several decades in various adult surveys as a means to obtain a subjective estimate of poverty level (e.g., Citro & Michael, 1995).

A number of measurement challenges could hinder development of a valid measure of the subjective SES of students, particularly for 4th-graders. For example, the meaning of subjective SES may vary based on geographic location. A subjective SES measure also might not capture distinctions between high earners with modest educational backgrounds and highly educated middle-level earners. A subjective SES measure could be susceptible to reference group effects (Crede, Bashshur, & Niehorster, 2010), that is, differences in how students see themselves due to the reference group to whom they are comparing themselves. For example, students from homogeneous neighborhoods might interpret objectively small neighbor-to-neighbor differences as large because their reference group is the neighborhood in which they live. (Effects of school heterogeneity on self evaluations has been studied in international surveys [Lafontaine & Monseur, 2007], but it seems that comparable studies have not been conducted with neighborhood heterogeneity.) However, if a valid measure of subjective SES could be developed, it might prove useful as a way to capture whether the child perceives that they have the resources to succeed. This would not, however, be consistent with a measure of SES that indexes access to actual resources of various types.

Panel Recommendations: Identifying Components and Correlates of SES

1. The primary components of SES are the “big 3” variables—family income, parental educational attainment, and parental occupational status.

2. Additional components of an expanded SES measure could include neighborhood and school SES.

3. Psychological variables and some subjective measures of SES may be useful contextual and potentially explanatory variables that could help interpret NAEP scores.

6. APPROACHES TO MEASURING SES COMPONENTS

The purpose of this section is to review ways of measuring the SES components identified in the previous section. The focus is on measuring the “big 3” and neighborhood and school SES. This section, reviews existing measures of each of the SES components,, including school records, the NAEP student background questionnaire, the NAEP 2012 pilot student background questionnaire, and the American Community Survey (ACS). The ACS measures are included because they provide alternative socioeconomic measures, and they may be useful in characterizing geographic areas.

Family income

As reviewed previously, NSLP eligibility, obtained through school records, is a measure of income (adjusted for family composition), and is featured prominently in NAEP reporting. The NAEP student background questionnaire also includes items yielding data that could be understood as reflecting family income:

- Books in the home
- Encyclopedia in the home
- Magazines in the home
- Computer in the home

The 2012 NAEP pilot student background questionnaire includes additional items that may yield data pertaining to family income:

- Home possessions (internet access, clothes dryer, dishwasher, more than one bathroom, your own bedroom)

The ACS includes items pertaining to income:

- Income (9 questions, total) (for each member of the household)
- Home possessions (8 items)
- Rooms in the home (2 items)

Other indirect measures of family income

Several other variables could be considered indirect measures of family income, but are not currently measured in NAEP background questionnaires. These include:

- Housing tenure (rent or own)
- Number of moves in the past year
- Presence of household member needing healthcare assistance
- Immigration status (recency of immigration)
- School resources
- Student's perceived level of support (home, school, neighborhood)

Housing tenure (owning as opposed to renting one's place of residence) is an indicator of income and wealth and of residential stability. In addition, there is considerable evidence regarding its relationship to age-grade retardation and high school dropout (Frederick & Hauser, 2008; Hauser, Frederick, & Andrew, 2007; Hauser, Pager, & Simmons, 2004; Hauser,

Simmons, & Pager, 2004), *Number of moves in the past year* serves as an indirect measure of housing tenure, and also as a measure of instability and high risk status. *Presence of household member needing healthcare assistance* can drain family financial resources. *Immigration status* is an indirect indicator of English language proficiency, social capital, and wealth. *School resources* is not typically thought of as measuring family income, but could be considered indirect measures of family resources, and school resources reflect resources available to the student. *Student's perceived level of support* at home, at school, and in the neighborhood also reflects the availability of resources to the student. Many of these measures could be collected through the student (and teacher and school) questionnaires, and some might be obtainable through school and Census (ACS) records.

The 2012 NAEP pilot student background questionnaire includes an item that may yield data indicating students' English language proficiency, social capital, and wealth:

- How long have you lived in the United States?

Household composition

Household composition—number of parents and siblings—should be included when measuring family income. Partly this is due to the fact that family income has to be distributed across the members of the household, and so financial resources available to the individual student will be a function of both family income and the number of individuals that income is spread across. One or two parents in the household will have an opposite effect, as two parents may provide more social and emotional support than one. NSLP eligibility itself implicitly includes household composition, as its Income Eligibility Guidelines (based on the federal income poverty guidelines) are stated by household size. There are no additional questions on household composition in the NAEP student questionnaire.

The 2012 NAEP pilot student background questionnaire includes the following household structure questions:

- Size of household (total, number of adults)
- Household structure (single- vs. dual-parent, and other relatives)

There may be some ACS variables that could be added to this list, such as number of workers in the household and number of earners in the family.

Parental educational attainment

The NAEP student questionnaire includes two parental educational attainment questions:

- Mother's educational attainment (8th and 12th grade only)
- Father's educational attainment (8th and 12th grade only)

The ACS includes educational attainment questions for each member of the household:

- Whether currently attending school (level and type)
- Educational attainment
- Major (for bachelor's degree holders)

Parental Occupational Status and Employment Status

The NAEP student questionnaire does not include any questions about parental occupation and employment status, nor is such information available from school records. Therefore this SES component has been absent from NAEP reporting.

The 2012 NAEP pilot student background questionnaire includes the following question about parental and household resident employment status:

- How many adults living in your home have a job?

The ACS includes the following employment status and occupation questions for each member of the household:

- Employment status (working for pay or not, part-time vs. full-time, etc.; 22 questions total)
- Occupation (6 questions total)

The National Education Longitudinal Study (NELS:88) (Ingels, 1990) asked 8th-graders for their mother's and father's occupation in an open-ended question. But it also included a closed (multiple-choice) question: "what kind of work do you expect to be doing when you are 30 years old?" The response choices included categories such as craftsperson or operator, farmer or farm manager, professional business or managerial, and so on.

Cognitive laboratory studies must be conducted on various question types for collecting student reports on parental occupation. If questions could be developed to provide reliable information on parental occupation, then it would be useful to use these data in creating a better measure of SES, even if such information does not reach the same reliability and validity level as other questionnaire responses.

There are upcoming opportunities to collect data on new SES component measures. For example, the Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-

K:2011) (NCES, 2012a) wave will test 4th-graders in 2014. That study will collect data from both parents and students. Certain questions concerning parental educational attainment, occupation, home possessions, or any other SES-related questions could be inserted into the parent and the student ECLS-K:2011 questionnaires for the 2014 study. A comparison could be made between the responses to evaluate the validity of 4th-grade student data.

Neighborhood SES

There are currently no direct measures of neighborhood SES from either the NAEP student questionnaire or school records. However, the 2012 NAEP pilot student background questionnaire includes the self-reported ZIP code item, from which neighborhood information could be obtained.

There are a variety of ways to measure neighborhood SES. In addition to measures such as percentage below poverty, unemployed, and with low educational attainment, other variables include the availability of parks and libraries in the neighborhood, the absence of abandoned buildings, proportion of single-parent households, and the proportion of households in which English is not spoken well.

Some neighborhood SES information could be obtained through the extended school questionnaire. Some items from the student questionnaire and from school records could also be aggregated so as to serve as neighborhood SES measures.

Additionally, the ACS includes a number of items that could be treated as measures of neighborhood SES, including unemployment, education, and income levels, household overcrowding, poverty, home ownership, and perhaps some indicators of vulnerability. ACS data would be suitable for neighborhood measures of SES, though there are some challenges in using ACS data, such as heterogeneity in neighborhoods at the tract level and above. Information obtained from the ACS 5-year estimates (with the least sampling error and provided at the smallest geographical unit) would not reflect rapid changes in a neighborhood, but neighborhoods tend to be very stable, and this is not likely to be a problem. Another challenge is that the size and boundaries of a neighborhood for the purposes of creating a neighborhood SES variable are undefined. Perhaps Census blocks (the smallest geographic area for which data are collected and tabulated), or block groups (optimal size of 1,500 people), or possibly even Census tracts (optimal size of 4,000 people), or ZIP code tabulation areas could serve as neighborhood boundaries for this purpose. Using Census blocks or tracts would require obtaining more precise location information for student households than student ZIP codes and would require special arrangements with the Census Bureau.

School SES

As noted above, school SES is most commonly measured by Title 1 status and percentage of students who are eligible for NSLP. However, additional school SES variables could be formed as aggregations of student-level variables obtained through school records or student questionnaires, such as percentage of English language learners, average level of parental educational attainment, average home possessions, and so forth. In addition, characteristics of the schools and school climate could be obtained through teacher and school questionnaires, and these *could* be treated as part of a school SES variable. Because most students attend neighborhood schools, it is likely that school and neighborhood SES measures would correlate highly, but it also may be that there is unique information in the school and neighborhood SES measures.

Panel Recommendation: Review Data Collection and Measurement Approaches

1. Additional variables could be studied as indirect measures associated with family income, including housing tenure (rent or own), number of moves in the past year, presence of a household member needing healthcare assistance, immigration status (and recency of immigration), school resources, and student's perceived level of support (home, school, neighborhood).

2. Family/household composition and structure—size of household and whether single or dual-parent—are also important variables to consider both because single parenthood is generally considered a disadvantage and because household resources are diluted in large households.

3. Parental education is currently measured through the NAEP questionnaire, but only for 8th- and 12th-graders. The ACS includes parental educational attainment questions, which could be used to obtain this measure for 4th-graders. However, a strategy would have to be developed to determine how to link ACS data to NAEP.

4. Cognitive laboratory studies should be conducted on various question types for collecting student reports on parental occupation. If a proper format could be identified for collecting reliable information on parental occupation, then it might be useful to include such questions in future questionnaires even if the reliability and validity level were not as high as is expected for other questionnaire responses.

5. There are currently no direct measures of neighborhood SES from either the NAEP student questionnaire or school records. However, if student ZIP code could be obtained, it may be possible for NAEP data to be linked to ACS data in order to compute neighborhood SES

measures for the students' residential neighborhoods (such as unemployment, education, and income levels, household overcrowding, poverty, home ownership, and perhaps some indicators of vulnerability). Research is needed to determine whether ZIP-code defined "neighborhoods" will yield useful additional components for an expanded SES measure. Data from student questionnaires and information from school records also could be aggregated to serve as neighborhood SES measures, although only for neighborhood schools.

6. School SES can be measured using Title 1 status and percentage of students eligible for NSLP. Additional school SES variables could be computed as aggregations of student-level variables, obtained through school records or through student questionnaires, such as percentage of English language learners, average level of parental educational attainment, and average home possessions. School characteristic and climate variables could be obtained through teacher and school questionnaires and these *could* be part of a school SES variable.

7. An ideal opportunity to inform SES measurement is available through participation in the upcoming Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011) (NCES, 2012a), which tests 4th-graders in 2014. The study will collect data from both parents and students, enabling a comparison of parent and 4th-grade student reports to test the validity of the student reports.

7. SES COMPOSITE

In the literature and in official reports SES is sometimes reported as a single variable, such as parental educational attainment level or NSLP eligibility, and sometimes as a composite variable with several component variables summed together. Initially, the panel was to consider alternatives and make recommendations on how an SES composite could be formed. However, during panel deliberations the scope was widened to include the possibility of the use of a single variable (or multiple single variables) rather than a composite to measure SES. Thus an objective for the panel was to consider the pros and cons of an SES composite vs. single-variable measure of SES. The charge was also to consider various issues in how to form a composite, such as how to weight the components of a composite, and whether to vary or keep weights constant across grades, whether to adjust weights (such as income) for locality, whether to change weights every year, or periodically, and so forth.

An advantage of treating SES as a single variable is that the meaning of a single variable is typically clear and easy to communicate. For example, audiences know what it means to have a parent who completed college, or to be eligible for a free lunch through the NSLP. A related advantage is that the meaning of different SES levels when defined as a single variable can be

transparent—it is clear what the differences are between groups of students who are eligible for a free lunch, eligible for a reduced price lunch, or not eligible. It is also clear what the differences are for students whose parents completed high school versus completed college.

A disadvantage to treating SES as a single variable is that SES is typically understood as having multiple components, including family income, parental occupational status, and parental educational attainment. Treating SES as only one of these components is at odds with the conventional definition of SES. There also is more measurement error in a single variable compared to a composite variable. Some of these problems could be overcome by treating SES as multiple separate variables. However, doing so complicates reporting and interpretation. The separate variables constitute separate operational definitions of SES, which could lead to potentially conflicting data about the relationship between achievement and SES, defined in different ways. For example, achievement for levels of parental educational attainment might show greater variability than achievement for different levels of income.

A composite variable combines information from all the components in a single variable, maintaining simplicity in reporting and avoiding conflicting stories about relationships to achievement. This could mask differences between components in their relationship to achievement, of course, and that could be a potential disadvantage to a composite variable. Nevertheless, the advantages of a composite variable generally outweigh the disadvantages. The remainder of this chapter focuses on a composite measure of SES.

General model

There are several ways to think about what a composite SES variable might look like and how it could be formed. In the psychometrics literature there are two kinds of latent variable models, formative and reflective (Bollen, 2002; Edwards & Bagozzi, 2000). A *reflective* measurement model is one in which the latent variable is assumed to be the cause of the measures or indicators (i.e., which are commonly called reflective or effect indicators, Blalock, 1964). For example, cognitive ability and personality are commonly assumed to cause responses to particular tests or test items. Changes in the latent variable cause changes in the indicator variables. Cronbach's (1959) alpha, factor analysis, and classical test theory are all reflective measurement models—covariation among indicators is assumed to be caused by an underlying latent variable. A *formative* measurement model is one in which the latent variable is assumed to be caused by the indicators (i.e., which are commonly called formative or cause indicators). That is, changes in the indicators cause changes in the latent variable. SES is commonly understood as a latent variable in a formative measurement sense because SES does not cause income, educational attainment, or occupational status; rather, income, educational attainment, and occupational status cause (or determine) SES. A fundamental difference between reflective and formative measurement is that latent variables in reflective

measurement are defined by the degree to which indicators covary, and in fact the pattern of covariances can be used to compute latent factor scores. The situation is different with SES and other latent variables in formative measurement. Here, indicator variables have no necessary relationship with one another and can be uncorrelated or negatively correlated with each other. An example of formative measurement could be variables such as the stress scale (Holmes & Rahe, 1967), which is simply a count of the number of stress-inducing events experienced by an individual within a relatively short period of time (e.g., a year). Indicators are life events, such as death of a spouse, imprisonment, personal injury, or pregnancy, none of which have any necessary relationship to each other (i.e., they can correlate positively, negatively, or be uncorrelated). However, the stress scale predicts future events, such as subsequent illness. In the same way, SES is useful for its ability to predict present and future academic achievement and other life outcomes.

To produce a composite index or score in formative measurement, some scheme must be used to weight the components in some fashion. Below is a review of several ways to do this.

Arbitrary weighting

An infinite number of arbitrary weighting schemes are possible for forming an SES composite. For example, the number of years of parental educational attainment could be added to annual family income and a rating of job status to form an SES composite. However, the weight of these measures would be related to the variance of the components, and somewhat arbitrary re-scalings of components (e.g., changing from income in dollars to income in cents or to a three-level value, such as 1 = high; 2 = medium; 3 = low) could have dramatic effects on the composite. Putting components on the same scale (e.g., through the use of standard scores, or *z* scores) would be a way to avoid this problem. An analyst or policy maker might believe that parental educational attainment is the most important component of SES in an educational application such as NAEP, and so parental educational attainment could be given more weight (e.g., twice the weight) than the other components in forming an SES composite.

An advantage of arbitrary weighting is that it is easy to communicate the rules by which components are combined to form a composite. For example, the "misery index" is the sum of the employment rate plus the inflation rate; the United Nations Development Program's "human development index" is a more complicated geometric mean of normalized indices, but is nevertheless arbitrary. However, the arbitrariness of these indexes is a visible feature, in that their makeup is clear and transparent. A disadvantage of arbitrary weighting is that it is arbitrary. There is no reason to prefer one set of weights to another, and different weights might give different answers to substantive questions (such as, what is the relationship between SES and achievement?).

Empirical weighting

If SES were treated as a latent variable with reflective indicators then component weights could be developed using factor analysis or principal component analysis of the indicators. A rationale for treating SES as a latent variable with reflective indicators is that its components correlate. Treating SES as a latent variable with reflective indicators implies that changing SES would result in a change in income, parental occupational status, and parental education, which seems implausible. Still, the reflective indicators assumption and approach to identifying and weighting an SES composite is used in PISA (see e.g., OECD, 2010a). In PISA, SES (PISA's Economic, Social, and Cultural Status index, or ESCS index) is computed from highest parental educational attainment (in number of years of education), highest parental occupation (converted to a status index), and number of home possessions (summing over 20 items), including books in the home. Weighting these three components is determined by a principal component analysis (conducted separately for each participating country) based on the covariances among the three components, and it has typically yielded approximately equal weights for the three components (although with job status given the most weight, education the second most, and home possessions the smallest). Thus PISA treats SES, at least partly, as a latent variable with reflective indicators.

However, as discussed previously, SES is more commonly thought of as a latent variable with formative indicators, because it is assumed that SES is caused by its indicators rather than the other way around. If SES is treated as a latent variable with formative indicators, then weights cannot be assigned by a covariance-based approach (e.g., principal component analysis) that only considers the components of SES. Such a system weights components according to their centrality (similarity or correlation) with respect to each other, but components do not have to be correlated with each other in a formative measurement model. Instead, under a formative variable assumption, an approach to forming a composite with non-arbitrary weights would be to compute weights through multiple regression analysis using an outcome variable. An outcome variable, such as NAEP Mathematics scores, could be regressed on the SES component variables, and the estimated weights could be used to form an SES composite optimized for predicting NAEP Mathematics scores for that particular grade and year. There are potential drawbacks to producing composite scores this way. One perspective is that SES emerged as a construct because of its predictive relationship with educational outcomes. It is therefore fair and reasonable to weight SES components according to the regression weights of those components when predicting educational outcomes, such as NAEP scores. A complexity related to this perspective is that regression weights will, in general, change depending on which NAEP scores are being predicted (e.g., 4^{th}-grade reading, 8^{th}-grade mathematics, 12^{th}-grade civics, etc.), and in what year they are being predicted (e.g., 2011, 2013). This issue is revisited in the next section of this chapter.

An alternative viewpoint is that SES should have an identity separate from its ability to predict particular achievement outcomes. This viewpoint reflects the perspective that little is learned by studying the relationship between SES and achievement if SES becomes little more than a set of variables optimally weighted to predict achievement. Instead, the relationship between SES and achievement should be a finding rather than an optimization exercise.

There are two ways out of this impasse between a tailored (i.e., regression-weighted) and independent (weights determined without regard to the composite's prediction of achievement) SES composite. If a composite predicts achievement equally well under a range of composite weights (e.g., unit weights, weights determined by regression with 4th-grade mathematics, or with 12th-grade reading), then the distinction between the two composite weighting approaches is of little practical importance. There is some evidence for this perspective (Noel-Miller & Hauser, 2011; Wilks, 1938).

Another way to define an SES composite empirically without tuning component weights to maximize prediction with NAEP scores is to consider additional outcome measures. This was an approach originally suggested by Hauser and Goldberger (1971) as the multiple indicator multiple cause (MIMIC) model (for a recent discussion regarding how this strategy helps identify formative latent variable models, see MacCallum & Browne, 1993). For example, SES is a widely used construct in the health literature, and a health outcome (e.g., absenteeism due to illness) could be used as an additional outcome variable that could be regressed on the SES components. Such a model could be estimated using a structural equation modeling (SEM) approach. Doing so would lead to an SES composite that was not being tuned specifically to the prediction of NAEP achievement. (Multiple NAEP population groups, for example, with various achievement scores at different grade levels, could also be used for this purpose, but the generalization would be to NAEP achievement, not to life outcomes in general.) At the same time, component weights would not be arbitrary, but would be based on the predictiveness of the SES composite across diverse outcomes.

Measurement invariance goals

In developing an SES composite, regardless of whether arbitrary or empirical weights are used and whether SES is a latent variable with formative or reflective indicators, there is an issue concerning the degree to which the composite should be defined in a consistent (i.e., invariant) way (i.e., given the same component weights) across situations; that is, across grades, across NAEP subject areas, across time, across locations, and so forth. One (extreme) option would be to have a specific SES composite for every measurement occasion. That is, there could be a 4th-grade mathematics SES composite for New York in 2013, and a separate SES composite for New Jersey, and separate SES composites for each grade, for each subject, and for each testing year. There could be other categories by which SES composites could be separately formed,

such as urban, suburban, and rural, or by cost-of-living areas (a given family income, say $40,000/year, might indicate different socioeconomic status depending on whether the family resided in Manhattan or the rural south). A National Research Council report on a new poverty measure (Citro & Michael, 1995) recommended adjustment for geographic differences in the cost of housing and insurance.

However, there is a benefit of having an SES composite that maintains the same component weights across all measurement occasions (i.e., across grades, subjects, locations, and years), namely, consistent SES measurement can simplify reporting and interpretation. In current reporting, SES indicators, such as parental educational attainment, NSLP eligibility, and home possessions, are measured the same across all contexts. That is, these indicators are invariant in the raw (manifest) metric across grades, subjects, locations, and years.

The issue of how and the degree to which the SES composite can be kept invariant is affected by the form of the SES composite. For example, with arbitrary weights for the components (sum of unit weighted parental educational attainment, parental occupational status, family income), educational attainment, occupational status, and income could be standardized separately within 4^{th}, 8^{th}, and 12^{th} grade, or they could be standardized across grades (e.g., parental educational attainment could be placed on a common scale across all three grades, or on separate scales, one for each grade). With regression-based weighting for SES components, weights could be obtained from a regression analysis for one grade and one subject, and applied to other grades and subjects, or separate regression analyses for each grade and subject could be conducted, or weights could be averaged across subjects or grades (c.f., Noel-Miller & Hauser, 2011).

Weights could be identical or similar across subjects and years, which might make interpretation and reporting simpler. However, this might not be possible across grades due to the differences in information that can be collected from 4^{th}-graders versus 8^{th}- or 12^{th}-graders. Based on prior research (e.g., Dawes, 1979; Noel-Warren & Hauser, 2011; Wilks, 1938), the weighting scheme might not have much impact on the identity of SES (i.e., applying two sets of component weights to construct an SES composite would likely result in two versions of SES that were highly correlated). Therefore a simpler approach, such as using unit weights or average weights (where average weights involve averaging the component weights obtained in one context with comparable weights obtained in another context), might be advisable. Examining the existing literature (Cohen, 1990) and closely reviewing the quality of data (once it was determined what data would be collected and from what sources) is the most appropriate course of action in determining weighting.

Missing data issues

Dealing with the issue of missing data may be more critical in the case of composite variables compared to single variables such as parental educational attainment (or NSLP), simply because there are more opportunities for data to be missing (e.g., through skips by the respondent). If casewise deletion were invoked any time any of the component items for an SES composite were missing, that could result in both a relatively high number of missing values, and the introduction of bias if data were not missing completely at random (using the standard terminology from Little & Rubin, 1987).

However, there are probably no special problems associated with imputing missing data in the case of computing the SES composite. For example, a standard practice (e.g., used in PISA) is to impute missing values for students with missing data for one of the SES components using data from the other two components. In general, either a maximum likelihood approach for handling missing data in the context of modeling the data, or a multiple imputation approach similar to that used for handling missing achievement data, could be used and would be worth exploring for this purpose (Enders, 2010).

Panel Recommendation: Create an SES Composite

1. The advantages of treating SES as a composite of several variables rather than as a single variable or multiple single variables outweigh the disadvantages.

2. The formative-reflective measurement model distinction was important in considering how to combine SES components into a composite measure. The literature and data quality should be examined before proposing a recommendation on a component weighting scheme.

3. Further study is necessary to address missing data issues in SES measurement.

8. IMPLICATIONS

Adopting a new measure of SES would have various implications on the reporting of NAEP scores. To begin with, a new measure would have to be clearly explained and communicated, because a new measure of SES might show greater achievement differences between low and high SES groups, compared to free lunch versus non-subsidized lunch groups. A sudden change in how SES was defined might therefore disrupt trends in the relationship between SES and NAEP achievement scores, which would create significant challenges to interpreting SES estimates over time.

Reporting and Implications for Trend

As reviewed in chapter 2, achievement scores are disaggregated in NAEP reports by individual SES proxy variables, most notably eligibility for NSLP (*not eligible*, *eligible for a reduced-price lunch*, and *eligible for a free lunch*). Eligibility for a free or reduced-price lunch is a variable with three categories, which is convenient for reporting. A new measure of SES could, and likely would, be a continuous variable. In that case, a decision would have to be made about whether to transform the continuous variable into a categorical variable, or treat it in some other fashion. If it were transformed into a categorical variable, a decision would have to be made about how many categories it could be reported by (e.g., three, more?) and how these categories would be labeled (e.g., low, medium, and high SES).

A new measure of SES would not have to be treated as categorical, however. In PISA (OECD 2010a, Figure II.1.3, p. 32), for example, SES data are reported on a continuous scale, with scatter plots of achievement scores and the PISA index of economic, social, and cultural status (ESCS), and a regression line of achievement on ESCS. With ESCS presented as a continuous variable, PISA reporting makes considerable use of presentations (e.g., tables and scatter plots of ESCS against a variety of variables), and the use of ESCS as a control variable in examining factors such as single-parent families and the like. PISA also computes "socio-economic gradients" that characterize the within-country relationships between ESCS and achievement, facilitating country-to-country comparisons on that measure.

A continuous SES variable could be used in NAEP reporting, but it would not have to be limited to presentations in scatter plots, or as a gradient index. For example, expected SES achievement at, say, the 20^{th}, 50^{th}, and 80^{th} percentiles of SES could be presented, or at the mean SES and at a level one standard deviation above and below the mean of SES. These displays would take a form similar to that taken by NSLP eligibility.

For understanding trends in variables undergoing changes, as SES would be if a new measure were adopted, it is useful to conduct bridge studies, such as those conducted as a result of new race/ethnicity classifications introduced in Census 2000 (Parker, Schenker, Ingram, Weed, Heck, & Madans, 2004). For SES, a carefully constructed study enabling bridging to NSLP eligibility could be useful for understanding trends. For example, for a reporting cycle or two, both SES and NSLP eligibility could be reported as the audience became familiar with the new scale. This would allow readers to compare SES effect sizes (on achievement) with SES measured by NSLP eligibility versus SES measured by a new composite.

Data conditioning

NAEP uses a balanced incomplete block design for administering only subsets of the item pool to particular students (i.e., each student only takes 2 of 11 blocks of items). Background information, including SES, along with data from the items actually administered, is used to estimate scores on the items that are not administered to a particular student, a process referred to as conditioning (Mislevy, 1991). Changing the measure of SES would likely lead to changes in the conditioning model, and changes in the posterior distributions of student responses from which plausible values that secondary analysts use are drawn. A question is how severe a difference in the conditioning model would likely result from a change in the makeup of SES (e.g., from NSLP to a new SES index). There is literature suggesting that the demographic variables are the most important background variables affecting the conditioning model (Thomas, 2002). And SES is likely to be among the more important demographic variables. That same literature, however, suggests that background variables are not as important to the conditioning model as the cognitive variables themselves (i.e., the responses to the cognitive items that are administered). As with the proposed bridge study, current variables, such as NSLP, could be retained, and differences in the conditioning model due to the inclusion of a new SES measure could be studied.

Use by other units, departments, agencies

The focus of the present effort is developing a new SES measure for NAEP. A new SES measure could have direct effects in reporting NAEP scores, such as providing a more valid estimate of the relationship between SES and achievement. In addition, SES is used for the conditioning model in NAEP to assist in the estimation of proficiency scores, and a better measure of SES could be more predictive of proficiency scores and thereby more useful for data conditioning. The quality of NAEP data reported could therefore improve as a result of a better SES measure. In addition to these specific benefits for NAEP reporting, there would be additional benefits based on secondary analysis of NAEP. SES, or proxy measures such as NSLP, is widely used in secondary analysis of NAEP data (e.g., Harwell & LeBeau, 2010; Sirin, 2005).

NCES programs beyond NAEP might benefit from the work conducted in defining and developing a new SES measure. NCES Fast Facts (2012b) provides a list of NCES surveys, many of which use SES measures of some kind. These include adult literacy surveys (National Assessments of Adult Literacy [NAAL], the Program for International Assessment of Adult Competencies [PIAAC]), international comparative surveys (Trends in International Mathematics and Science Study [TIMSS], Progress in International Reading Literacy [PIRLS], the Program for International Student Assessment [PISA]), longitudinal surveys (the Early Childhood Longitudinal Study [ECLS], Baccalaureate and Beyond [B&B], Beginning

Postsecondary Students Longitudinal Study [BPS]), and so forth. For some of these studies there could be a fairly direct transfer of findings on improving SES measurement. For other studies some of the research, methods, and lessons learned in developing an improved SES measure could be incorporated into future study designs.

Investigating new methods for measuring SES could produce benefits that extend beyond NCES and the U.S. Department of Education. For example, in the health sector, there is an extensive literature that relates SES to women's health, public health, and psychological health (APA, 2007b); to specific conditions, such as cancer (Singh, Miller, Hankey, Edwards, 2003) and cardiovascular disease (Winkleby, Jatulis, Frank, Fortmann, 1992); and to other health and wellness issues. Agencies such as the National Institutes of Health (NIH) and the Center for Disease Control (CDC) may benefit from research conducted for NAEP by NCES in developing improved measures of SES.

Anticipated effects and unanticipated side effects

Developing a new SES measure is likely to involve both anticipated effects and unanticipated side effects. It is reasonable to assume that developing the new measure will involve an interagency agreement and collaboration between NCES and the Census Bureau. Such interagency collaborations are beneficial, but often introduce scheduling complications, new costs, and other challenges that require flexibility, patience, and a willingness to consider a variety of approaches to solving potential problems. Another relatively minor change will be a requirement to collect ZIP code information from respondents, perhaps on the NAEP questionnaire. Privacy issues are also likely to be important to resolve.

It is always difficult to anticipate the unanticipated side effects of measurement changes, but as with any assessment, new measurement is often accompanied by the element of consequential validity (Messick, 1995). Consequential validity refers to the changes in practice or culture that accompany changes in assessment. For example, introducing writing assessments can lead to an increased emphasis on writing instruction in the schools; introducing a high-stakes noncognitive skills assessment can lead to more emphasis on developing noncognitive skills. It is not entirely predictable what changes might accompany the introduction of a new SES measure, but if such a measure proves to be more valid than current measures, it is possible that more attention could be given to the importance of the SES-achievement relationship and to a more equitable distribution of educational resources.

Panel Recommendation: Consider Implications of a New Measure of SES

1. There are reporting and psychometric implications that should be considered before implementation of a new SES measure. They include whether and how to characterize SES

levels, whether to conduct a bridge study linking new and old measures of SES, and studying the implications of a new SES measure on the conditioning model used by NAEP to generate plausible values.

9. DISCUSSION

The goal of this panel was to provide recommendations for a new measure of SES that could be used in NAEP. The role of the white paper was to serve as technical documentation of the panel deliberations and to bring this work to the attention of stakeholders and the research community to engage discussion about SES and its measurement.

NAEP is required by law to report scores by SES. Current SES measures, such as NSLP eligibility and parental educational attainment, are single proxy variables, which are limited in several ways. Historically, SES has been defined as a composite measure reflecting resources available to the individual, as expressed in family income, parental educational attainment, parental occupational status, and sometimes neighborhood resources. A common view, as reflected in other large-scale educational assessments such as PISA, is that composite measures that include all of the SES components may be more informative than single measures.

A second limitation of current SES measurement concerns the quality of the data. Student reports of some SES components (such as parental educational attainment) may be unreliable and biased, and reports on variables like these by 4th-graders are likely to be particularly unreliable. This is not to say that they are unusable. Attempts to collect data from 4th-graders on parental educational attainment and perhaps even parental occupational status should be revisited. However, additional data sources such as NCES and state assessment databases and private data sources should also be considered to help bolster the quality of an SES measure.

Perhaps the most critical data quality issue in current SES measurement concerns NSLP eligibility. Measures of NSLP eligibility have several problems, including large errors in eligibility certification and jurisdiction-wide eligibility which fails to differentiate poverty levels within schools or jurisdictions where everyone is declared NSLP eligible (Harwell & LeBeau, 2010; Hauser, 1994). Most importantly, that trend is likely to continue and even get worse.

Given the current limitations of how NAEP measures SES, a major contribution of the panel was to devise a consensus definition of SES, based on a review of various perspectives on SES:

> *SES can be defined broadly as one's access to financial, social, cultural, and human capital resources. Traditionally a student's SES has included, as components, parental educational attainment, parental occupational status, and household or family income, with appropriate*

adjustment for household or family composition. An expanded SES measure could include measures of additional household, neighborhood, and school resources.

Note that this definition outlines and provides a justification for both a core SES measure, which should be the subject of immediate focus for operational reporting, and a more expanded measure, which could be treated as a research project intended to illuminate some of the more contextual and explanatory aspects of SES.

There are other potential components of SES, such as subjective SES and psychological factors. These are best understood as contextual and explanatory variables that could help in the interpretation of SES-achievement relationships, but these contextual factors should not be considered part of a core SES meeting the charge of a congressionally mandated reporting variable. A research program studying these variables, however, could be critical for understanding the importance of measuring SES in the context of an educational achievement survey.

The panel reviewed existing and proposed new measures of SES components from sources including school records, the student questionnaire, additional potential NAEP questionnaire items that were pilot-tested in 2009, 2011, and 2012, and questions from the American Community Survey (ACS). Measures reviewed included ones pertaining to family income and home possessions, parental educational attainment, parental occupational status, and neighborhood wealth and resource indicators. Additional measures that might be related to family income and resources, such as housing tenure, number of residence moves, household members' healthcare needs, immigration status, and household composition measures were also considered. Some of these can be obtained from ACS data. Although NAEP 4th-grade questionnaires do not ask students to indicate parental educational attainment in the questionnaire due to low data quality, such information can be obtained from ACS data. Occupational information is not asked about in the NAEP questionnaires, again due to concerns with low data quality, but such information can be obtained from ACS data. An extensive amount of neighborhood SES data could be obtained from the ACS, including neighborhood poverty levels, unemployment, educational attainment, presence of parks and libraries, abandoned buildings, single-parent households, and non-English speaking households. However, there are challenges in obtaining these kinds of data from ACS and for linking ACS data to NAEP data, such as determining how best to aggregate data in linking datasets.

There are a wide variety of ways to combine all the information on components of SES. A composite can be assembled by summing variables reflecting family income, parental educational attainment, parental occupational status, and neighborhood SES indicators. The primary distinction is in whether the summing would occur by arbitrarily weighting the

components (e.g., unit weighted), or by allowing the components to be weighted to best predict some outcome, such as student achievement. There are advantages and disadvantages to both approaches. Another important consideration would be how to maintain the meaning of SES across grades, across locations (e.g., varying cost-of-living regions), and across time. However, with respect to the issue of component weighting, there is some evidence that this may be merely academic and that practically how variables are weighted might not make much difference in what SES is (Noel-Miller & Hauser, 2011). That is, an SES composite with a set of weights determined from one context is likely to be highly correlated with an SES composite based on a set of weights determined from a different context, given that the components themselves tend to be highly correlated, and neither will be much different from a unit-weighted composite, as has been long known (Wilks, 1938).

Developing a new SES measure for NAEP has implications for reporting and elsewhere. If a new measure were developed, it might be useful to report achievement results disaggregated by SES, measured both by the current measures (e.g., NSLP eligibility, parental educational attainment) and by the new composite measure. While it may be valuable to treat SES as a continuous variable, it could also be treated as a categorical variable (e.g., low, medium, and high SES). The research and findings resulting from developing the new SES measure for NAEP would benefit other federal programs both within and outside NCES.

Key Recommendations

Summarized below are the panel's key recommendations for improving the measurement and reporting of SES.

Recommendation 1. Family income and other indicators of home possessions and resources, parental educational attainment, and parental occupational status (the "big 3"), should all be considered components of a core SES measure; that is, part of the measurement of a core SES variable. The core SES measure should be the subject of immediate focus for operational reporting. This recommendation reflects the academic literature on SES.

Recommendation 2. An expanded SES measure could include additional variable components besides family income, parental educational attainment, and parental occupational status. These additional components could include resources available in the student's neighborhood or community and resources available at school. Consideration should be given to the development of an expanded SES measure in addition to the core SES measure.

Recommendation 3. The advantages of treating SES as a composite—e.g., a single summary for reporting, greater reliability, and representation of the full range of SES factors—outweigh the disadvantages, especially because the use of the composite would not preclude using and reporting on single measures. Therefore, attempts should be made to develop an SES composite measure.

Recommendation 4. The validity of the most widely used measure of SES—NSLP eligibility—has been decreasing due to jurisdiction-wide eligibility and other factors, and that trend is likely to continue. There will be growing pressure to replace NSLP eligibility with a new, more valid measure. Burden issues prohibit a longer questionnaire, and there is concern about the reliability of student reports on SES components, particularly educational attainment (for 4th-graders) and occupation (for all grades). Because of data quality issues, along with burden considerations, attempts should be made to explore the possibility of linking to Census data on SES components. Studies should be conducted with the U.S. Census Bureau to determine the feasibility of linking Census data to NAEP and to evaluate the quality of the data that would result from various linking strategies.

REFERENCES

American Psychological Association. (2007a). *Office on socioeconomic status*. Retrieved from http://www.apa.org/pi/ses/homepage.html.

American Psychological Association. (2007b). *Report of the APA Task Force on socioeconomic status*. Washington, DC: Author. Retrieved from http://www2.apa.org/pi/SES_task_force report.pdf.

Barton, P. E. (2003). *Parsing the achievement gap: Baselines for tracking progress.* Princeton, NJ: Policy Information Center, Educational Testing Service.

Blau, P. M., & Duncan, O. D. (1967). *The American Occupational Structure*. New York: Wiley and Sons.

Blalock, H.M. (1964) *Causal Inferences in Nonexperimental Research*, Chapel Hill, NC: University of North Carolina Press.

Bollen, K.A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology, 53(1),* 605-634.

Bryant, E. C., Glazer, E., Hansen, M. A., & Kursch, A. (1974). *Associations between educational outcomes and background variables [Monograph]*. Denver, CO: National Assessment of Educational Progress.

Burton, L., Price-Spratlen, T., & Spencer, M. B. (1997). On ways of thinking about measuring neighborhoods: Implications for studying context and developmental outcomes for children. In J. Brooks-Gunn, G. Duncan, & J. Lawrence Aber (Eds.), *Neighborhood poverty: Context and consequences for children* (Vol. 2, pp. 132–144). New York: Russell Sage Foundation Press.

Chapin, F. S. (1933). *The measurement of social status, 3,* University of Minnesota Press.

Citro, C.F. & Michael, R.T. (1995). *Measuring poverty: A new approach. Panel on Poverty and Family Assistance: Concepts, Information Needs, and Measurement Methods, Committee on National Statistics, Commission on Behavioral and Social Sciences and Education, National Research Council.* Washington, DC: National Academy Press.

Cohen, Jacob. (1990). Things I have learned (so far). *American Psychologist, 45*, 1304-1312. doi:10.1037/0003-066X.45.12.1304

Coleman, J. S. (1988). Social capital in the creation of human capital. American Journal of Sociology, 94, S95–S120.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfield, F. D., et al. (1966). *Equality of educational opportunity (2 vols.)*. Washington, DC: U.S. Government Printing Office.

Cowan C.D., Sellman S.W. (2008). *Improving the quality of NAEP socioeconomic status information: Report on research activities.* Alexandria, VA: Human Resources Research Organization (HumRRO). (Retrieved from http://www.humrro.org/corpsite/publication/improving-quality-naep-socioeconomic-status-information-report-research-activities, February, 2012.)

Credé, M., Bashshur, M. & Niehorster, S. (2010) Reference group effects in the measurement of personality and attitudes. *Journal of Personality Assessment, 92,* 390 – 399.

Cuff, N. B. (1934). The vectors of socio-economic status. Pe*abody Journal of Education, 12 (3),* 114-117.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist 34,* 571-582. doi:10.1037/0003-066X.34.7.571.

Demakakos, P., Nazroo, J., Breeze, E., & Marmot, M. (2008). Socioeconomic status and health: The role of subjective socioeconomic status. *Social Science and Medicine, 67,* 330-340.

Duncan, G.J., & Magnuson, K. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps? *The Future of Children, 15,* 35-54.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of the relationship between constructs and measures. *Psychological Methods, 5,* 155-174.

Enders, Craig K. (2010). *Applied Missing Data Analysis*. New York: Guildford Press.

Fertig, M. (2003). Educational Production, Endogenous Peer Group Formation and Class Composition – Evidence from the PISA 2000 Study, IZA DP No. 714, Bonn, Germany: Institute for the Study of Labor.

Food and Nutrition Service. (1990). Study of Income Verification in the National School Lunch Program, Final report (Vol. 1, Rep. No. FNS- 53–3198-6–44, prepared by R. St. Pierre, M. Puma, M. Battaglia, & J. Layzer). Alexandria, VA: U.S. Department of Agriculture. Retrieved from http://www.fns.usda.gov/oane/MENU/Published/CNP/cnp-archive.htm.

Frederick, C., & Hauser, R. M. (2008). Have We Put an End to Social Promotion? Changes in School Progress among Children Aged 6 to 17 from 1972 to 2005, *Demography*, 719-740.

Ganzeboom, H. B., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research 21(1),* 1-56.

Harwell, M. R., & LeBeau, B. (2010). Student eligibility for a free lunch as an SES Measure in education research. *Educational Researcher, 39,* 120-131.

Hauser, Robert M. (1969) Schools and the Stratification Process. *American Journal of Sociology, 74 (6),* 587-611.

Hauser, R. M. (1994). Measuring socioeconomic status in studies of child development. *Child Development, 65*, 1541–1545.

Hauser, R. M., & Andrew, M. (2007). *Reliability of student and parent reports of socioeconomic status in NELS-88*. Working paper presented at ITP seminar at the University of Wisconsin–Madison. Retrieved from http://www.wcer.wisc.edu/itp/Spring%20 08%20seminar/HauserNELS-SES%20measurement_070607a.pdf.

Hauser, R. M., Frederick, C., & M. Andrew (2007). Grade Retention in the Age of Standards-Based Reform. In A. Gamoran (Ed.) *No Child Left Behind and Poverty* (Pp. 120-53). Washington, DC: Brookings Institution.

Hauser, R. M., & Goldberger, A.S. (1971). The treatment of unobservable variables in path analysis. In H. L. Costner (Ed.), *Sociological Methodology (pp. 81-117).* San Francisco: Jossey-Bass.

Hauser, R. M., Pager, D. I., & Simmons, S. J. (2004). Race-Ethnicity, Social Background, and Grade Retention. In H. Walberg, A. J. Reynolds, & M. C. Wang (Eds), *Can Unlike Students Learn Together? Grade Retention, Tracking, and Grouping.* (pp. 97-114).Greenwich, CT: Information Age Publishing.

Hauser, R. M., Simmons, S. J., & Pager, D. I. (2004). High School Dropout, Race-Ethnicity, and Social Background from the 1970s to the 1990s. In G. Orfield (Ed.), *Dropouts in America: Confronting the Graduation Rate Crisis* (Pp. 85-106). Cambridge, MA: Harvard Education Press, 2004.

Hauser, R. M., & Warren, J. R. (1997). Sociological indexes for occupations: A review, update, and critique. *Sociological Methodology, 27*, 177–298.

Holley, C. E. (1916). *The relationship between persistence in school and home conditions*. Chicago, IL: University of Chicago Press.

Holmes, T.H., & Rahe, R.H. (1967). The Social Readjustment Rating Scale. *Journal of Psychosomatic Research, 11(2),* 213–218.

Ingels, S. J. (1990). *National education longitudinal study of 1988: base year: student component data file user's manual.* Washington, D.C.: U.S. Dept. of Education, Office of Educational Research and Improvement. http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2002323.

Kurki, A., Boyle, A., & Aladjem, D.K. (2005*). Beyond free lunch: Alternative poverty measures in educational research and program evaluation.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Lafontaine, D., Monseur, C. (August, 2007). Why do non-cognitive variables better predict mathematics achievement in some countries than in others? A methodological study of PISA 2003. Earli Conference, Budapest

Levin, H. M. & Belfield, C. R. (2002). Families as contractual partners in education. *UCLA Law Review, 49,* 1799-1824.

Little, R.J.A., & Rubin, D.B. (1987) *Statistical analysis with missing data.* New York: Wiley.

Lynd, R. S., & Lynd, H. M. (1929). *Middletown: A study in American culture*. New York: Harcourt Brace.

MacArthur (2011). *Research Network on SES & Health: Research, Psychosocial Notebook.* http://www.macses.ucsf.edu/research/psychosocial/usladder.php (retrieved 4 August 2011).

MacCallum, R.C., & Browne, M.W. (1993). The use of causal indicators in covariance structure models: Some practical issues. *Psychological Bulletin, 114(3),* 533-541.

Mayer, Susan E. 1997. *What money can't buy: Family income and children's life chances.* Cambridge, MA: Harvard University Press

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50,* 741–749.

Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika, 56,* 177-196.

National Assessment Governing Board. (2003). *Background information framework for the National Assessment of Educational Progress.* Washington, DC: U.S. Department of Education, National Assessment Governing Board.

National School Lunch Program. (2008). *Fact sheet*. Retrieved from http://www.fns.usda.gov/cnd/governance/notices/naps/NAPs.htm.

NCES (2011). The Nation's Report Card: Science 2009 (NCES 2011–451). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

NCES (2012a). *Early Childhood Longitudinal Program*. Retrieved from http://nces.ed.gov/ecls/Kindergarten.asp.

NCES (2012b). Fast Facts. http://nces.ed.gov/fastfacts/display.asp?id=11

No Child Left Behind Act of 2001. Pub. L. No. 107-110, 115 Stat. 1425 (2002).

Noel-Miller, C. & Hauser, R. M. (2011). Unpublished manuscript.

Office of Research, Nutrition, and Analysis. (1994). *School lunch eligible non-participants: Final report* (Rep. No. 53–3198–018, prepared by F. Glantz, R. Berg, D. Porcari, E. Sackoff, & S. Pazer). Alexandria, VA: U.S. Department of Agriculture. Retrieved from http:// www.fns.usda.gov/oane/MENU/Published/CNP/cnp-archive.htm

OECD (2010a). *PISA 2009 Results: Overcoming Social Background – Equity in Learning Opportunities and Outcomes (Volume II).* Retrieved from http://dx.doi.org/10.1787/9789264091504-en, December 6, 2011.

OECD (2010b). *Education at a glance 2010: OECD indicators*. Retrieved from http://www.oecd.org/document/52/0,3746,en_2649_39263238_45897844_1_1_1_1,00.html.

Parker, J.D., Schenker, N., Ingram, D.D., Weed, J.A., Heck, K.E., Madans, J.H. (2004). Bridging between two standards for collecting information on race and ethnicity: An application to Census 2000 and vital rates. *Public Health Reports (119),* 192-205.

Reynolds, A. J., & Walberg, H. J. (1992). A structural model of science achievement and attitude: An extension to high school. *Journal of Educational Psychology, 84,* 371-382.

Rivas, S, & Hauser, R. M., (2008). *Progress Report: Toward Improved Socioeconomic Measurement in NAEP.* University of Wisconsin. Unpublished manuscript.

Sewell, W.H. (1940). The construction and standardization of a scale for the measurement of socio-economic status of Oklahoma farm families. *Oklahoma Agricultural and Mechanical College Experimental Station, Technical Bulletin No. 9.*

Sewell, W. H., Haller, A., & Portes, A. (1969). The educational and early occupational attainment process. *American Sociological Review,* 34, 82-92.

Sims V. M. (1927).*The measurement of socioeconomic status.* Bloomington, IL: Public School Printing Co.

Singh, G.K., Miller, B.A., Hankey, B.F., Edwards, B.K. (2003). *Area Socioeconomic Variations in U.S. Cancer Incidence, Mortality, Stage, Treatment, and Survival, 1975–1999.* NCI Cancer Surveillance Monograph Series (4), NIH Publication No. 03-5417. Bethesda, MD: National Cancer Institute.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*, 417–453.

Spaeth, J. L. (1976). Cognitive complexity: A dimension underlying the socioeconomic achievement process. In W. H. Sewall, R. M. Hauser, & D. L. Featherman (Eds.), *Schooling and achievement in American society* (pp. 103-131). New York: Academic Press.

Taussig, F. W. (1920). *Principles of economics.* Newcastle: Cambridge Scholars Publishing.

Thomas, N. (2002). The role of secondary covariates when estimating latent trait population distributions. *Psychometrika, 67(1),* 33-48.

Trends in International Mathematics and Science Study (2003). *Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*. Retrieved from http://pirls.bc.edu/timss2003i/PDF/t03_af_book.pdf

Tucker-Drob, E. M., & Harden, K. P. (2012). Intellectual interest mediates gene-by-socioeconomic status interaction on adolescent academic achievement. *Child Development, 83,* 743-757.

van Ewijk, R. & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review, 5 (2),* 134-150.

Watkins, T.J. (1997). Teacher communications, child achievement, and parent traits in parent involvement models. *Journal of Educational Research, 91(1),* 3-14.

Walpole, M. (2003). Socioeconomic status and college: How SES affects college experiences and outcomes. *Review of Higher Education, 27,* 45–73.

Wenglinsky, H. (1998). Finance equalization and within-school equity: The relationship between education spending and the social distribution of achievement. *Education Evaluation and Policy Analysis, 20(4),* 269-283.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91,* 461–481.

Wilks, S. S. (1938). "Weighting systems for linear functions of correlated variables when there is no dependent variable". *Psychometrika, 3,* 23–40.

Winkleby, M.A., Jatulis, D.E., Frank, E., Fortmann, S.P. (1992). Socioeconomic status and health: How education, income, and occupation contribute to risk factors for cardiovascular disease. *American Journal of Public Health, 82(6),* 816-820.

EDUCATION WEEK
Published Online: December 11, 2012
Published in Print: December 12, 2012, as **NAEP Seeks to Test New Measure of Student Poverty**

New Student-Poverty Measures Proposed for National Tests

Proposed indicators go broader, deeper

By Sarah D. Sparks

Washington

Back to Story

Aiming to get a clearer picture of how students' home and community resources affect their academic achievement, America's best-known K-12 education barometer, the National Assessment of Educational Progress, is building a comprehensive new way to **gauge socioeconomic status**.

The new measure, being developed by the National Assessment Governing Board and the National Center for Education Statistics, is intended to look beyond a traditional measure of family income to a child's family, community, and school supports for learning.

"This issue has just been on the burner for so, so long," said Maria V. Ferguson, the executive director of the Washington-based Center on Education Policy. "When NAGB starts talking about it, that does elevate it to a place where it could be part of a bigger policy debate," she said. "I wonder if the folks at NAGB are hoping this could be an opening salvo into a bigger conversation about how [different SES measures] might affect other programs."

The governing board commissioned eight researchers in education, economics, statistics, human development, and sociology that have been working on the new indicators since 2010. The panel released its initial proposal at a NAGB meeting here Nov. 29.

"We rapidly learned that socioeconomic status contains multiple dimensions and categories that don't neatly collapse back to 'low' versus 'high,' " said Charles D. Cowan, the chief executive officer of the San Antonio-based research group Analytic Focus and a member of the governing board's expert panel. "Over the last 10 to 15 years, there's been an explosion in the data available" on student characteristics, Mr. Cowan said. "Perhaps now is the time to think about alternative measures of SES simply because now we are able to think about it."

Beyond Free Lunch

For decades, the universal proxy for students' socioeconomic status—for NAEP and nearly every federal education and child-health program—has been just such a high-low indicator: **eligibility for subsidized meals** under the National School Lunch Program.

Federal food aid does capture a huge swath of students in poverty: The school lunch program alone provides meals for more than 31 million children, at reduced cost to those living at or below 185 percent of the **federal poverty line**, and free to those who are at or below 130 percent of the poverty line

Socioeconomic Status Reconsidered

The National Assessment Governing Board is considering a new method of identifying a student's socioeconomic status when disaggregating the results of the National Assessment of Educational Progress. NAEP researchers now rely primarily on a student's eligibility for the National School Lunch Program—which as of 2011 provided free or low-cost meals to more than 31 million students in poverty each day—as a proxy for socioeconomic status. This

or who are homeless, in foster care, or in certain other programs. In 2012 in the lower 48 states and the District of Columbia, children living in a family of four on $40,000 or less a year would be eligible for reduced price meals; the free-lunch cut-off for the same family would be $30,000.

From a research and policy perspective, however, experts say food-aid eligibility gives an incomplete picture of the resources of students in poverty, and no information about students who don't qualify. Moreover, those poverty counts notoriously underrepresent students as they get older and more self-conscious about applying for free or reduced-price lunch.

"There are many problems regarding the use of free and reduced-cost lunch," said Henry M. Levin, a research panelist and an economics and education professor at Teachers College, Columbia University, who is now on sabbatical at Peking University in Beijing.

"It does not distinguish in a sensitive way differences along the entire spectrum of SES," he noted in an email to Education Week. "Even for the poor or relatively poor, there are large differences" within the range of free-lunch eligibility.

The governing board has tried in the past to fill in the gaps using the background questionnaire students complete along with NAEP, according to William Ward, a senior research scientist for assessment at NCES, which administers NAEP. But some of those questions have become outdated or have not been found to be relevant to a child's real socioeconomic status.

traditional indicator is bolstered by background questions on home possessions, such as washing machines, encyclopedias, and mobile phones.

Proposed New "Core" SES Indicators

- Family income and indicators of home possessions and that have been shown to be linked to educational access, such as Internet availability and number of books in the home
- Parents' educational attainment
- Parents' occupational status

Potential "Expanded" SES Indicators

- **Family:** For example, family structure, stability, and the presence of extended family and other supportive adults
- **Neighborhood:** Including the concentration of poverty or linguistic isolation, the percentage of unemployed adults, and the availability of museums, parks, or safe walking routes
- **School:** The aggregate SES composition of students at the school the child attends, as distinct from the neighborhood SES level

Potential Additional Context Indicators

- **Physical stressors:** Local rates of illness or environmental problems
- **Psychological stressors:** Levels of crime in the school and community
- **Psychological protectors:** Student perception of parent involvement and expectations

SOURCE: National Assessment Governing Board

"We used to ask, 'Do you have a washer-dryer?' but now everyone has a washer-dryer," Mr. Cowan said. "We used to ask, 'Do you have a cellphone?' Now, do any of your students *not* have a cellphone?"

More Than Income

The updated measure of socioeconomic status will look at broader resources and learning supports, Mr. Cowan said.

It will start with the "big three": the family's income, parents' level of educational attainment, and whether and where they are employed. This year's administration of NAEP has also tried

out new background questions, including how long the child has lived in the United States, how many family members live with the child, and how many adults in the home have a job.

Because elementary students in particular may have difficulty identifying these, the governing board is considering supplementing the data with information from the U.S. Census Bureau's **American Community Survey**, an annual study of a representative sample of 3.5 million households nationwide that asks about family structure, employment and income, transportation, and other details.

The NAEP student survey would still include questions about home possessions that research has shown to be related to student achievement, such as access to the Internet and the number of books in the home, Mr. Ward said. But the board is considering supplementing the "core" SES measures with other indicators of resources in the child's neighborhood and school that could highlight differences between students living at the same income level in different areas.

For example, an 8th grader in New York's Spanish Harlem neighborhood could still have access to libraries and museums, while a peer in rural southern Utah may have no local library but live a bike ride away from national parks.

Indicators of school and neighborhood supports also may be pulled from administrative data and from the Census Bureau, such as the degree of concentration of poverty or linguistic isolation, the average educational degree earned, and the employment levels in the neighborhood.

The governing board panel plans to present its proposed socioeconomic indicators at the annual meeting of the American Educational Research Association in April before piloting their use in 2014 and reporting the results in 2015.

Vol. 32, Issue 14, Pages 6-7

NCES
National Center for Education Statistics

Background Variables on International Assessments: PISA, TIMSS, and PIRLS

The United States regularly participates in three major international assessments:

- **PISA**—Program for International Student Assessment , an assessment of 15-year-old students in mathematics, reading, and science literacy (as well as occasional other subjects)
- **TIMSS**—Trends in International Mathematics and Science Study, an assessment of mathematics and science at grades 4 and 8
- **PIRLS**—Progress in International Reading Literacy Study, an assessment of reading at grade 4

All three assessments include contextual questionnaires that address topics similar to those found in NAEP questionnaires. The topics include students' family backgrounds and home learning resources, their attitudes and experiences concerning the subjects assessed, as well as school questionnaires that address demographic, resource, and governance aspects of students' schools. PIRLS and TIMSS also include teacher questionnaires addressing teachers' demographic backgrounds, preparation for teaching, perceptions of their schools and jobs, and some aspects of their instruction and their attitudes and beliefs about teaching and learning.

At the meeting of the Ad Hoc Committee, NCES will discuss how the international questionnaires compare with those used by NAEP and how background data are used in reporting the international assessments.

Illustrative materials include examples of reports and background questionnaires from PISA and TIMMS and a NAEP student questionnaire for comparison:

- PISA 2009 Results: Executive Summary
- PISA in Focus: Equity
- TIMSS 2011 report chapter on school climate and discipline
- PISA student questionnaire
- PISA school questionnaire
- General description of TIMSS contextual questionnaires
- TIMSS student questionnaire
- NAEP student questionnaire for grade 8 mathematics

PISA 2009 Results: Executive Summary

This work is published on the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Organisation or of the governments of its member countries.

Please cite this publication as:
OECD (2010), *PISA 2009 Results: Executive Summary*

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Photo credits:
Getty Images © Ariel Skelley
Getty Images © Geostock
Getty Images © Jack Hollingsworth
Stocklib Image Bank © Yuri Arcurs

THE OECD PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT (PISA)

PISA focuses on young people's ability to use their knowledge and skills to meet real-life challenges. This orientation reflects a change in the goals and objectives of curricula themselves, which are increasingly concerned with what students can do with what they learn at school and not merely with whether they have mastered specific curricular content. PISA's unique features include its:

- Policy orientation, which highlights differences in performance patterns and identifies features common to high-performing students, schools and education systems by linking data on learning outcomes with data on student characteristics and other key factors that shape learning in and outside of school.
- Innovative concept of "literacy", which refers both to students' capacity to apply knowledge and skills in key subject areas and to their ability to analyse, reason and communicate effectively as they pose, interpret and solve problems in a variety of situations.
- Relevance to lifelong learning, which goes beyond assessing students' competencies in school subjects by asking them to report on their motivation to learn, their beliefs about themselves and their learning strategies.
- Regularity, which enables countries to monitor their progress in meeting key learning objectives.
- Breadth of geographical coverage and collaborative nature, which, in PISA 2009, encompasses the 34 OECD member countries and 41 partner countries and economies.

To learn more about PISA and to download our publications and data, please visit our website: *www.pisa.oecd.org*

To learn more about the OECD, please visit *www.oecd.org*

Executive Summary

WHAT STUDENTS KNOW AND CAN DO: STUDENT PERFORMANCE IN READING, MATHEMATICS AND SCIENCE

PISA's conception of reading literacy encompasses the range of situations in which people read, the different ways written texts are presented, and the variety of ways that readers approach and use texts, from the functional and finite, such as finding a particular piece of practical information, to the deep and far-reaching, such as understanding other ways of doing, thinking and being. Research shows that these kinds of reading literacy skills are more reliable predictors of economic and social well-being than the number of years spent in school or in post-formal education.

Korea and Finland are the highest performing OECD countries, with mean scores of 539 and 536 points, respectively. However, the partner economy Shanghai-China outperforms them by a significant margin, with a mean score of 556.

Top-performing countries or economies in reading literacy include Hong Kong-China (with a mean score of 533), Singapore (526), Canada (524), New Zealand (521), Japan (520) and Australia (515). The Netherlands (508), Belgium (506), Norway (503), Estonia (501), Switzerland (501), Poland (500), Iceland (500) and Liechtenstein (499) also perform above the OECD mean score of 494, while the United States, Sweden, Germany, Ireland, France, Denmark, the United Kingdom, Hungary, Portugal, and partner economy Chinese Taipei have scores close to the OECD mean.

The lowest performing OECD country, Mexico, has an average score of 425. This means that the gap between the highest and lowest performing OECD countries is 114 points – the equivalent of more than two school years. And the gap between the highest and lowest performing partner country or economy is even larger, with 242 score points – or more than six years of formal schooling – separating the mean performance of Shanghai-China and Kyrgyzstan (314).

Differences ***between*** countries represent, however, only a fraction of overall variation in student performance. Addressing the educational needs of such diverse populations and narrowing the gaps in student performance that have been observed remains a formidable challenge for all countries.

In 18 participating countries, including Mexico, Chile and Turkey, the highest reading proficiency level achieved by most students was the baseline Level 2.

Level 2 is considered a baseline level of proficiency, at which students begin to demonstrate the reading skills that will enable them to participate effectively and productively in life. Students who do not reach Level 2 have difficulties locating basic information that meets several conditions, making comparisons or contrasts around a single feature, working out what a well-defined part of a text means when the information is not prominent, or making connections between the text and outside knowledge by drawing on personal experience and attitudes. The proportion of 15-year-olds in this situation varies widely across countries, from fewer than one in 10 in four countries and economies to the majority of students in 10 countries. Even in the average OECD country, where nearly one student in five does not reach Level 2, tackling such low performance remains a major challenge.

At the other end of the proficiency spectrum, an average of 7.6% of students attain Level 5, and in Singapore, New Zealand and Shanghai-China the percentage is more than twice the OECD average.

However, for some countries, developing even a small corps of high-performing students remains an aspiration: in 16 countries, fewer than 1% of students reach Level 5. Students at this level are able to retrieve information requiring the reader to locate and organise several pieces of deeply embedded information, inferring which information in the text is relevant. They can critically evaluate information and build hypotheses drawing on specialised knowledge, develop a full and detailed understanding of a text whose content or form is unfamiliar, and understand concepts that are contrary to expectations.

Results from the PISA 2009 assessment show that nurturing high performance and tackling low performance need not be mutually exclusive. The countries with the very highest overall reading performance in PISA 2009, Finland and Korea, as well as the partner economies Hong Kong-China and Shanghai-China, also have among the lowest variation in student scores. Equally importantly, Korea has been able to raise its already-high reading performance even further, by more than doubling the percentage of students reaching Level 5 or higher since 2000.

Korea, with a country mean of 546 score points, performed highest among OECD countries in the PISA 2009 mathematics assessment. The partner countries and economies Shanghai-China, Singapore and Hong Kong-China rank first, second and third, respectively.

In the PISA 2009 mathematics assessment, the OECD countries Finland, Switzerland, Japan, Canada, the Netherlands, New Zealand, Belgium, Australia, Germany, Estonia, Iceland, Denmark, Slovenia and the partner countries and economies Chinese Taipei, Liechtenstein and Macao-China also perform significantly above the OECD average in mathematics.

Shanghai-China, Finland, Hong Kong-China and Singapore are the four highest performers in the PISA 2009 science assessment.

In science, New Zealand, Canada, Estonia, Australia, the Netherlands, Germany, Switzerland, the United Kingdom, Slovenia, Poland, Ireland and Belgium as well as the partner country and economies Chinese Taipei, Liechtenstein and Macao-China also perform significantly above the OECD average.

Some 14.6% of students in Shanghai-China and 12.3% of students in Singapore attain the highest levels of proficiency in all three assessment subjects.

High-level skills are critical for innovation and, as such, are key to economic growth and social development. On average, across OECD countries, 16.3% of students are top performers in at least one of the subject areas of science, mathematics or reading. However, only 4.1% of 15-year-old students are top performers in all three assessment subject areas.

Girls outperform boys in reading skills in every participating country.

Throughout much of the 20th century, concern about gender differences in education focused on girls' underachievement. More recently, however, the scrutiny has shifted to boys' underachievement in reading. In the PISA 2009 reading assessment, girls outperform boys in every participating country by an average, among OECD countries, of 39 PISA score points – equivalent to more than half a proficiency level or one year of schooling.

On average across OECD countries, boys outperform girls in mathematics by 12 score points while gender differences in science performance tend to be small, both in absolute terms and when compared with the large gender gap in reading performance and the more moderate gender gap in mathematics. The ranks of top-performing students are filled nearly equally with girls and boys. On average across OECD countries, 4.4% of girls and 3.8% of boys are top performers in all three subjects, and 15.6% of girls and 17.0% of boys are top performers in at least one subject area. While the gender gap among top-performing students is small in science (1% of girls and 1.5% of boys), it is significant in reading (2.8% of girls and 0.5% of boys) and in mathematics (3.4% of girls and 6.6% of boys).

Countries of similar prosperity can produce very different educational results.

The balance of proficiency in some of the richer countries in PISA looks very different from that of some of the poorer countries. In reading, for example, the ten countries in which the majority of students are at Level 1 or below, all in poorer parts of the world, contrast starkly in profile with the 34 OECD countries, where on average a majority attains at least Level 3. However, the fact that the best-performing country or economy in the 2009 assessment is Shanghai-China, with a GDP per capita well below the OECD average, underlines that low national income is not incompatible with strong educational performance. Korea, which is the best-performing OECD country, also has a GDP per capita below the OECD average. Indeed, while there is a correlation between GDP per capita and educational performance, this predicts only 6% of the differences in average student performance across countries. The other 94% of differences reflect the fact that two countries of similar prosperity can produce very different educational results. Results also vary when substituting spending per student, relative poverty or the share of students with an immigrant background for GDP per capita.

The following table summarises the key data of this volume. For each country, it shows the average score of 15-year-olds in reading, mathematics and science as well as on the subscales that were used to measure reading skills in greater detail. Cells shaded in light blue indicate values above the OECD average. Cells shaded in medium blue indicate values below the OECD average. Cells shaded in dark blue indicate values that are not statistically different from the OECD average.

.....→

■ Figure I. ■
COMPARING COUNTRIES' AND ECONOMIES' PERFORMANCE

Statistically significantly **above** the OECD average
Not statistically significantly different from the OECD average
Statistically significantly **below** the OECD average

	On the overall reading scale	On the reading subscales					On the mathematics scale	On the science scale
		Access and retrieve	*Integrate and interpret*	*Reflect and evaluate*	*Continuous texts*	*Non-continuous texts*		
OECD average	493	495	493	494	494	493	496	501
Shanghai-China	556	549	558	557	564	539	600	575
Korea	539	542	541	542	538	542	546	538
Finland	536	532	538	536	535	535	541	554
Hong Kong-China	533	530	530	540	538	522	555	549
Singapore	526	526	525	529	522	539	562	542
Canada	524	517	522	535	524	527	527	529
New Zealand	521	521	517	531	518	532	519	532
Japan	520	530	520	521	520	518	529	539
Australia	515	513	513	523	513	524	514	527
Netherlands	508	519	504	510	506	514	526	522
Belgium	506	513	504	505	504	511	515	507
Norway	503	512	502	505	505	498	498	500
Estonia	501	503	500	503	497	512	512	528
Switzerland	501	505	502	497	498	505	534	517
Poland	500	500	503	498	502	496	495	508
Iceland	500	507	503	496	501	499	507	496
United States	500	492	495	512	500	503	487	502
Liechtenstein	499	508	498	498	495	506	536	520
Sweden	497	505	494	502	499	498	494	495
Germany	497	501	501	491	496	497	513	520
Ireland	496	498	494	502	497	496	487	508
France	496	492	497	495	492	498	497	498
Chinese Taipei	495	496	499	493	496	500	543	520
Denmark	495	502	492	493	496	493	503	499
United Kingdom	494	491	491	503	492	506	492	514
Hungary	494	501	496	489	497	487	490	503
Portugal	489	488	487	496	492	488	487	493
Macao-China	487	493	488	481	488	481	525	511
Italy	486	482	490	482	489	476	483	489
Latvia	484	476	484	492	484	487	482	494
Slovenia	483	489	489	470	484	476	501	512
Greece	483	468	484	489	487	472	466	470
Spain	481	480	481	483	484	473	483	488
Czech Republic	478	479	488	462	479	474	493	500
Slovak Republic	477	491	481	466	479	471	497	490
Croatia	476	492	472	471	478	472	460	486
Israel	474	463	473	483	477	467	447	455
Luxembourg	472	471	475	471	471	472	489	484
Austria	470	477	471	463	470	472	496	494
Lithuania	468	476	469	463	470	462	477	491
Turkey	464	467	459	473	466	461	445	454
Dubai (UAE)	459	458	457	466	461	460	453	466
Russian Federation	459	469	467	441	461	452	468	478
Chile	449	444	452	452	453	444	421	447
Serbia	442	449	445	430	444	438	442	443
Bulgaria	429	430	436	417	433	421	428	439
Uruguay	426	424	423	436	429	421	427	427
Mexico	425	433	418	432	426	424	419	416
Romania	424	423	425	426	423	424	427	428
Thailand	421	431	416	420	423	423	419	425
Trinidad and Tobago	416	413	419	413	418	417	414	410
Colombia	413	404	411	422	415	409	381	402
Brazil	412	407	406	424	414	408	386	405
Montenegro	408	408	420	383	411	398	403	401
Jordan	405	394	410	407	417	387	387	415
Tunisia	404	393	393	427	408	393	371	401
Indonesia	402	399	397	409	405	399	371	383
Argentina	398	394	398	402	400	391	388	401
Kazakhstan	390	397	397	373	399	371	405	400
Albania	385	380	393	376	392	366	377	391
Qatar	372	354	379	376	375	361	368	379
Panama	371	363	372	377	373	359	360	376
Peru	370	364	371	368	374	356	365	369
Azerbaijan	362	361	373	335	362	351	431	373
Kyrgyzstan	314	299	327	300	319	293	331	330

Source: OECD, *PISA 2009 Database*.
StatLink http://dx.doi.org/10.1787/888932343342

OVERCOMING SOCIAL BACKGROUND: EQUITY IN LEARNING OPPORTUNITIES AND OUTCOMES

The best performing school systems manage to provide high-quality education to all students.

Canada, Finland, Japan, Korea and the partner economies Hong Kong-China and Shanghai-China all perform well above the OECD mean performance and students tend to perform well regardless of their own background or the school they attend. They not only have large proportions of students performing at the highest levels of reading proficiency, but also relatively few students at the lower proficiency levels.

Disadvantaged students may have access to more teachers, but not necessarily to the best teachers.

With the exception of Turkey, Slovenia, Israel and the United States, where socio-economically disadvantaged schools also tend to be deprived in terms of basic resources, such as larger student-staff ratios, OECD countries place at least an equal, if not a larger, number of teachers into socio-economically disadvantaged schools as they do in advantaged schools. But despite this fact, disadvantaged schools still report great difficulties in attracting qualified teachers. In other words, in disadvantaged schools, quantity of resources does not necessarily translate into quality of resources since, in general, more advantaged students attend schools that have a higher proportion of full-time teachers who have an advanced university degree. Findings from PISA suggest that, in terms of teacher resources, many students face the double liability of coming from a disadvantaged background and attending a school with lower quality resources. Many countries also show a strong relationship between the socio-economic background of students and their success at school and, in some of these countries, these disparities are magnified by large variations in the schools' socio-economic backgrounds, that is, in the backgrounds of the students' peers.

Home background influences educational success, and schooling often appears to reinforce its effects. Although poor performance in school does not automatically follow from a disadvantaged socio-economic background, the socio-economic background of students and schools does appear to have a powerful influence on performance.

Socio-economic disadvantage has many facets and cannot be ameliorated by education policy alone, much less in the short term. The educational attainment of parents can only gradually improve, and average family wealth depends on the long-term economic development of a country and on a culture that promotes individual savings. However, even if socio-economic background itself is hard to change, PISA shows that some countries succeed in reducing its impact on learning outcomes.

While most of the students who perform poorly in PISA are from socio-economically disadvantaged backgrounds, some peers from similar backgrounds excel in PISA, demonstrating that overcoming socio-economic barriers to achievement is possible. Resilient students come from the bottom quarter of the distribution of socio-economic background in their country and score in the top quarter among students from all countries with similar socio-economic background. In Finland, Japan, Turkey, Canada and Portugal and the partner country Singapore, between 39% and 48% of disadvantaged students are resilient. In Korea and the partner economy Macao-China, 50% and 56% of disadvantaged students can be considered resilient, and this percentage is 72% and 76% in partner economies Hong Kong-China and Shanghai-China, respectively.

Across OECD countries, a student from a more socio-economically advantaged background (among the top one seventh) outperforms a student from an average background by 38 score points, or about one year's worth of education, in reading. In New Zealand, France, the partner country Bulgaria and the partner economy Dubai (UAE), the gap between advantaged and disadvantaged students is more than 50 score points. On average across OECD countries, 14% of the differences in student reading performance within each country is associated with differences in students' socio-economic background. In Hungary and the partner countries Peru, Bulgaria and Uruguay, more than 20% of the differences in student performance is associated with differences in background.

Regardless of their own socio-economic background, students attending schools with a socio-economically advantaged intake tend to perform better than those attending schools with more disadvantaged peers.

In the majority of OECD countries, the effect of the school's economic, social and cultural status on students' performance far outweighs the effects of the individual student's socio-economic background. And the magnitude of the differences is striking. In Japan, the Czech Republic, Germany, Belgium and Israel and the partner countries Trinidad and Tobago and Liechtenstein, the performance gap between two students with similar socio-economic backgrounds, one of whom attends a school with an average socio-economic background and the other attends a school with an advantaged socio-economic background (among the top 16% in the country), is equivalent to more than 50 score points, on average, or more than a year's worth of education.

Across OECD countries, first-generation students – those who were born outside the country of assessment and who also have foreign-born parents – score, on average, 52 score points below students without an immigrant background

In New Zealand, Canada and Switzerland, 20% to 25% of students are from an immigrant background while the proportions are even higher in Liechtenstein (30%), Hong Kong-China (39%), Luxembourg (40%) and Qatar (46%). In Macao-China and Dubai (UAE), that percentage is at least 70%. There is no positive association between the size of the immigrant student population and average performance at the country or economy level, and there is also no relationship between the proportion of students with an immigrant background and the performance gaps between native and immigrant students. These findings contradict the assumption that high levels of immigration will inevitably lower the mean performance of school systems.

Students in urban schools perform better than students in other schools, even after accounting for differences in socio-economic background.

In Turkey, the Slovak Republic, Chile, Mexico and Italy, as well as the partner countries Peru, Tunisia, Albania, Argentina and Romania, the performance gap between students in urban schools and those in rural schools is more than 45 score points after accounting for differences in socio-economic background. This is more than one year of education across OECD countries. That gap is 80 score points or more – or two years of schooling – in Hungary and in the partner countries Bulgaria, Kyrgyzstan and Panama. However, this pattern is not observed in Belgium, Finland, Germany, Greece, Iceland, Ireland, Israel, the Netherlands, Poland, Sweden, the United Kingdom and the United States.

On average across the OECD, 17% of students come from single-parent families and they score five score points lower than students from other types of families after accounting for socio-economic background.

Among OECD countries, the gap is particularly large in the United States where, after accounting for socio-economic background, the performance difference between students from single-parent families and those from other types of families stands at 23 score points. In Ireland, Poland and Mexico, the gap is 13 score points and in Belgium, Japan and Luxembourg it is 10 score points, double the average among OECD countries. Among partner countries and economies, students from single-parent families score 10 points lower than peers from other types of families after accounting for socio-economic background.

Parents' engagement with their children's reading life has a positive impact on their children's reading performance. Students whose parents reported that they had read a book with their child "every day or almost every day" or "once or twice a week" during the first year of primary school performed higher in PISA 2009 than students whose parents reported that they had done this "never or almost never" or "once or twice a month". On average across the 14 countries that had collected information on this question, the difference is 25 score points, but it ranges from 4 score points in the partner country Lithuania to 63 score points in New Zealand. Also, 15-year-olds whose parents discuss political or social issues once a week or more score 28 score points higher than those whose parents do not, or who talk about these issues less often. The performance advantage was largest in Italy, at 42 score points, and smallest in the partner economy Macao-China, and it is observed across all countries.

The following table summarises key data. For each country, it shows the average score of 15-year-olds in reading and seven equity measures from PISA: i) and ii) two measures focusing on those who achieve the baseline level of proficiency in PISA: the proportion of boys and girls who score below Level 2; iii) a measure of those who overcome socio-economic disadvantaged and do best given their weak prospects, the proportion of resilient students; iv) and v) two measures of the relationship between student background and performance: the percentage of variation in student performance explained by the student's socio-economic background and the slope of the socio-economic gradient, the average gap in performance between students from different socio-economic backgrounds; and vi) and vii) two measures of equality in the distribution of educational resources, namely the quality and quantity of teachers. For the first five measures, cells shaded in light blue indicate values of quality or equity above the OECD average. Cells shaded in medium blue indicate values of equity below the OECD average. Cells shaded in dark blue indicate values that are not statistically different from the OECD average. In the last two columns, cells shaded in light blue indicate that disadvantaged schools are more likely to have more or better resources. Cells shaded in medium blue that advantaged schools are more likely to have more or better resources. Cells shaded in dark blue indicate values where disadvantaged and advantaged schools are equally likely to have more or better resources. In these two last columns, estimates in bold indicate that they are statistically different from the OECD average.

.....→

■ Figure II. ■
SUMMARY OF PISA MEASURES OF EDUCATIONAL EQUITY

Higher quality or equity than OECD average

At OECD average (no statistically significant difference)

Lower quality or equity than OECD average

Disadvantaged schools are more likely to have more or better resources, in **bold** if relationship is statistically different from the OECD average

Within country correlation is not statistically significant

Advantaged schools are more likely to have more or better resources, in **bold** if relationship is statistically different from the OECD average

		Mean reading score	Percentage of boys below proficiency Level 2	Percentage of girls below proficiency Level 2	Percentage of resilient students	Percentage of variance in student performance explained by students' socio-economic background	Slope of the socio-economic gradient	Correlation between the socio-economic background of schools and the percentage of teachers with university-level (ISCED 5A) among all full-time teachers	Correlation between socio-economic background of schools and the student/teacher ratio
	OECD average	493	25	13	8	14	38	0.15	-0.15
OECD	Korea	539	9	2	14	11	32	-0.03	0.30
OECD	Finland	536	13	3	11	8	31	-0.01	0.08
OECD	Canada	524	14	6	10	9	32	0.03	0.09
OECD	New Zealand	521	21	8	9	17	52	0.07	0.11
OECD	Japan	520	19	8	11	9	40	0.20	**0.38**
OECD	Australia	515	20	9	8	13	46	0.02	-0.07
OECD	Netherlands	508	18	11	8	13	37	**0.62**	**0.38**
OECD	Belgium	506	21	14	8	19	47	**0.58**	**0.66**
OECD	Norway	503	21	8	6	9	36	0.15	0.19
OECD	Estonia	501	19	7	9	8	29	0.00	**0.43**
OECD	Switzerland	501	22	11	8	14	40	0.24	0.06
OECD	Poland	500	23	7	9	15	39	-0.05	0.01
OECD	Iceland	500	24	10	7	6	27	**0.30**	**0.40**
OECD	United States	500	21	14	7	17	42	0.10	**-0.17**
OECD	Sweden	497	24	10	6	13	43	-0.04	0.12
OECD	Germany	497	24	13	6	18	44	-0.02	**0.28**
OECD	Ireland	496	23	11	7	13	39	-0.08	**0.49**
OECD	France	496	26	14	8	17	51	w	w
OECD	Denmark	495	19	11	6	15	36	0.16	**0.27**
OECD	United Kingdom	494	23	14	6	14	44	-0.03	-0.10
OECD	Hungary	494	24	11	6	26	48	0.07	0.02
OECD	Portugal	489	25	11	10	17	30	0.04	**0.39**
OECD	Italy	486	29	13	8	12	32	0.13	**0.50**
OECD	Slovenia	483	31	11	6	14	39	**0.55**	**-0.25**
OECD	Greece	483	30	13	7	12	34	0.24	0.25
OECD	Spain	481	24	15	9	14	29	m	**0.45**
OECD	Czech Republic	478	31	14	5	12	46	**0.37**	0.08
OECD	Slovak Republic	477	32	13	5	15	41	**-0.21**	0.00
OECD	Israel	474	34	19	6	13	43	0.20	**-0.20**
OECD	Luxembourg	472	33	19	5	18	40	**0.39**	**0.28**
OECD	Austria	470	35	20	5	17	48	**0.64**	-0.07
OECD	Turkey	464	33	15	10	19	29	0.04	**-0.26**
OECD	Chile	449	36	25	6	19	31	**0.25**	-0.05
OECD	Mexico	425	46	34	7	14	25	-0.04	0.03
Partners	Shanghai-China	556	7	2	19	12	27	**0.32**	**-0.13**
Partners	Hong Kong-China	533	11	5	18	5	17	0.12	0.02
Partners	Singapore	526	16	9	12	15	47	**0.22**	**-0.14**
Partners	Liechtenstein	499	21	9	9	8	26	**0.57**	**0.70**
Partners	Chinese Taipei	495	22	10	10	12	36	0.29	-0.07
Partners	Macao-China	487	21	9	13	2	12	**-0.18**	**0.17**
Partners	Latvia	484	27	9	8	10	29	0.19	**0.38**
Partners	Croatia	476	31	13	7	11	32	**0.28**	0.32
Partners	Lithuania	468	35	13	5	14	33	0.19	0.21
Partners	Dubai (UAE)	459	41	21	3	14	51	**-0.01**	**-0.27**
Partners	Russian Federation	459	36	19	5	11	37	0.31	**0.29**
Partners	Serbia	442	43	23	4	10	27	0.06	0.11
Partners	Bulgaria	429	52	29	2	20	51	0.17	0.21
Partners	Uruguay	426	51	34	4	21	37	0.08	0.13
Partners	Romania	424	51	30	2	14	36	0.11	-0.02
Partners	Thailand	421	55	33	7	13	22	0.16	-0.02
Partners	Trinidad and Tobago	416	55	34	5	10	38	**0.56**	**0.38**
Partners	Colombia	413	50	45	6	17	28	-0.08	-0.14
Partners	Brazil	412	56	43	6	13	28	0.03	**-0.20**
Partners	Montenegro	408	61	37	2	10	31	**0.38**	**0.33**
Partners	Jordan	405	62	34	3	8	24	-0.02	0.06
Partners	Tunisia	404	58	43	7	8	19	0.20	-0.02
Partners	Indonesia	402	65	42	6	8	17	0.16	**-0.16**
Partners	Argentina	398	59	45	3	20	40	0.22	-0.02
Partners	Kazakhstan	390	67	50	1	12	38	**0.34**	**0.44**
Partners	Albania	385	69	44	3	11	31	**0.38**	0.15
Partners	Qatar	372	72	54	1	4	25	**-0.07**	0.11
Partners	Panama	371	72	59	2	18	31	-0.13	0.03
Partners	Peru	370	70	60	1	27	41	**0.48**	-0.02
Partners	Azerbaijan	362	78	68	1	7	21	**0.44**	0.23
Partners	Kyrgyzstan	314	88	78	0	15	40	**0.35**	0.27

Countries are ranked in descending order of the mean score in reading, separately for OECD and partner countries and economies.
Source: OECD, *PISA 2009 Database*, Tables I.2.2., II.1.1., II.2.3., II.3.2 and II.3.3.
StatLink http://dx.doi.org/10.1787/888932343684

LEARNING TO LEARN: STUDENT ENGAGEMENT, STRATEGIES AND PRACTICES

PISA results show that mastering strategies that assist learning, such as methods to remember and understand or summarise texts and reading widely, are essential if students are to become proficient readers. Practicing reading by reading for enjoyment is most closely associated with better outcomes when it is accompanied by high levels of critical thinking and strategic learning. Across OECD countries, students who have low levels of awareness about which strategies are most effective for understanding, remembering and summarising information are less proficient readers than those who have high levels of awareness about these strategies, regardless of their reading habits.

In all countries, students who enjoy reading the most perform significantly better than students who enjoy reading the least.

There has been considerable debate as to what type of reading may be most effective in fostering reading skills and improving reading performance. The results from PISA suggest that, although students who read fiction are more likely to achieve high scores, it is students who read a wide variety of material who perform particularly well in reading. Compared with not reading for enjoyment at all, reading fiction for enjoyment appears to be positively associated with higher scores in the PISA 2009 reading assessment, while reading comic books is associated with little improvement in reading proficiency in some countries, and with lower overall reading performance in other countries. Also, students who are extensively engaged in online reading activities, such as reading e-mails, chatting on line, reading news on line, using an online dictionary or encyclopaedia, participating in online group discussions and searching for information on line, are generally more proficient readers than students who do little online reading.

On average across OECD countries, 37% of students – and 45% or more in Austria, the Netherlands, and Luxembourg – report that they do not read for enjoyment at all.

In all but a few countries, students who use appropriate strategies to understand and remember what they read, such as underlining important parts of the texts or discussing what they read with other people, perform at least 73 points higher in the PISA assessment – that is, one full proficiency level or nearly two full school years – than students who use these strategies the least. In Belgium, Switzerland and Austria, the quarter of students who use these strategies the most score an average of 110 points higher than the quarter of students who use them the least. That translates into a difference of roughly one-and-a-half proficiency levels or nearly three years of formal schooling.

In all countries, boys are not only less likely than girls to say that they read for enjoyment, they also have different reading habits when they do read for pleasure.

Most boys and girls in the countries that took part in PISA 2009 sit side by side in the same classrooms and work with similar teachers. Yet, PISA reveals that in OECD countries, boys are on average 39 points behind girls in reading, the equivalent of one year of schooling. PISA suggests that differences in how boys and girls approach learning and how engaged they are in reading account for most of the gap in reading performance between boys and girls, so much so that this gap could be predicted to shrink by 14 points if boys approached learning as positively as girls, and by over 20 points if they were as engaged in reading as girls. This does not mean that if boys' engagement and awareness of learning strategies rose by this amount the increase would automatically translate into respective performance gains, since PISA does not measure causation. But since most of the gender gap can be explained by boys being less engaged, and less engaged students show lower performance, then policy makers should look for more effective ways of increasing boys' interest in reading at school or at home.

PISA reveals that, although girls have higher mean reading performance, enjoy reading more and are more aware of effective strategies to summarise information than boys, the differences within genders are far greater than those between the genders. Moreover, the size of the gender gap varies considerably across countries, suggesting that boys and girls do not have inherently different interests and academic strengths, but that these are mostly acquired and socially induced. The large gender gap in reading is not a mystery: it can be attributed to differences that have been identified in the attitudes and behaviours of boys and girls.

Girls are more likely than boys to be frequent readers of fiction, and are also more likely than boys to read magazines. However, over 65% of boys regularly read newspapers for enjoyment and only 59% of girls do so. Although relatively few students say that they read comic books regularly, on average across OECD countries, 27% of boys read comic books several times a month or several times a week, while only 18% of girls do so.

High-performing countries are also those whose students generally know how to summarise information.

Across OECD countries, the difference in reading performance between those students who know the most about which strategies are best for summarising information and those who know the least is 107 score points. And students who say that they begin the learning process by figuring out what they need to learn, then ensure that

they understand what they read, figure out which concepts they have not fully grasped, try to remember the most important points in a text and look for additional clarifying information when they do not understand something they have read, tend to perform better on the PISA reading scale than those who do not.

While factors such as predisposition, temperament, peer pressure and socialisation may contribute to boys having less interest in reading than girls, boys could be encouraged to enjoy reading more and to read more for enjoyment.

PISA results suggest that boys would be predicted to catch up with girls in reading performance if they had higher levels of motivation to read and used effective learning strategies. In Finland, for example, if boys were equally aware as girls of the most effective ways of summarising complex information in their reading, their scores in the PISA assessment would be predicted to be 23 points higher. Similarly, in most of the countries that participated in PISA 2009, if the most socio-economically disadvantaged students had the same levels of awareness about these strategies as their most advantaged peers, their reading performance would be predicted to be at least 15 points higher.

Across OECD countries, if socio-economically disadvantaged students were as aware of effective strategies to summarise information as advantaged students, the performance gap between the two groups of students could be 20% narrower. The poor reading proficiency seen among socio-economically disadvantaged boys is of particular concern because, without the ability to read well enough to participate fully in society, these students and their future families will have fewer opportunities to escape a cycle of poverty and deprivation. On average in the OECD area, socio-economically disadvantaged boys would be predicted to perform 28 points higher in reading if they had the same level of awareness of effective summarising strategies as socio-economically advantaged girls and 35 points higher if they enjoyed reading as much as socio-economically advantaged girls.

In recent years, the gender gap in reading engagement has widened, as has the gender gap in reading performance.

Changing students' attitudes and behaviours may be inherently more difficult than providing equal access to high quality teachers and schools, two of the factors that explain the low performance of socio-economically disadvantaged students – an area where PISA shows that over the past decade, some countries have achieved significant progress.

The following table provides selected results.

- The first column shows students' mean reading scores.
- The second column shows the percentage of students who reported high levels of awareness about effective learning strategies and who regularly read a wide range of materials, including fiction and non-fiction books or at least magazines and newspapers, for enjoyment (considered 'wide and deep' or 'narrow and deep' readers).
- The third column shows the score point differences in reading between boys and girls, with negative numbers indicating an advantage for boys and positive numbers indicating an advantage for girls.
- The fourth column shows gender differences in the percentage of 'wide and deep' or 'narrow and deep' readers.
- The fifth column shows the portion of the gender gap that would be predicted to be closed if boys had the same level of enjoyment of reading as girls.
- The sixth column shows the score point difference between the top and bottom quarters of the socio-economic distribution of students.
- The seventh column shows the differences in the share of students who are 'wide and deep' or 'narrow and deep' readers between the top and bottom quarters of the socio-economic distribution of students. Larger numbers indicate a higher share of 'wide and deep' or 'narrow and deep' readers among socio-economically advantaged students.
- The last column shows the portion of the socio-economic gap in reading performance that would be predicted to be closed if socio-economically disadvantaged students had the same level of awareness of effective reading strategies (here, summarising strategies) as socio-economically advantaged students.

Values that are larger than the OECD average are shown in light blue; while values that are smaller than the OECD average are shown in medium blue and values that are not statistically different from the OECD average are shown in dark blue.

.....→

■ Figure III. ■
COMPARING THE CONTRIBUTION OF STUDENTS' ENGAGEMENT IN READING AND APPROACHES TO LEARNING TO READING PERFORMANCE AND EQUITY

Statistically significantly **above** the OECD average
Not statistically significantly different from the OECD average
Statistically significantly **below** the OECD average

		Mean Reading Score	Percentage of "wide and deep" and "narrow and deep readers"	Difference in reading performance (G – B)	Difference in the percentage of girls and boys who can be considered "wide and deep" and "narrow and deep" readers (G – B)	Proportion of the overall gender gap that could be closed if boys enjoyed reading as much as girls	Socio-economic differences in reading performance (top – bottom quarter of ESCS)	Socio-economic differences in the percentage of students who are "wide and deep" and "narrow and deep" (top – bottom quarter of ESCS)	Proportion of the socio-economic gap that could be closed if socio-economically disadvanatged students had values on the index of summarising as socio-economically advantaged students
		Mean Score	%	Dif.	Dif.	%	Dif.	Dif.	%
	OECD average	493	45	39	11	61	89	17	20
OECD	Korea	539	35	35	5	30	70	32	27
OECD	Finland	536	60	55	20	64	62	17	27
OECD	Canada	524	37	34	14	86	68	15	13
OECD	New Zealand	521	37	46	11	63	102	14	20
OECD	Japan	520	54	39	6	33	73	18	25
OECD	Australia	515	35	37	9	76	91	16	22
OECD	Netherlands	508	34	24	9	102	83	23	23
OECD	Belgium	506	46	27	3	81	116	23	27
OECD	Norway	503	56	47	14	52	70	17	22
OECD	Estonia	501	61	44	14	65	60	12	17
OECD	Switzerland	501	54	39	11	76	94	22	24
OECD	Poland	500	50	50	20	49	88	17	20
OECD	Iceland	500	49	44	20	58	62	12	18
OECD	United States	500	30	25	7	95	105	12	14
OECD	Sweden	497	43	46	16	68	91	19	18
OECD	Germany	497	41	40	0	80	105	21	23
OECD	Ireland	496	45	39	14	48	86	5	15
OECD	France	496	46	40	1	54	110	20	21
OECD	Denmark	495	48	29	8	75	80	21	20
OECD	United Kingdom	494	40	25	10	90	91	11	19
OECD	Hungary	494	52	38	15	65	118	20	20
OECD	Portugal	489	43	38	9	61	87	17	24
OECD	Italy	486	39	46	7	56	85	15	20
OECD	Slovenia	483	45	55	16	42	87	15	20
OECD	Greece	483	34	47	1	54	90	18	13
OECD	Spain	481	38	29	6	73	83	22	15
OECD	Czech Republic	478	47	48	14	59	84	12	23
OECD	Slovak Republic	477	52	51	16	35	87	13	18
OECD	Israel	474	36	42	17	44	102	14	19
OECD	Luxembourg	472	50	39	8	70	114	16	19
OECD	Austria	470	50	41	10	70	102	20	23
OECD	Turkey	464	38	43	12	25	92	16	11
OECD	Chile	449	37	22	17	57	91	19	15
OECD	Mexico	425	36	25	6	27	82	16	17
Partners	Shanghai-China	556	59	40	5	31	74	21	11
Partners	Hong Kong-China	533	41	33	7	44	46	15	14
Partners	Singapore	526	59	31	11	81	98	19	17
Partners	Liechtenstein	499	49	32	14	76	62	25	34
Partners	Chinese Taipei	495	44	37	6	53	76	24	17
Partners	Macao-China	487	44	34	11	38	25	18	23
Partners	Latvia	484	45	47	20	52	63	16	19
Partners	Croatia	476	53	51	19	40	74	17	19
Partners	Lithuania	468	53	59	21	47	83	20	17
Partners	Dubai (UAE)	459	56	51	10	38	102	15	19
Partners	Russian Federation	459	46	45	16	43	78	9	16
Partners	Serbia	442	43	39	16	37	67	18	24
Partners	Bulgaria	429	42	61	20	27	132	22	16
Partners	Uruguay	426	35	42	15	30	116	15	20
Partners	Romania	424	44	43	13	23	85	16	17
Partners	Thailand	421	40	38	12	22	63	15	8
Partners	Trinidad and Tobago	416	49	58	19	26	92	10	19
Partners	Colombia	413	46	9	10	41	89	12	19
Partners	Brazil	412	37	29	14	34	83	13	16
Partners	Montenegro	408	42	53	8	30	80	18	15
Partners	Jordan	405	34	57	14	12	66	12	9
Partners	Tunisia	404	45	31	11	0	63	12	4
Partners	Indonesia	402	43	37	11	8	45	18	13
Partners	Argentina	398	40	37	14	24	122	15	15
Partners	Kazakhstan	390	46	43	13	-1	84	12	12
Partners	Albania	385	50	62	17	38	77	15	10
Partners	Qatar	372	42	50	8	23	56	9	14
Partners	Panama	371	37	33	13	10	108	10	13
Partners	Peru	370	50	22	9	19	129	20	14
Partners	Azerbaijan	362	32	24	6	21	50	12	4
Partners	Kyrgyzstan	314	34	53	7	10	94	18	14

Countries are ranked by their mean reading score.
Source: OECD, *PISA 2009 Database*.
StatLink http://dx.doi.org/10.1787/888932360309

WHAT MAKES A SCHOOL SUCCESSFUL? RESOURCES, POLICIES AND PRACTICES

Since school is where most learning happens, what happens in school has a direct impact on learning. In turn, what happens in school is influenced by the resources, policies and practices approved at higher administrative levels in a country's education system.

Successful school systems – those that perform above average and show below-average socio-economic inequalities – provide all students, regardless of their socio-economic backgrounds, with similar opportunities to learn.

Systems that show high performance and an equitable distribution of learning outcomes tend to be comprehensive, requiring teachers and schools to embrace diverse student populations through personalised educational pathways. In contrast, school systems that assume that students have different destinations with different expectations and differentiation in terms of how they are placed in schools, classes and grades often show less equitable outcomes without an overall performance advantage.

Earlier PISA assessments showed these expectations to be mirrored in how students perceived their own educational future. The results of these differences can also be seen in the distribution of student performance within countries and in the impact that socio-economic background has on learning outcomes:

- In countries, and in schools within countries, where more students repeat grades, overall results tend to be worse.
- In countries where more students repeat grades, socio-economic differences in performance tend to be wider, suggesting that people from lower socio-economic groups are more likely to be negatively affected by grade repetition.
- In countries where 15-year-olds are divided into more tracks based on their abilities, overall performance is not enhanced, and the younger the age at which selection for such tracks first occurs, the greater the differences in student performance, by socio-economic background, by age 15, without improved overall performance.
- In school systems where it is more common to transfer weak or disruptive students out of a school, performance and equity both tend to be lower. Individual schools that make more use of transfers also perform worse in some countries.

These associations account for a substantial amount of the differences in the outcomes of schooling systems. For example, the frequency with which students are transferred across schools is associated with a third of the variation in country performance. This does not necessarily mean that if transfer policies were changed, a third of country differences in reading performance would disappear, since PISA does not measure cause and effect. Transferring pupils who do badly may be partly a symptom, rather than a cause, of schools and school systems that are not producing satisfactory results, especially for lower-achieving students. It is worth noting that the schools with lower transfer rates tend to have greater autonomy and other means of addressing these challenges. The cluster of results listed above suggests that, in general, school systems that seek to cater to different students' needs through a high level of differentiation in the institutions, grade levels and classes have not succeeded in producing superior overall results, and in some respects they have lower-than-average and more socially unequal performance.

Most successful school systems grant greater autonomy to individual schools to design curricula and establish assessment policies, but these school systems do not necessarily allow schools to compete for enrolment.

The incentive to deliver good results for all students is not just a matter of how a school's student body is defined. It also depends on the ways in which schools are held accountable for their results and what forms of autonomy they are allowed to have – and how that could help influence their performance. PISA has looked at accountability both in terms of the information that is made available about performance and in terms of the use made of that information – whether by administrative authorities through rewards or control systems, or by parents, for example through their choice of school. Thus the issues of autonomy, evaluation, governance and choice interact in providing a framework in which schools are given the incentives and the capacity to improve. PISA 2009 finds that:

- In countries where schools have greater autonomy over what is taught and how students are assessed, students tend to perform better.
- Within countries where schools are held to account for their results through posting achievement data publicly, schools that enjoy greater autonomy in resource allocation tend to do better than those with less autonomy. However, in countries where there are no such accountability arrangements, the reverse is true.
- Countries that create a more competitive environment in which many schools compete for students do not systematically produce better results.

- Within many countries, schools that compete more for students tend to have higher performance, but this is often accounted for by the higher socio-economic status of students in these schools. Parents with a higher socio-economic status are more likely to take academic performance into consideration when choosing schools.
- In countries that use standards-based external examinations, students tend to do better overall, but there is no clear relationship between performance and the use of standardised tests or the public posting of results at the school level. However, performance differences between schools with students of different social backgrounds are, on average, lower in countries that use standardised tests.

After accounting for the socio-economic and demographic profiles of students and schools, students in OECD countries who attend private schools show performance that is similar to that of students enrolled in public schools.

On average, socio-economically disadvantaged parents are over 13 percentage points more likely than socio-economically advantaged parents to report that they consider "low expenses" and "financial aid" as very important determinants in choosing a school. If children from socio-economically disadvantaged backgrounds cannot attend high-performing schools because of financial constraints, then school systems that offer parents more choice of schools for their children will necessarily be less effective in improving the performance of all students.

School systems considered successful tend to prioritise teachers' pay over smaller classes.

School systems differ in the amount of time, human, material and financial resources they invest in education. Equally important, school systems also vary in how these resources are spent:

- At the level of the school system and net of the level of national income, PISA shows that higher teachers' salaries, but not smaller class sizes, are associated with better student performance. Teachers' salaries are related to class size in that if spending levels are similar, school systems often make trade-offs between smaller classes and higher salaries for teachers. The findings from PISA suggest that systems prioritising higher teachers' salaries over smaller classes tend to perform better, which corresponds with research showing that raising teacher quality is a more effective route to improved student outcomes than creating smaller classes.
- Within countries, schools with better resources tend to do better only to the extent that they also tend to have more socio-economically advantaged students. Some countries show a strong relationship between schools' resources and their socio-economic and demographic background, which indicates that resources are inequitably distributed according to schools' socio-economic and demographic profiles.
- In other respects, the overall lack of a relationship between resources and outcomes does not show that resources are not important, but that their level does not have a systematic impact within the prevailing range. If most or all schools have the minimum resource requirements to allow effective teaching, additional material resources may make little difference to outcomes.

In more than half of all OECD countries, over 94% of 15-year-old students reported that they had attended pre-primary school for at least some time.

Students who had attended pre-primary school tend to perform better than students who have not. This advantage is greater in school systems where pre-primary education lasts longer, where there are smaller pupil-to-teacher ratios at the pre-primary level and where there is higher public expenditure per pupil at that level of education. Across all participating countries, school systems with a higher proportion of students who had attended pre-primary education tend to perform better.

Schools with better disciplinary climates, more positive behaviour among teachers and better teacher-student relations tend to achieve higher scores in reading.

Across OECD countries, 81% of students report that they feel they can work well in class most of the time, 71% report that they never, or only in some classes, feel that other students don't listen, and 72% say that their teacher never, or only in some lessons, has to wait a long time before students settle down to learn.

Meanwhile, 28% of students in OECD countries are enrolled in schools whose principals report that their teaching staff's resistance to change negatively affects students or that students' needs are not met; 23% attends schools whose principals report that students are not encouraged by teachers in the school; 22% attend schools whose principals believe that learning is hindered by low teacher expectations; and 17% of students attend schools whose principals say that teacher absenteeism hampers learning.

■ Figure IV.a ■

Comparing school systems…

…some characteristics are related to better reading performance and/or greater equity throughout the system. *For example…*

		Reading performance	Strength of the relationship between students' socio-economic background and reading performance	Percentage of students in schools that transfer students to other schools due to low achievement, behavioural problems or special learning needs	Index of school responsibility for curriculum and assessment[1]	Existence of standards-based external examinations[2]
				The lower the rate of student transfers, the better the performance and the more equitable the educational opportunities.	The greater the school autonomy, the better the performance.	The greater the prevalence of standards-based external examinations, the better the performance.
		Mean score	% variance explained	%	Mean index	Proportion
OECD	Australia	515	12.7	2.9	0.17	0.81
	Austria	470	16.6	52.2	-0.31	0.00
	Belgium	506	19.3	48.3	-0.17	0.00
	Canada	524	8.6	12.6	-0.66	0.51
	Chile	449	18.7	24.0	-0.09	0.00
	Czech Republic	478	12.4	22.1	0.92	1.00
	Denmark	495	14.5	6.1	0.05	1.00
	Estonia	501	7.6	9.9	0.22	1.00
	Finland	536	7.8	1.7	-0.15	1.00
	France	496	16.7	w	w	w
	Germany	497	17.9	24.0	-0.25	0.35
	Greece	483	12.5	42.2	-1.25	0.00
	Hungary	494	26.0	14.1	0.11	1.00
	Iceland	500	6.2	0.0	0.23	1.00
	Ireland	496	12.6	0.8	0.01	1.00
	Israel	474	12.5	23.1	-0.01	1.00
	Italy	486	11.8	20.1	0.20	1.00
	Japan	520	8.6	8.4	1.06	1.00
	Korea	539	11.0	6.3	0.79	1.00
	Luxembourg	472	18.0	68.0	-0.86	1.00
	Mexico	425	14.5	32.8	-0.92	0.00
	Netherlands	508	12.8	15.2	1.04	1.00
	New Zealand	521	16.6	2.7	0.81	1.00
	Norway	503	8.6	1.4	-0.57	1.00
	Poland	500	14.8	8.2	0.31	1.00
	Portugal	489	16.5	0.8	-0.93	0.00
	Slovak Republic	477	14.6	30.2	0.08	1.00
	Slovenia	483	14.3	21.8	-0.38	1.00
	Spain	481	13.6	7.3	-0.48	0.00
	Sweden	497	13.4	3.4	0.21	0.00
	Switzerland	501	14.1	21.3	-0.62	0.00
	Turkey	464	19.0	35.1	-1.04	1.00
	United Kingdom	494	13.7	2.2	0.83	1.00
	United States	500	16.8	12.6	-0.20	0.07
	OECD average	493	14.0	17.6	-0.06	0.66
Partners	Albania	385	10.7	17.4	-0.42	m
	Argentina	398	19.6	14.6	-0.57	0.00
	Azerbaijan	362	7.4	15.4	-0.64	1.00
	Brazil	412	13.0	13.8	-0.56	0.00
	Bulgaria	429	20.2	34.5	-0.91	1.00
	Colombia	413	16.6	41.3	-0.21	1.00
	Croatia	476	11.0	18.3	-0.93	1.00
	Dubai (UAE)	459	14.2	20.1	0.15	1.00
	Hong Kong-China	533	4.5	12.1	0.92	1.00
	Indonesia	402	7.8	40.5	0.13	1.00
	Jordan	405	7.9	46.0	-1.20	1.00
	Kazakhstan	390	12.0	12.9	-0.98	m
	Kyrgyzstan	314	14.6	38.2	-0.25	1.00
	Latvia	484	10.3	14.7	-0.54	1.00
	Liechtenstein	499	8.4	0.0	-0.05	1.00
	Lithuania	468	13.6	6.8	0.13	1.00
	Macao-China	487	1.8	47.7	0.86	0.00
	Montenegro	408	10.0	7.1	-0.97	1.00
	Panama	371	18.1	32.4	-0.60	0.00
	Peru	370	27.4	26.8	-0.19	0.00
	Qatar	372	4.0	44.9	-0.61	0.00
	Romania	424	13.6	40.1	-0.36	0.78
	Russian Federation	459	11.3	13.6	-0.36	1.00
	Serbia	442	9.8	29.8	-1.03	0.26
	Shanghai-China	556	12.3	15.3	-0.09	1.00
	Singapore	526	15.3	0.7	-0.09	1.00
	Chinese Taipei	495	11.8	37.2	0.38	1.00
	Thailand	421	13.3	9.8	0.76	0.79
	Trinidad and Tobago	416	9.7	13.8	-0.61	1.00
	Tunisia	404	8.1	26.0	-1.29	0.00
	Uruguay	426	20.7	8.5	-0.99	0.00

1. In this index, 0 is the OECD average. The higher the value, the greater the autonomy granted to schools in that country/economy.
2. Values between 0 and 1 indicate that standards-based external examinations exist in some parts of the school system, but not throughout.

■ Figure IV.b ■

Comparing schools within each country/economy…

…the climate in class is also associated with reading performance. *For example...*

Students in schools with better displinary climates tend to perform better.

Change in the reading score per unit of the index of disciplinary climate[1]

	Average index[1]
Japan	0.75
Singapore	0.12
Romania	0.43
United States	0.16
Dubai (UAE)	0.13
Italy	0.03
Australia	-0.07
Croatia	-0.13
Kyrgyzstan	0.35
Shanghai-China	0.45
United Kingdom	0.11
New Zealand	-0.12
Chinese Taipei	0.09
Montenegro	0.28
Peru	0.19
Serbia	-0.02
Slovenia	-0.11
Colombia	0.19
Lithuania	0.30
Macao-China	0.11
Hong Kong-China	0.37
Bulgaria	0.02
Trinidad and Tobago	-0.02
Qatar	-0.02
Hungary	-0.02
Iceland	-0.05
Kazakhstan	0.78
Austria	0.11
Azerbaijan	0.57
Ireland	-0.03
OECD average	0.00
Albania	0.53
Czech Republic	-0.18
Norway	-0.24
Turkey	0.03
Brazil	-0.18
Korea	0.38
Thailand	0.33
Luxembourg	-0.21
Denmark	0.01
Estonia	0.05
France	-0.20
Switzerland	0.09
Canada	-0.08
Panama	0.04
Mexico	0.11
Slovak Republic	-0.02
Sweden	-0.03
Greece	-0.40
Spain	0.09
Israel	0.08
Jordan	0.23
Poland	0.07
Belgium	-0.07
Chile	-0.10
Uruguay	-0.01
Russian Federation	0.44
Portugal	0.19
Latvia	0.25
Germany	0.25
Netherlands	-0.28
Liechtenstein	0.13
Finland	-0.29
Indonesia	0.26
Argentina	-0.26
Tunisia	-0.19

0 10 20 30 40
Score points

In most countries/economies, students in schools with better teacher-student relations tend to perform better.

Change in the reading score per unit of the index of teacher-student relations[2]

	Average index[2]
Australia	0.11
Norway	-0.17
Japan	-0.42
New Zealand	0.19
Iceland	0.17
Ireland	-0.08
Finland	-0.16
Estonia	-0.04
Denmark	0.18
Sweden	0.15
United Kingdom	0.12
Qatar	0.18
Jordan	0.26
Canada	0.32
Albania	0.67
France	-0.15
United States	0.32
Russian Federation	0.07
Shanghai-China	0.21
Czech Republic	-0.24
Latvia	-0.03
Hong Kong-China	-0.03
OECD average	0.00
Chinese Taipei	0.03
Korea	-0.27
Greece	-0.18
Trinidad and Tobago	0.16
Singapore	0.24
Brazil	0.19
Slovak Republic	-0.16
Poland	-0.35
Dubai (UAE)	0.36
Lithuania	0.14
Mexico	0.14
Chile	0.09
Portugal	0.37
Spain	-0.03
Tunisia	0.02
Netherlands	-0.11
Hungary	-0.01
Slovenia	-0.42
Romania	0.02
Peru	0.29
Turkey	0.44
Macao-China	-0.24
Liechtenstein	0.08
Switzerland	0.24
Belgium	-0.04
Israel	0.05
Italy	-0.06
Austria	0.00
Thailand	0.10
Luxembourg	-0.04
Azerbaijan	0.53
Germany	0.01
Kazakhstan	0.41
Croatia	-0.17
Kyrgyzstan	0.27
Panama	0.46
Bulgaria	-0.01
Colombia	0.34
Indonesia	0.13
Uruguay	0.03
Serbia	0.16
Montenegro	0.13
Argentina	0.04

-20 -10 0 10 20 30
Score points

Note: Darker bars show statistically significant associations.
1. In this index, zero is the OECD average and one unit is the standard deviation across OECD countries. The higher the value, the better the school climate.
2. In this index, zero is the OECD average and one unit is the standard deviation across OECD countries. The higher the value, the better the teacher-student relations.

LEARNING TRENDS: CHANGES IN STUDENT PERFORMANCE SINCE 2000

The design of PISA does not just allow for a comparison of the relative standing of countries in terms of their learning outcomes; it also enables each country to monitor changes in those outcomes over time. Such changes indicate how successful education systems have been in developing the knowledge and skills of 15-year-olds.

Indeed, some countries have seen impressive improvements in performance over the past decade, sometimes exceeding the equivalent of an average school year's progress for the entire 15-year-old student population. Some of these countries have been catching up from comparatively low performance levels while others have been advancing further from already high levels. All countries seeking to improve their results can draw encouragement – and learn lessons – from those that have succeeded in doing so in a relatively short period of time.

Changes in student performance over time prove that a country's performance in reading is not set in stone. In both absolute and relative terms, educational results can improve, and they cannot be regarded either as part of fixed "cultural" differences between countries or as inevitably linked to each country's state of economic development.

Since both PISA 2000 and PISA 2009 focused on reading, it is possible to track how student performance in reading changed over that period. Among the 26 OECD countries with comparable results in both assessments, Chile, Israel, Poland, Portugal, Korea, Hungary and Germany as well as the partner countries Peru, Albania, Indonesia, Latvia, Liechtenstein and Brazil all improved their reading performance between 2000 and 2009, while performance declined in Ireland, Sweden, the Czech Republic and Australia.

Between 2000 and 2009, the percentage of low performers in Chile dropped by more than 17 percentage points, while the share of top performers in Korea grew by more than 7 percentage points.

In many countries, improvements in results were largely driven by improvements at the bottom end of the performance distribution, signalling progress towards greater equity in learning outcomes. Among OECD countries, variation in student performance fell by 3%. On average across the 26 OECD countries with comparable data for both assessments, 18% of students performed below the baseline reading proficiency Level 2 in 2009, while 19% did so in 2000. Among countries where between 40% and 60% of students performed below Level 2 in 2000, Chile reduced that proportion by the largest amount, and Mexico and the partner country Brazil also show important decreases in their share of low performers. Among countries where the proportion of students performing below Level 2 was smaller than 40% but still above the OECD average of 19%, the partner country Latvia reduced the proportion by 13 percentage points, while Portugal, Poland, Hungary, Germany, Switzerland and the partner country Liechtenstein reduced the share by smaller amounts. In Denmark, the percentage of students below Level 2 fell from an already below-average level.

The share of top performers – those students who attain reading proficiency Level 5 or 6 in reading – increased in Japan, Korea and the partner economy Hong Kong-China such that these countries now have the largest proportions of high-achieving students among the countries participating in the 2009 assessment. Several countries that had above-average proportions of top performers in 2000 saw those proportions decrease in 2009. Notable among them was Ireland, where the proportion of top performers fell from 14% to 7%, which is below the OECD average.

Between 2000 and 2009, Poland, Portugal, Germany, Switzerland and the partner countries Latvia and Liechtenstein raised the performance of their lowest-achieving students while maintaining the performance level among their highest-achieving students. Korea, Israel and the partner country Brazil raised the performance of their highest-achieving students while maintaining the performance level among their lowest-achieving students. Chile and the partner countries Indonesia, Albania and Peru showed improvements in reading performance among students at all proficiency levels.

On average, OECD countries narrowed the gap in scores between their highest- and lowest-performing students between 2000 and 2009; some also improved overall performance. In Chile, Germany, Hungary, Poland, Portugal, and the partner countries Indonesia, Latvia and Liechtenstein, overall performance improved while the variation in performance decreased. In many cases, this was the result of improvements among low-achieving students.

The gender gap in reading performance did not narrow in any country between 2000 and 2009.

The gender gap in reading performance widened in Israel, Korea, Portugal, France and Sweden, and in the partner countries and economies Romania, Hong Kong-China, Indonesia and Brazil between 2000 and 2009. The fact that girls outperform boys in reading is most evident in the proportion of girls and boys who perform below baseline

proficiency Level 2. Across OECD countries, 24% of boys perform below Level 2 compared to only 12% of girls. The proportion of girls performing below this level decreased by two percentage points between 2000 and 2009, while the share of low-achieving boys did not change during the period.

Across the OECD area, the percentage of students with an immigrant background increased by an average of two percentage points between 2000 and 2009. The performance gap between students with and without an immigrant background remained broadly similar over the period. However, some countries noted large reductions in the performance advantage of students without an immigrant background. In Belgium, Switzerland and Germany, the gap narrowed by between 28 and 38 score points due to improvements in reading proficiency among students with an immigrant background. However, the gap is still relatively wide in these countries.

Across OECD countries, overall performance in mathematics remained unchanged between 2003 and 2009, as did performance in science between 2006 and 2009.

In mathematics, students in Mexico, Turkey, Greece, Portugal, Italy, Germany and the partner countries Brazil and Tunisia improved their mathematics scores considerably, while students in the Czech Republic, Ireland, Sweden, France, Belgium, the Netherlands, Denmark, Australia and Iceland saw declines in their performance. On average across the 28 OECD countries with comparable results in the 2003 and 2009 assessments, the share of students below mathematics proficiency Level 2 remained broadly similar over the period, with a minor decrease from 21.6% to 20.8%. Among the OECD countries in which more than half of students performed below mathematics proficiency Level 2 in 2003, Mexico shrunk this proportion by 15 percentage points, from 66% to 51%, by 2009 while Turkey reduced it from 52% to 42% during the same period. Meanwhile, the percentage of top performers in mathematics in those 28 OECD countries decreased slightly, from 14.7% in 2003 to 13.4% in 2009. Portugal showed the largest increase – four percentage points – in top performers in mathematics.

In science, 11 of the 56 countries that participated in both the 2006 and 2009 assessments show improvements in student performance. Turkey, for example, saw a 30 score point increase, nearly half a proficiency level, in just three years. Turkey also reduced the percentage of students below science proficiency Level 2 by almost 17 percentage points, from 47% to 30%. Portugal, Chile, the United States, Norway, Korea and Italy all reduced the share of lowest performers in science by around five percentage points or more, as did the partner countries Qatar, Tunisia, Brazil and Colombia. Performance in science declined considerably in five countries.

On average across OECD countries, the percentage of students who report reading for enjoyment daily dropped by five percentage points.

Enjoyment of reading tends to have deteriorated, especially among boys, signalling the challenge for schools to engage students in reading activities that 15-year-olds find relevant and interesting. On average across OECD countries, the percentage of students who said they read for enjoyment every day fell from 69% in 2000 to 64% in 2009. On the other hand, changes in student-teacher relations and classroom climate have generally been favourable or, at least, they have not deteriorated as many would have expected. Generally, students have become more confident that they can get help from their teachers. Across the 26 OECD countries that participated in both assessments, 74% of students in 2000 agreed or strongly agreed with the statements, "If I need extra help, I will receive it from my teachers" or "Most of my teachers treat me fairly", while in 2009, 79% of students agreed or strongly agreed with those statements. Overall, aspects of classroom discipline have also improved. Thus there is no evidence to justify the notion that students are becoming progressively more disengaged from school.

■ Figure V. ■
COMPARING A SUMMARY OF CHANGES IN READING PERFORMANCE

Mean score in reading 2009 is statistically significantly above the OECD average. Changes in reading and in the share of students at proficiency Level 5 or above are statistically significantly positive. Changes in the share of students below proficiency Level 2 and in the association of socio-economic background with reading is statistically significantly negative.

Mean score in reading 2009 is not statistically significantly different from the OECD average. Changes in reading, in the share of students at proficiency Level 5 or above, in the share of students below proficiency Level 2 and in the association of socio-economic background with reading are not statistically significantly different.

Mean score in reading 2009 is statistically significantly below the OECD average. Changes in reading and in the share of students at proficiency Level 5 or above are statistically significantly negative. Changes in the share of students below proficiency Level 2 and in the association of socio-economic background with reading is statistically significantly positive.

		Change in reading performance between 2000 to 2009					
	Mean score in reading 2009	All students	Boys	Girls	Share of students below proficiency Level 2	Share of students at proficiency Level 5 or above	Association of socio-economic background with reading performance
Peru	370	43	35	50	-14.8	0.4	0.1
Chile	449	40	42	40	-17.6	0.8	-7.6
Albania	385	36	35	39	-13.7	0.1	-9.9
Indonesia	402	31	23	39	-15.2		-6.9
Latvia	484	26	28	23	-12.5	-1.2	-11.0
Israel	474	22	9	35	-6.7	3.3	-8.4
Poland	500	21	14	28	-8.2	1.3	-1.5
Portugal	489	19	12	26	-8.6	0.6	-4.7
Liechtenstein	499	17	16	17	-6.4	-0.4	-13.3
Brazil	412	16	9	21	-6.2	0.8	-0.6
Korea	539	15	4	25	0.0	7.2	8.5
Hungary	494	14	11	17	-5.1	1.0	-4.2
Germany	497	13	10	15	-4.2	-1.2	-7.7
Greece	483	9	3	13	-3.1	0.6	2.0
Hong Kong-China	533	8	0	17	-0.8	2.9	-8.6
Switzerland	501	6	1	10	-3.6	-1.1	-2.3
Mexico	425	3	1	6	-4.0	-0.5	-7.3
Belgium	506	-1	0	-5	-1.2	-0.8	0.7
Bulgaria	429	-1	-8	6	0.7	0.6	-4.5
Italy	486	-1	-5	2	2.1	0.5	3.2
Denmark	495	-2	-5	-1	-2.7	-3.4	-3.2
Norway	503	-2	-5	-1	-2.5	-2.8	0.4
Russian Federation	459	-2	-6	1	-0.1	-0.0	1.4
Japan	520	-2	-6	3	3.5	3.6	c
Romania	424	-3	-18	11	-0.9	-1.5	10.7
United States	500	-5	-2	-6	-0.3	-2.4	-9.2
Iceland	500	-7	-10	-6	2.3	-0.5	5.4
New Zealand	521	-8	-8	-8	0.6	-3.0	4.9
France	496	-9	-15	-4	4.6	1.1	7.0
Thailand	421	-9	-6	-10	5.8	-0.2	-0.7
Canada	524	-10	-12	-10	0.7	-4.0	-6.4
Finland	536	-11	-12	-8	1.2	-4.0	5.8
Spain	481	-12	-14	-10	3.3	-0.9	1.5
Australia	515	-13	-17	-13	1.8	-4.9	-1.4
Czech Republic	478	-13	-17	-6	5.6	-1.9	-11.4
Sweden	497	-19	-24	-15	4.9	-2.2	7.7
Argentina	398	-20	-15	-22	7.7	-0.7	-1.7
Ireland	496	-31	-37	-26	6.2	-7.3	5.8

Note: Countries are ranked in descending order of the change in reading performance between 2000 and 2009 for all students.
Source: OECD, *PISA 2009 Database*, Tables V.2.1, V.2.2, V.2.4 and V.4.3

StatLink http://dx.doi.org/10.1787/888932359948

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where governments work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The European Commission takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

OECD PUBLISHING, 2, rue André-Pascal, 75775 PARIS CEDEX 16

PISA IN FOCUS 25

education policy education policy education policy education policy education policy education policy education policy

Are countries moving towards more equitable education systems?

- PISA results show that no country or economy has reached the goal of creating a completely equitable education system, but some are much closer than others.
- Some countries and economies have shown that improvements in equity can be achieved at the same time as improvements in overall performance, and in a relatively short time.

PISA has consistently shown a link between disadvantage and poor performance in school…

Ideally, school systems provide high-quality educational opportunities for all students, irrespective of the students' backgrounds. Students from socio-economically advantaged families and those from disadvantaged families should be equally likely to succeed in school. That is the ideal, anyway. In most countries, the reality looks a lot different. PISA results have consistently shown that socio-economic disadvantage is linked to poor performance in school. In fact, on average across OECD countries, disadvantaged students are twice as likely to be among the poorest performers in reading compared to advantaged students. On average, a socio-economically advantaged student scores 88 points higher on the PISA reading test than a socio-economically disadvantaged student, a difference that is equivalent to more than two years of schooling.

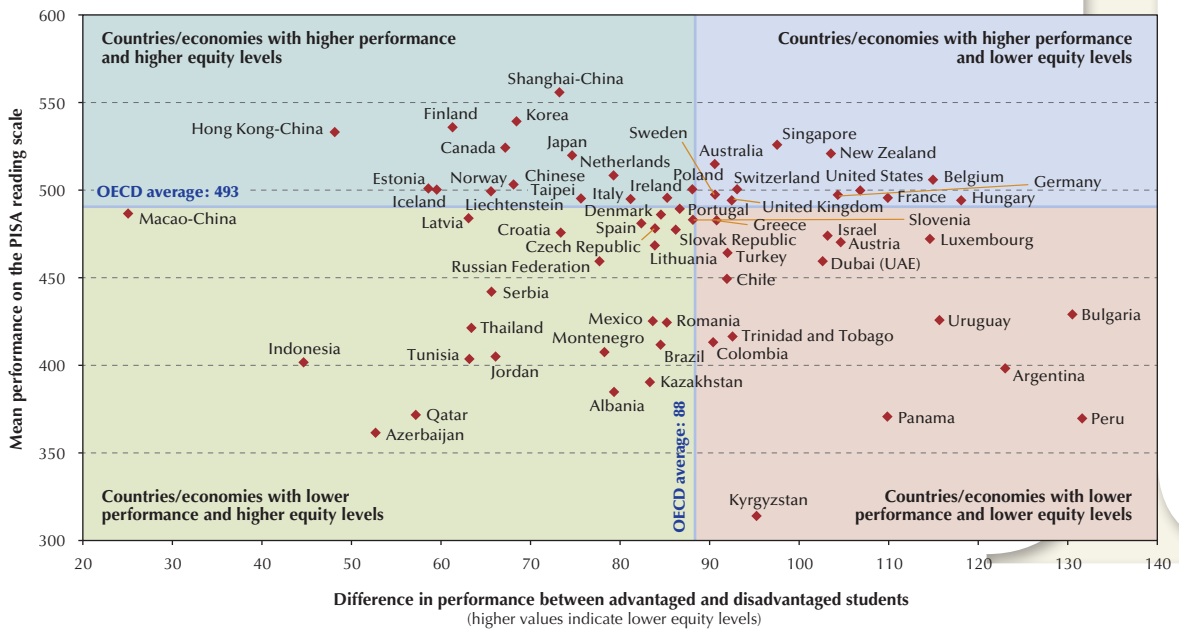
Yet the fact that countries and economies vary in the degree to which learning outcomes are linked to socio-economic background demonstrates that social background is not destiny, and that policy and practice can make a difference. Moreover, countries can pursue equitable learning outcomes while also moving towards high student performance. In the PISA 2009 survey, many of the countries and economies with the greatest equity in student outcomes are also top performers. Students in Canada, Estonia, Finland, Hong Kong-China, Iceland, Korea, Liechtenstein and Norway score above the OECD average in reading, and the difference in performance between advantaged and disadvantaged students is less than 70 score points. Other countries and economies also achieve equitable learning outcomes, but their students do not perform as well.

Azerbaijan, Indonesia, Jordan, Macao-China, Qatar, Serbia, Thailand and Tunisia are as equitable as the preceding group of countries, but their students score below the OECD average in reading. The difference in reading performance between advantaged and disadvantaged students is highest – more than 100 score points – in Argentina, Austria, Belgium, Bulgaria, Dubai (UAE), France, Germany, Hungary, Israel, Luxembourg, New Zealand, Panama, Peru, the United States and Uruguay.

…but that doesn't mean that the link is unbreakable.

The fact that school systems vary in the degree to which they allow socio-economic differences to become differences in performance is an indication that any obstacles posed by disadvantaged social background can be overcome. While some of those differences may be attributable to culture, the fact that the relationship has changed significantly in some countries suggests that policy and practice can make a difference.

Country/economy's equity and performance in PISA 2009



Sources: OECD (2010), *PISA 2009 Results: What Students Know and Can Do*, Volume I, OECD Publishing, Table I.2.3; and OECD (2010), *PISA 2009 Results: Overcoming Social Background*, Volume II, OECD Publishing, Table II.3.1.

Many countries and economies have made notable progress in narrowing the performance gaps between advantaged and disadvantaged students while simultaneously improving overall performance. This shows that education systems can reduce the extent to which differences in socio-economic background relate to student performance while promoting learning for all students. Comparing results from PISA 2000 and PISA 2009 reveals that, in Albania, Chile, Germany and Latvia the relationship between students' socio-economic status and their reading performance weakened and students' overall reading performance improved. In Germany, for example, the performance gap between advantaged and disadvantaged students narrowed by more than 25 score points and the average reading performance improved by 13 points. In Chile, average performance in reading improved by 40 score points and the performance gap between advantaged and disadvantaged students narrowed by more than 15 score points.

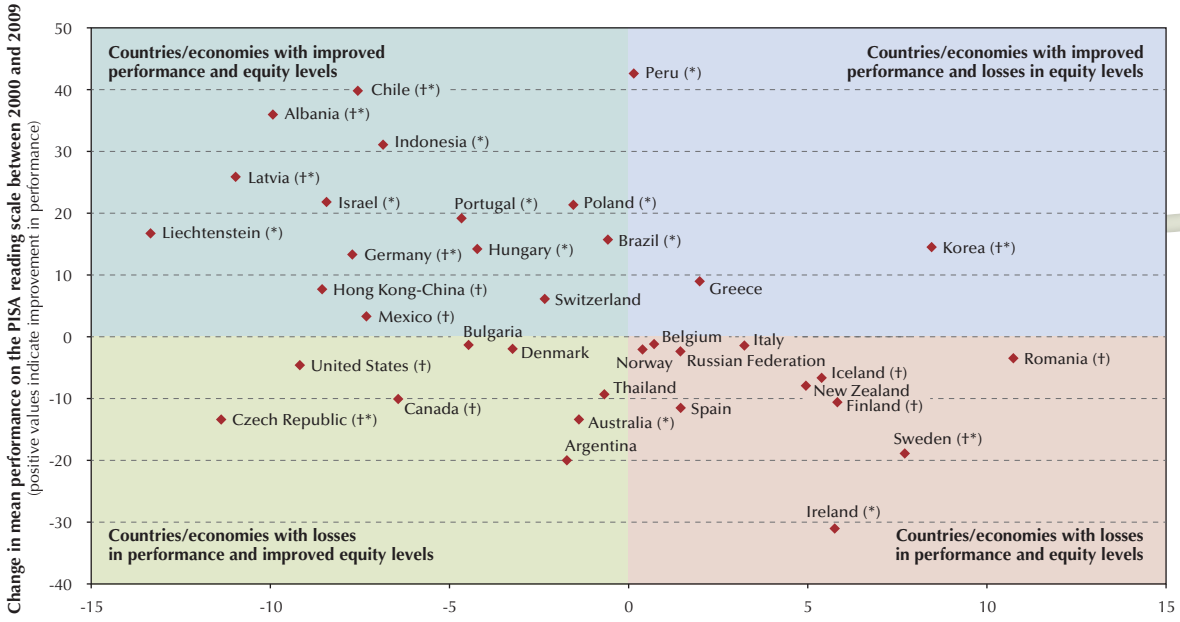
Other countries and economies, namely Canada, the Czech Republic, Hong Kong-China, Mexico and the United States, also improved their equity levels without a concurrent improvement in overall performance.

The right policies can help to weaken the link.

In Albania, Chile, Germany and Latvia, the association between socio-economic status and reading performance weakened between 2000 and 2009 while overall performance improved. Equity and performance improved simultaneously in these countries either because students from all backgrounds improved, with largest improvements seen among disadvantaged students (e.g. in Albania and Chile) or because although the performance of advantaged students did not change significantly, that of disadvantaged students did (for example, in Germany and Latvia).

Education policies that can foster improvements in equity and performance can include giving more and better support to disadvantaged students, who start school with deficits in their education; ensuring that all schools provide high-quality instruction; and offering additional educational opportunities to disadvantaged students, as their parents might not be able provide them. More support could be given to schools with large disadvantaged student populations or to disadvantaged students within schools. Broader social policies that help to ensure that the life experiences of advantaged and disadvantaged students are not that different, either at home – crucially, before students enter formal education – or in school can also promote both equity and high performance.

Change in equity and performance levels between 2000 and 2009



Note: The change in reading performance associated with a one-unit increase in a student's *PISA index of economic, social and cultural status* is usually referred to as the slope of the socio-economic gradient and is the slope of a regression of socio-economic status on student reading performance. The difference in this change between 2009 and 2000 is presented in the horizontal axis.

† Change in equity is statistically significant * Change in performance is statistically significant

Source: OECD (2010), *PISA 2009 Results: Learning Trends*, Volume V, OECD Publishing, Tables V.2.1 and V.4.3.

All these policies that promote equity and overall performance in a school system are based on the premise that students should compete on a level playing field, and that if socio-economic disadvantage deprives students of having the same opportunities other students enjoy, these opportunities need to be provided by the school system.

Socio-economic status in PISA is measured through the *PISA index of economic, social and cultural status* (ESCS). This index combines students' responses on their parents' occupations and educational attainment, and their reports on the cultural possessions and educational resources available in the students' home. **Advantaged students** are those in the top quarter of the ESCS distribution within their country; **disadvantaged students** are those in the bottom quarter. **Poor-perfoming students** are those in the bottom quarter of a country's performance distribution.

The bottom line: Although no country or economy has achieved the goal of developing a completely equitable education system, with the right policies, some have been able to weaken the link between socio-economic background and performance. This proves that it should not be considered inevitable that students from disadvantaged backgrounds perform poorly in school. All students should have the same opportunities to succeed in school, regardless of their backgrounds.

For more information

Contact Guillermo Montt *(Guillermo.MONTT@oecd.org)*

See OECD (2010), *PISA 2009 Results: Overcoming Social Background: Equity in Learning Opportunities and Outcomes*, Volume II, PISA, OECD Publishing,
OECD (2010), *PISA 2009 Results: Learning Trends: Changes in Student Performance Since 2000*, Volume V, PISA, OECD Publishing,
Supporting data.

Visit
www.pisa.oecd.org
www.oecd.org/pisa/infocus

Coming next month
Grade expectations

Chapter 6

School Climate

Students with the highest mathematics achievement typically attend schools that emphasize academic success, as indicated by rigorous curricular goals, effective teachers, students that desire to do well, and parental support. In contrast, schools with discipline and safety problems are not conducive to high achievement. Students that attended schools with disorderly environments and reported more frequent bullying had much lower achievement than their counterparts in safe and orderly schools.

The school's educational values are reflected by the teachers, school leadership, the students themselves, and their parents. A school with a positive atmosphere toward high achievement and a rigorous academic program can overcome resource shortages and encourage students toward excellent performance. By contrast, a school with more disciplinary problems is not conducive to higher student achievement. When students are fearful and worried about their safety, for example, it is difficult to focus on academics. Chapter 6 presents the TIMSS 2011 results about positive and negative aspects of the atmosphere in schools around the world.

Schools Emphasize Academic Success

Studies of academic optimism show that a positive school atmosphere emphasizing academic achievement can even overcome socioeconomic disadvantages (McGuigan & Hoy, 2006). There are several dimensions of academic optimism, including a school communicating its academic emphasis through clear and rigorous academic goals. However, because individuals are the actors within schools, the effect on achievement is greatest when there is a collective influence. This includes a school administration and teachers that support and trust in students' achievement. In addition to making it clear that academic success is important, principals and teachers need to emphasize it can be achieved. Parents' support for their children's learning also contributes to a schools' collective efficacy or belief that the school's academic goals can be implemented.

School Emphasis on Academic Success

The TIMSS 2011 School Emphasis on Academic Success scale characterizes five aspects of academic optimism:

- Teachers' understanding of the school's curricular goals;
- Teachers' degree of success in implementing the school's curriculum;
- Teachers' expectations for student achievement;
- Parental support for student achievement; and
- Students' desire to do well in school.

This set of questions was given to both students' principals and teachers, with the respective responses used to create scales.

Exhibit 6.1 shows the principals' reports on the School Emphasis on Academic Success scale for the TIMSS 2011 fourth grade assessment. As might be anticipated, principals had very positive attitudes about the emphasis on

academics in their schools, so the three regions of the scale have been described as **Very High**, **High**, and **Medium**. Students were scored according to their principals' characterization of their school in terms of the five aspects. Students in schools with **Very High Emphasis** on academic success had principals characterizing three of the five aspects as "very high" and the other two as "high," on average. Students in **Medium Emphasis** schools had principals characterizing three of the five aspects as "medium" and the other two as "high," on average. All other students attended schools with a **High Emphasis** on academic success.

On average, across the fourth grade countries, 8 percent of the students attended schools where the principal reported a **Very High Emphasis** on academic success, 58 percent a school with a **High Emphasis**, and 34 percent a school with a **Medium Emphasis**. Although the results were not entirely consistent from country to country, internationally at the fourth grade, on average, there was a direct correspondence between average mathematics achievement and principals' reports, with higher emphasis on academic success related to higher average mathematics achievement. The results were similar for the sixth grade and benchmarking participants.

Exhibit 6.2 shows the principals' reports on the School Emphasis on Academic Success scale for the TIMSS 2011 eighth grade assessment. Although similar to the fourth grade results, principals of the eighth grade schools reported slightly less emphasis on academic success, with 7 percent of the students attending a school where the principal reported a **Very High Emphasis** on academic success, 53 percent a school with a **High Emphasis**, and 41 percent a school with a **Medium Emphasis** (compared to 8%, 58%, and 34%, respectively, at the fourth grade). There was also a somewhat greater difference in average mathematics achievement (46 points) between students attending **Very High Emphasis** schools and students attending **Medium Emphasis** schools (495 vs. 449).

Exhibits 6.3 and 6.4 show the teachers' reports on the School Emphasis on Academic Success scale for the fourth and eighth grade assessments, respectively. The teachers' reports were remarkably similar to those of the principals for both assessments, and with each reported decrease in academic emphasis, the students had progressively lower average mathematics achievement. Similar to their principals, the eighth grade mathematics teachers reported a little less emphasis on academic success than the fourth grade teachers, but the achievement gap between students in very high and medium emphasis schools was greater at the eighth grade (54 points) than at the fourth grade (26 points).

Exhibit 6.1: School Emphasis on Academic Success - Principal Reports

TIMSS 2011 Mathematics 4th Grade

Reported by Principals

Students were scored according to their principals' responses characterizing five aspects on the *School Emphasis on Academic Success* scale. Students in schools where their principals reported a **Very High Emphasis** on academic success had a score on the scale of at least 13.1, which corresponds to their principals characterizing three of the five aspects as "very high" and the other two as "high," on average. Students in schools with a **Medium Emphasis** on academic success had a score no higher than 8.9, which corresponds to their principals characterizing three of the five aspects as "medium" and the other two as "high," on average. All other students attended schools with a **High Emphasis** on academic success.

Country	Very High Emphasis		High Emphasis		Medium Emphasis		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Northern Ireland	33 (4.2)	577 (4.9)	60 (4.3)	558 (4.1)	7 (2.5)	540 (13.6)	12.0 (0.19)
Qatar	31 (2.9)	435 (10.3)	54 (3.2)	411 (5.3)	15 (2.4)	374 (8.2)	11.6 (0.14)
Ireland	28 (4.0)	543 (4.8)	67 (3.9)	523 (3.7)	4 (1.7)	508 (9.6)	11.9 (0.17)
United States	22 (2.5)	561 (4.4)	60 (2.7)	543 (2.9)	18 (2.1)	519 (4.7)	11.2 (0.13)
New Zealand	22 (3.0)	506 (4.7)	67 (3.3)	487 (3.4)	11 (2.1)	448 (11.0)	11.5 (0.14)
Korea, Rep. of	22 (3.5)	612 (4.4)	58 (4.3)	606 (2.7)	20 (3.4)	594 (3.3)	11.1 (0.19)
United Arab Emirates	21 (1.6)	463 (5.7)	61 (2.0)	429 (3.4)	18 (1.7)	401 (6.2)	11.2 (0.09)
Chinese Taipei	17 (3.0)	592 (5.7)	71 (3.7)	592 (2.4)	12 (2.5)	584 (4.6)	11.3 (0.15)
Australia	16 (3.0)	544 (7.6)	64 (3.8)	519 (3.7)	21 (3.0)	488 (5.6)	10.9 (0.14)
Malta	13 (0.1)	507 (3.8)	69 (0.1)	502 (1.5)	18 (0.1)	466 (3.1)	11.1 (0.00)
Bahrain	11 (2.5)	483 (9.7)	68 (3.7)	433 (4.6)	21 (2.8)	418 (7.6)	10.6 (0.16)
England	10 (2.9)	554 (6.0)	72 (4.7)	546 (4.9)	17 (3.8)	517 (9.9)	10.8 (0.18)
Iran, Islamic Rep. of	9 (2.0)	442 (14.6)	70 (3.4)	436 (4.6)	21 (2.7)	408 (6.1)	10.6 (0.12)
Saudi Arabia	9 (2.7)	453 (23.3)	59 (4.1)	412 (4.4)	32 (3.4)	394 (11.2)	10.2 (0.18)
Croatia	9 (2.5)	499 (6.4)	70 (3.8)	492 (2.2)	21 (3.4)	479 (5.2)	10.7 (0.14)
Sweden	9 (2.7)	522 (10.0)	59 (4.8)	505 (2.8)	32 (4.9)	497 (3.8)	10.3 (0.17)
Kuwait	9 (2.0)	349 (12.4)	65 (3.8)	348 (4.5)	27 (3.8)	327 (6.9)	10.4 (0.17)
Oman	9 (1.8)	376 (7.9)	73 (3.0)	383 (3.9)	18 (2.2)	362 (6.5)	10.6 (0.10)
Austria	8 (2.1)	511 (8.5)	75 (4.4)	511 (2.4)	17 (3.9)	493 (7.4)	10.4 (0.14)
Singapore	8 (0.0)	627 (12.2)	62 (0.0)	610 (4.3)	31 (0.0)	591 (6.3)	10.2 (0.00)
Finland	6 (1.9)	561 (2.1)	71 (4.2)	548 (2.5)	24 (4.2)	536 (5.9)	10.4 (0.16)
Lithuania	6 (2.0)	547 (13.2)	65 (3.6)	541 (3.0)	29 (3.4)	517 (5.6)	10.0 (0.13)
Kazakhstan	5 (1.9)	495 (26.2)	65 (4.4)	506 (6.3)	30 (4.1)	492 (8.5)	10.2 (0.12)
Chile	5 (1.9)	516 (17.1)	30 (3.3)	481 (5.8)	65 (3.8)	452 (4.2)	8.8 (0.19)
Denmark r	5 (1.3)	543 (4.5)	65 (3.6)	539 (3.6)	30 (3.3)	540 (3.9)	10.1 (0.11)
Portugal	4 (2.0)	543 (8.6)	64 (5.0)	537 (4.9)	31 (4.5)	522 (6.5)	10.0 (0.13)
Azerbaijan	4 (1.7)	478 (15.9)	44 (3.8)	467 (10.9)	53 (3.8)	456 (6.3)	9.2 (0.15)
Romania	4 (1.6)	547 (13.3)	55 (4.1)	495 (6.8)	41 (4.1)	459 (9.7)	9.5 (0.15)
Poland	3 (1.6)	525 (20.1)	70 (3.5)	484 (2.6)	26 (3.7)	470 (4.1)	9.8 (0.15)
Morocco	3 (1.0)	408 (16.9)	25 (3.1)	357 (10.5)	72 (3.0)	325 (5.5)	8.0 (0.14)
Yemen	2 (1.2)	~ ~	35 (4.2)	260 (8.5)	62 (4.5)	242 (8.5)	8.7 (0.18)
Tunisia	2 (1.3)	~ ~	37 (4.3)	372 (5.6)	60 (4.2)	350 (5.0)	8.8 (0.16)
Spain	2 (1.3)	~ ~	58 (4.1)	491 (3.1)	40 (3.9)	471 (5.2)	9.6 (0.12)
Turkey	2 (1.0)	~ ~	33 (3.3)	493 (8.4)	65 (3.1)	455 (5.7)	8.6 (0.14)
Thailand	2 (1.1)	~ ~	52 (4.8)	463 (5.7)	46 (4.8)	449 (7.9)	9.5 (0.14)
Serbia	2 (1.2)	~ ~	52 (4.0)	521 (4.0)	46 (4.0)	507 (4.5)	9.4 (0.13)
Slovenia	2 (0.8)	~ ~	63 (2.9)	514 (2.8)	35 (3.1)	511 (3.1)	9.6 (0.10)
Russian Federation	2 (0.9)	~ ~	50 (4.4)	547 (5.1)	48 (4.3)	538 (4.5)	9.2 (0.11)
Hong Kong SAR	1 (0.9)	~ ~	60 (4.5)	602 (3.5)	38 (4.6)	601 (6.6)	9.7 (0.16)
Japan	1 (1.0)	~ ~	48 (4.5)	592 (2.7)	51 (4.5)	579 (2.7)	9.0 (0.16)
Italy	1 (0.8)	~ ~	52 (3.7)	507 (3.6)	46 (3.7)	508 (3.9)	9.4 (0.10)
Hungary	1 (0.9)	~ ~	49 (3.9)	537 (4.1)	50 (3.9)	495 (6.0)	9.0 (0.13)
Czech Republic	1 (0.9)	~ ~	45 (3.9)	513 (4.3)	54 (4.0)	509 (2.9)	8.9 (0.13)
Armenia	1 (0.8)	~ ~	56 (4.3)	457 (4.3)	43 (4.3)	446 (6.0)	9.6 (0.12)
Norway	1 (0.1)	~ ~	64 (4.7)	500 (3.5)	34 (4.7)	484 (3.9)	9.8 (0.13)
Germany	1 (0.8)	~ ~	66 (3.4)	537 (2.3)	33 (3.3)	512 (4.1)	9.9 (0.11)
Netherlands r	1 (1.0)	~ ~	50 (6.0)	544 (3.0)	49 (6.0)	538 (3.2)	9.3 (0.18)
Georgia	1 (0.9)	~ ~	46 (3.9)	457 (7.1)	53 (3.6)	443 (4.9)	9.1 (0.11)
Slovak Republic	1 (0.7)	~ ~	41 (3.4)	520 (4.7)	58 (3.4)	496 (5.3)	8.8 (0.10)
Belgium (Flemish)	1 (0.0)	~ ~	70 (3.7)	553 (2.2)	30 (3.7)	543 (3.8)	9.9 (0.11)
International Avg.	8 (0.3)	511 (2.2)	58 (0.5)	496 (0.7)	34 (0.5)	477 (0.9)	

Centerpoint of scale set at 10.

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

A tilde (~) indicates insufficient data to report achievement.

An "r" indicates data are available for at least 70% but less than 85% of the students.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 6.1: School Emphasis on Academic Success - Principal Reports (Continued)

TIMSS 2011 Mathematics 4th Grade

Country	Very High Emphasis		High Emphasis		Medium Emphasis		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Sixth Grade Participants							
Honduras	10 (2.5)	385 (14.7)	61 (4.5)	395 (8.9)	29 (4.1)	403 (6.3)	10.2 (0.17)
Botswana	5 (1.8)	505 (24.5)	29 (3.8)	438 (8.1)	66 (4.1)	404 (3.7)	8.8 (0.18)
Yemen	2 (1.2)	~ ~	33 (4.2)	369 (9.0)	65 (4.2)	337 (7.1)	8.7 (0.17)
Benchmarking Participants							
Dubai, UAE	35 (0.3)	495 (3.1)	49 (0.5)	467 (2.3)	16 (0.4)	397 (5.2)	11.8 (0.01)
Alberta, Canada	31 (4.4)	515 (3.5)	58 (4.9)	507 (3.0)	12 (2.8)	490 (9.8)	11.8 (0.17)
Florida, US r	27 (5.0)	580 (6.7)	58 (5.3)	532 (4.4)	15 (4.4)	529 (7.2)	11.5 (0.27)
Abu Dhabi, UAE	17 (3.4)	435 (12.5)	68 (3.8)	413 (5.7)	15 (3.0)	393 (13.6)	11.0 (0.17)
Ontario, Canada	12 (2.9)	534 (6.5)	65 (4.3)	522 (3.8)	23 (4.1)	499 (4.2)	10.6 (0.20)
North Carolina, US	7 (4.2)	599 (6.9)	76 (7.1)	558 (5.3)	17 (5.6)	530 (9.0)	10.8 (0.27)
Quebec, Canada	5 (1.6)	563 (11.1)	75 (3.6)	535 (2.4)	21 (3.4)	519 (5.5)	10.4 (0.12)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

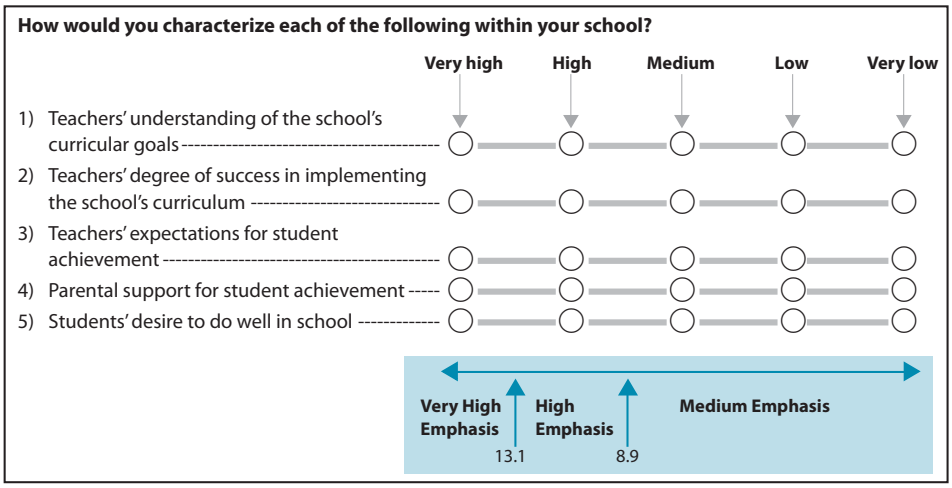


Exhibit 6.2: School Emphasis on Academic Success - Principal Reports

TIMSS 2011 Mathematics 8th Grade

Reported by Principals

Students were scored according to their principals' responses characterizing five aspects on the *School Emphasis on Academic Success* scale. Students in schools where their principals reported a **Very High Emphasis** on academic success had a score on the scale of at least 13.3, which corresponds to their principals characterizing three of the five aspects as "very high" and the other two as "high," on average. Students in schools with a **Medium Emphasis** on academic success had a score no higher than 9.2, which corresponds to their principals characterizing three of the five aspects as "medium" and the other two as "high," on average. All other students attended schools with a **High Emphasis** on academic success.

Country	Very High Emphasis		High Emphasis		Medium Emphasis		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Qatar	27 (0.3)	453 (4.6)	57 (0.3)	395 (5.0)	16 (0.1)	378 (5.7)	11.5 (0.02)
England	26 (3.5)	525 (12.3)	56 (4.7)	509 (8.2)	19 (3.4)	477 (14.7)	11.6 (0.18)
Australia	20 (2.7)	558 (15.8)	48 (3.8)	509 (5.9)	32 (3.1)	476 (7.4)	10.8 (0.16)
New Zealand	19 (3.8)	524 (9.2)	61 (4.9)	484 (7.1)	20 (3.3)	467 (6.7)	11.1 (0.15)
United Arab Emirates	17 (1.6)	497 (6.5)	63 (2.0)	453 (2.7)	20 (1.8)	426 (4.7)	11.1 (0.09)
Korea, Rep. of	16 (3.2)	637 (7.3)	56 (4.3)	613 (3.8)	28 (3.6)	597 (3.8)	10.7 (0.19)
United States	15 (2.0)	532 (8.0)	61 (2.7)	515 (3.7)	24 (2.1)	486 (5.4)	10.9 (0.09)
Chinese Taipei	12 (2.8)	657 (15.1)	81 (3.3)	605 (3.8)	7 (1.7)	579 (7.7)	11.4 (0.11)
Iran, Islamic Rep. of	12 (2.5)	462 (15.1)	62 (3.6)	418 (6.4)	27 (2.6)	387 (5.7)	10.7 (0.13)
Singapore	11 (0.0)	651 (11.2)	60 (0.0)	614 (4.2)	29 (0.0)	586 (7.8)	10.8 (0.00)
Israel	9 (2.4)	515 (18.7)	75 (3.6)	529 (5.1)	17 (3.0)	471 (13.6)	11.0 (0.13)
Indonesia	8 (2.2)	417 (18.8)	60 (4.8)	387 (6.7)	32 (4.4)	377 (5.9)	10.4 (0.16)
Oman	7 (1.4)	407 (10.7)	67 (2.8)	373 (3.8)	25 (2.6)	332 (4.7)	10.5 (0.10)
Saudi Arabia	7 (2.3)	442 (17.8)	48 (4.5)	396 (6.4)	45 (4.1)	383 (7.6)	9.9 (0.16)
Ghana	6 (1.7)	374 (8.7)	53 (4.6)	337 (7.0)	41 (4.3)	315 (5.7)	10.0 (0.13)
Malaysia	6 (1.9)	467 (25.6)	65 (3.1)	453 (6.7)	29 (2.7)	405 (8.4)	10.4 (0.12)
Kazakhstan	5 (1.8)	513 (22.8)	60 (4.2)	483 (6.3)	35 (4.1)	489 (6.8)	10.2 (0.13)
Jordan	5 (1.6)	439 (10.9)	56 (3.5)	415 (5.1)	39 (3.6)	389 (5.6)	10.0 (0.14)
Chile	5 (1.8)	467 (11.9)	27 (3.3)	451 (5.4)	68 (3.3)	401 (3.7)	8.7 (0.17)
Sweden r	5 (2.1)	488 (5.5)	62 (4.6)	491 (3.0)	34 (4.4)	475 (3.6)	10.3 (0.15)
Romania	4 (1.6)	531 (18.0)	55 (4.6)	473 (6.3)	41 (4.6)	432 (6.9)	9.8 (0.16)
Finland	4 (1.8)	530 (8.2)	71 (4.1)	517 (2.8)	25 (3.9)	501 (4.3)	10.4 (0.13)
Syrian Arab Republic	4 (1.7)	350 (20.6)	39 (3.7)	394 (6.1)	57 (3.9)	373 (5.8)	9.3 (0.19)
Bahrain	4 (0.1)	522 (8.0)	57 (0.3)	425 (2.4)	40 (0.3)	375 (3.3)	10.3 (0.01)
Macedonia, Rep. of	3 (1.1)	439 (18.9)	64 (3.6)	440 (6.1)	33 (3.7)	403 (11.7)	10.2 (0.15)
Morocco	3 (0.9)	450 (28.5)	26 (2.7)	393 (5.5)	71 (2.7)	360 (2.5)	8.7 (0.12)
Hong Kong SAR	3 (1.6)	662 (40.2)	51 (4.1)	608 (5.9)	47 (4.3)	554 (7.7)	9.8 (0.15)
Palestinian Nat'l Auth.	3 (1.4)	404 (10.8)	52 (4.1)	408 (5.0)	46 (4.2)	400 (6.5)	9.7 (0.14)
Thailand	3 (1.4)	445 (17.9)	47 (3.9)	436 (8.2)	50 (4.1)	418 (6.2)	9.7 (0.15)
Lebanon	2 (1.2)	~ ~	59 (4.1)	467 (5.4)	39 (3.9)	424 (5.0)	9.8 (0.16)
Slovenia	2 (1.1)	~ ~	62 (3.4)	508 (2.8)	35 (3.5)	499 (4.4)	9.8 (0.12)
Turkey	2 (0.9)	~ ~	33 (3.1)	495 (8.8)	65 (3.0)	429 (4.5)	8.9 (0.11)
Norway	2 (1.1)	~ ~	63 (4.6)	479 (3.3)	35 (4.5)	466 (3.3)	10.1 (0.13)
Lithuania	2 (1.1)	~ ~	56 (3.9)	512 (3.8)	42 (3.9)	489 (5.0)	9.7 (0.12)
Japan	2 (1.1)	~ ~	52 (4.4)	580 (4.0)	47 (4.3)	556 (3.8)	9.7 (0.14)
Hungary	1 (1.0)	~ ~	48 (4.2)	524 (4.1)	51 (4.1)	486 (6.0)	9.3 (0.15)
Tunisia	1 (0.4)	~ ~	18 (3.1)	443 (10.4)	82 (3.0)	421 (3.0)	8.0 (0.14)
Italy	0 (0.0)	~ ~	47 (3.6)	502 (3.6)	53 (3.6)	495 (4.0)	9.4 (0.13)
Armenia	0 (0.0)	~ ~	41 (4.2)	479 (5.3)	59 (4.2)	459 (4.5)	9.3 (0.10)
Georgia	0 (0.0)	~ ~	30 (3.3)	450 (10.7)	70 (3.3)	424 (4.5)	8.7 (0.11)
Russian Federation	0 (0.0)	~ ~	28 (3.0)	563 (7.8)	72 (3.0)	530 (4.0)	8.8 (0.08)
Ukraine	0 (0.0)	~ ~	31 (3.5)	505 (6.2)	69 (3.5)	468 (4.5)	9.0 (0.10)
International Avg.	7 (0.3)	495 (3.1)	53 (0.6)	477 (0.9)	41 (0.5)	449 (1.0)	

Centerpoint of scale set at 10.

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

A tilde (~) indicates insufficient data to report achievement.

An "r" indicates data are available for at least 70% but less than 85% of the students.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 6.2: School Emphasis on Academic Success - Principal Reports (Continued)

TIMSS 2011 Mathematics 8th Grade

Country	Very High Emphasis		High Emphasis		Medium Emphasis		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Ninth Grade Participants							
South Africa	4 (1.0)	470 (45.6)	31 (3.1)	371 (7.6)	66 (3.0)	335 (3.2)	8.9 (0.12)
Honduras	2 (1.0)	~ ~	52 (4.6)	340 (6.3)	47 (4.7)	333 (5.8)	9.4 (0.18)
Botswana	1 (0.8)	~ ~	20 (3.2)	416 (7.0)	79 (3.2)	389 (2.7)	8.2 (0.13)
Benchmarking Participants							
Dubai, UAE	28 (0.4)	519 (5.2)	59 (0.4)	470 (2.6)	13 (0.3)	418 (5.4)	11.8 (0.02)
Massachusetts, US	27 (6.1)	576 (12.9)	51 (6.7)	556 (9.3)	22 (5.8)	546 (14.3)	11.4 (0.34)
Connecticut, US r	22 (5.6)	542 (17.4)	54 (6.9)	531 (10.5)	24 (5.7)	471 (13.3)	11.2 (0.29)
Alberta, Canada	19 (3.1)	521 (7.5)	68 (4.0)	504 (2.9)	13 (2.7)	493 (5.5)	11.5 (0.15)
Colorado, US	18 (4.6)	546 (10.0)	52 (7.2)	520 (8.5)	30 (5.7)	495 (15.7)	10.9 (0.26)
California, US r	14 (3.0)	542 (11.7)	63 (5.9)	492 (7.5)	23 (4.9)	462 (10.1)	10.8 (0.21)
Indiana, US r	13 (5.6)	537 (16.5)	68 (7.0)	522 (7.8)	18 (5.5)	524 (10.6)	11.1 (0.32)
Abu Dhabi, UAE	13 (3.4)	495 (17.5)	64 (4.4)	449 (4.7)	22 (3.9)	422 (6.0)	10.9 (0.18)
Ontario, Canada	13 (3.1)	520 (6.5)	62 (4.4)	517 (2.9)	25 (3.6)	494 (6.9)	10.7 (0.17)
Minnesota, US	12 (5.1)	544 (33.9)	68 (6.0)	549 (5.8)	20 (5.4)	537 (10.7)	11.1 (0.24)
Alabama, US r	11 (2.8)	512 (27.1)	56 (9.0)	469 (8.2)	33 (9.0)	447 (10.0)	10.6 (0.27)
Florida, US	10 (4.9)	502 (39.8)	66 (8.2)	518 (9.5)	24 (6.9)	502 (10.6)	10.6 (0.31)
North Carolina, US	9 (4.2)	549 (12.7)	46 (7.4)	549 (10.4)	45 (6.6)	520 (11.4)	10.1 (0.25)
Quebec, Canada	7 (1.8)	576 (8.5)	62 (4.1)	537 (3.1)	31 (3.7)	512 (5.0)	10.4 (0.13)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

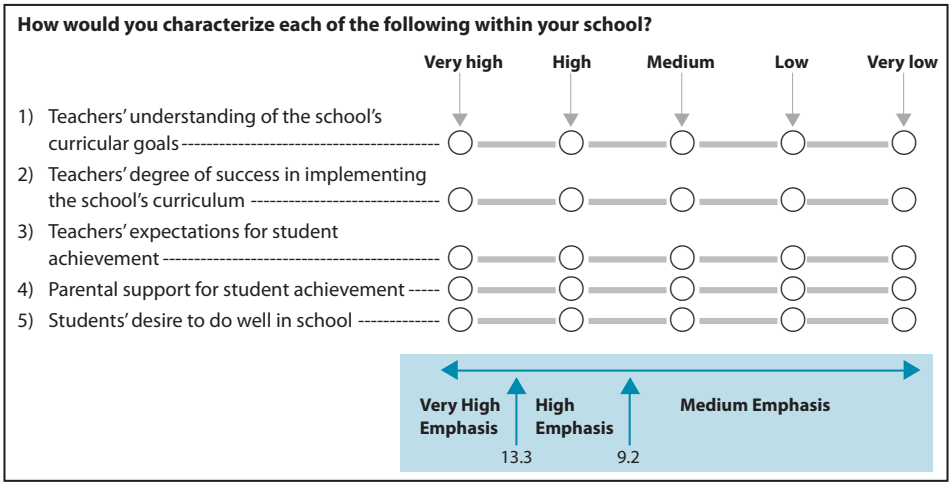


Exhibit 6.3: School Emphasis on Academic Success - Teacher Reports

TIMSS 2011 4th Grade Mathematics

Reported by Teachers

Students were scored according to their teachers' responses characterizing five aspects on the *School Emphasis on Academic Success* scale. Students in schools where their teachers reported a **Very High Emphasis** on academic success had a score on the scale of at least 13.1, which corresponds to their teachers characterizing three of the five aspects as "very high" and the other two as "high," on average. Students in schools with a **Medium Emphasis** on academic success had a score no higher than 8.8, which corresponds to their teachers characterizing three of the five aspects as "medium" and the other two as "high," on average. All other students attended schools with a **High Emphasis** on academic success.

Country	Very High Emphasis		High Emphasis		Medium Emphasis		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Northern Ireland r	31 (4.3)	573 (6.9)	65 (4.4)	559 (4.6)	5 (1.6)	550 (10.5)	11.9 (0.17)
Ireland	22 (3.4)	546 (5.1)	70 (3.5)	526 (3.6)	8 (1.8)	494 (7.6)	11.5 (0.15)
Croatia	21 (3.0)	490 (3.6)	69 (3.6)	489 (2.3)	10 (2.2)	496 (6.7)	11.4 (0.12)
United States	18 (2.1)	560 (4.6)	66 (2.5)	545 (2.2)	16 (1.8)	515 (5.1)	11.0 (0.10)
Korea, Rep. of	17 (3.4)	618 (5.8)	65 (3.8)	605 (2.4)	18 (3.5)	593 (3.3)	10.9 (0.19)
Australia r	16 (3.0)	550 (12.3)	63 (4.2)	519 (4.4)	20 (3.0)	495 (5.8)	10.8 (0.16)
England	16 (3.0)	563 (7.5)	67 (4.5)	546 (4.7)	17 (3.4)	522 (9.0)	11.1 (0.16)
United Arab Emirates	15 (1.8)	464 (7.6)	66 (2.5)	436 (2.3)	19 (2.1)	409 (8.2)	10.9 (0.10)
New Zealand	14 (2.0)	509 (8.0)	69 (2.8)	487 (3.0)	17 (2.5)	465 (5.5)	10.9 (0.11)
Kazakhstan	12 (2.3)	489 (10.9)	68 (3.4)	504 (5.8)	20 (2.9)	502 (9.6)	10.7 (0.13)
Qatar	11 (2.6)	439 (21.8)	63 (4.8)	414 (5.5)	25 (4.5)	397 (10.1)	10.6 (0.19)
Malta	11 (0.1)	515 (2.7)	70 (0.1)	498 (1.7)	19 (0.1)	477 (2.7)	10.5 (0.00)
Saudi Arabia	11 (2.4)	440 (11.1)	63 (3.8)	419 (6.9)	26 (3.3)	377 (9.7)	10.4 (0.15)
Bahrain	9 (3.0)	499 (14.6)	63 (5.2)	435 (5.5)	28 (4.1)	418 (3.9)	10.3 (0.17)
Iran, Islamic Rep. of	9 (1.8)	452 (13.3)	68 (3.5)	436 (5.1)	23 (3.0)	404 (6.6)	10.5 (0.13)
Romania	9 (2.3)	476 (22.8)	61 (3.7)	498 (5.7)	30 (3.3)	449 (11.6)	10.2 (0.16)
Kuwait	9 (2.4)	343 (14.9)	65 (3.8)	342 (4.1)	26 (3.4)	340 (6.5)	10.2 (0.14)
Austria	8 (1.9)	521 (5.4)	74 (2.8)	510 (3.2)	18 (2.5)	495 (5.7)	10.4 (0.13)
Denmark	8 (2.2)	553 (6.9)	69 (3.4)	543 (2.8)	23 (2.9)	528 (5.2)	10.3 (0.11)
Oman	8 (1.6)	414 (7.9)	65 (2.8)	390 (3.4)	27 (2.6)	365 (5.3)	10.2 (0.10)
Azerbaijan	8 (2.1)	488 (24.2)	39 (3.5)	468 (8.0)	53 (3.5)	458 (6.8)	9.5 (0.14)
Chinese Taipei	7 (1.9)	589 (8.1)	67 (3.8)	594 (2.4)	26 (3.6)	585 (4.8)	10.1 (0.16)
Poland	7 (2.0)	479 (6.4)	76 (3.2)	483 (2.6)	17 (2.8)	472 (3.8)	10.3 (0.12)
Spain	7 (2.1)	496 (7.0)	54 (4.4)	495 (3.0)	39 (4.1)	462 (4.1)	9.7 (0.16)
Sweden r	7 (1.7)	532 (6.4)	63 (4.3)	508 (2.7)	31 (4.3)	491 (4.3)	10.0 (0.16)
Chile	6 (2.0)	524 (8.0)	43 (3.7)	473 (4.5)	51 (4.0)	445 (4.9)	9.2 (0.16)
Serbia	5 (1.9)	558 (13.6)	69 (3.6)	521 (3.5)	25 (3.3)	493 (7.2)	10.1 (0.13)
Portugal	4 (1.7)	581 (17.2)	56 (4.7)	540 (3.3)	40 (4.6)	516 (6.0)	9.9 (0.18)
Finland	4 (1.6)	550 (9.8)	63 (3.2)	550 (2.4)	33 (3.4)	537 (4.3)	9.9 (0.12)
Turkey	4 (1.1)	532 (11.2)	39 (3.3)	490 (9.3)	57 (3.3)	450 (5.1)	8.8 (0.14)
Lithuania	3 (1.0)	532 (11.9)	74 (3.2)	536 (3.4)	23 (3.2)	526 (4.9)	10.2 (0.09)
Thailand	3 (1.4)	436 (8.6)	55 (4.2)	469 (6.5)	42 (4.3)	448 (7.8)	9.5 (0.16)
Singapore	3 (1.0)	619 (22.8)	61 (2.5)	610 (4.4)	36 (2.5)	597 (5.2)	9.6 (0.10)
Georgia	3 (1.2)	474 (21.5)	59 (4.0)	459 (4.4)	38 (4.0)	435 (6.9)	9.7 (0.13)
Armenia	3 (1.2)	452 (23.8)	57 (3.2)	458 (4.3)	40 (3.2)	445 (5.6)	9.6 (0.12)
Yemen	3 (1.5)	254 (89.1)	39 (4.5)	248 (9.7)	58 (4.4)	245 (7.6)	8.9 (0.18)
Hong Kong SAR	2 (1.3)	~ ~	73 (4.0)	606 (3.7)	25 (3.9)	590 (9.5)	9.9 (0.14)
Morocco	2 (0.8)	~ ~	16 (2.3)	382 (13.8)	82 (2.2)	328 (5.2)	7.6 (0.12)
Belgium (Flemish)	2 (1.1)	~ ~	67 (3.4)	554 (2.1)	31 (3.3)	539 (3.8)	9.8 (0.10)
Slovenia	2 (1.1)	~ ~	66 (3.7)	516 (2.7)	32 (3.5)	506 (3.2)	9.7 (0.10)
Czech Republic	2 (0.9)	~ ~	43 (4.5)	513 (3.3)	55 (4.5)	508 (3.6)	9.0 (0.14)
Tunisia	2 (0.8)	~ ~	42 (3.3)	371 (6.3)	57 (3.2)	349 (5.1)	8.9 (0.15)
Japan	1 (0.8)	~ ~	57 (3.5)	589 (2.1)	42 (3.5)	581 (2.6)	9.4 (0.12)
Norway	1 (0.7)	~ ~	74 (4.2)	496 (3.1)	24 (4.1)	488 (6.0)	9.9 (0.14)
Italy	1 (0.4)	~ ~	56 (3.8)	512 (3.6)	43 (3.9)	504 (4.1)	9.3 (0.12)
Germany	1 (0.8)	~ ~	61 (3.7)	540 (2.1)	38 (3.8)	510 (3.5)	9.4 (0.11)
Hungary	1 (0.7)	~ ~	59 (3.5)	533 (3.8)	40 (3.5)	486 (6.1)	9.4 (0.13)
Slovak Republic	1 (0.4)	~ ~	51 (3.5)	514 (3.4)	48 (3.5)	498 (6.0)	9.2 (0.12)
Russian Federation	1 (0.0)	~ ~	52 (3.9)	544 (3.7)	47 (4.0)	539 (5.9)	9.3 (0.12)
Netherlands r	0 (0.0)	~ ~	40 (4.2)	545 (3.9)	60 (4.2)	535 (2.7)	9.0 (0.13)
International Avg.	7 (0.3)	503 (3.3)	60 (0.5)	496 (0.7)	33 (0.5)	477 (0.9)	

Centerpoint of scale set at 10.

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

A tilde (~) indicates insufficient data to report achievement.

An "r" indicates data are available for at least 70% but less than 85% of the students.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 6.3: School Emphasis on Academic Success - Teacher Reports (Continued)

TIMSS 2011 Mathematics 4th Grade

Country	Very High Emphasis		High Emphasis		Medium Emphasis		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Sixth Grade Participants							
Honduras	12 (3.0)	438 (18.2)	52 (4.4)	397 (6.5)	37 (4.4)	382 (9.4)	10.2 (0.21)
Botswana	4 (2.0)	503 (31.9)	35 (4.1)	432 (6.8)	61 (4.0)	408 (4.7)	8.9 (0.18)
Yemen	2 (1.2)	~ ~	43 (4.4)	355 (9.1)	55 (4.5)	341 (7.7)	8.8 (0.17)
Benchmarking Participants							
Dubai, UAE	24 (2.3)	505 (6.4)	63 (2.7)	465 (3.4)	14 (1.4)	437 (12.0)	11.2 (0.07)
Florida, US r	20 (4.3)	563 (9.8)	59 (4.1)	542 (4.4)	22 (3.5)	531 (6.1)	10.9 (0.27)
Alberta, Canada r	18 (4.1)	508 (6.1)	70 (3.7)	511 (2.9)	11 (2.7)	478 (10.5)	11.2 (0.17)
Abu Dhabi, UAE	13 (3.0)	441 (15.4)	66 (4.1)	421 (4.4)	20 (4.2)	394 (15.9)	11.0 (0.18)
Ontario, Canada	11 (2.3)	531 (7.7)	62 (3.7)	522 (3.3)	26 (3.4)	507 (6.3)	10.4 (0.16)
North Carolina, US	7 (3.1)	587 (18.1)	67 (4.6)	554 (5.1)	25 (5.2)	539 (6.4)	10.3 (0.24)
Quebec, Canada	5 (1.9)	555 (11.7)	67 (4.1)	535 (2.8)	28 (4.1)	523 (3.6)	10.2 (0.15)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

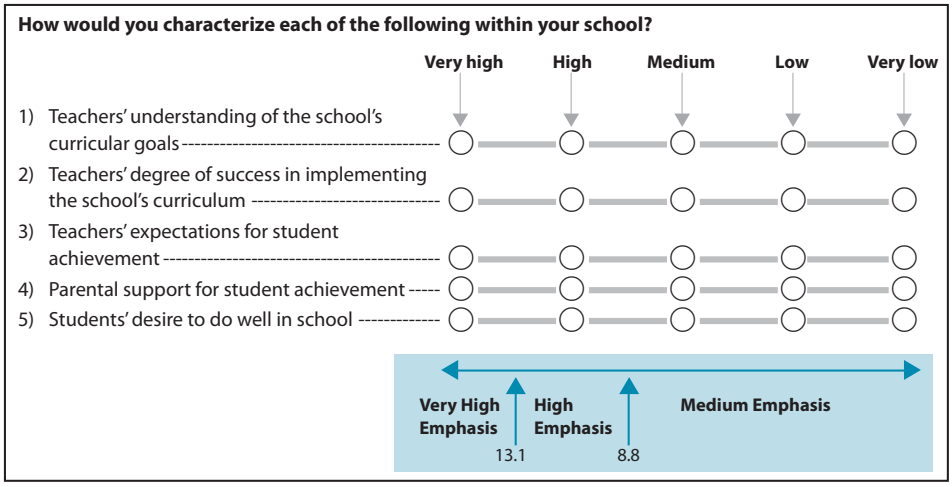


Exhibit 6.4: School Emphasis on Academic Success - Teacher Reports

TIMSS 2011 Mathematics 8th Grade

Reported by Teachers

Students were scored according to their teachers' responses characterizing five aspects on the *School Emphasis on Academic Success* scale. Students in schools where their teachers reported a **Very High Emphasis** on academic success had a score on the scale of at least 13.6, which corresponds to their teachers characterizing three of the five aspects as "very high" and the other two as "high," on average. Students in schools with a **Medium Emphasis** on academic success had a score no higher than 9.5, which corresponds to their teachers characterizing three of the five aspects as "medium" and the other two as "high," on average. All other students attended schools with a **High Emphasis** on academic success.

Country	Very High Emphasis		High Emphasis		Medium Emphasis		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Qatar	19 (4.2)	440 (15.5)	54 (4.5)	413 (6.0)	27 (3.2)	380 (8.4)	11.4 (0.17)
England	16 (2.4)	526 (11.0)	59 (4.1)	508 (7.3)	24 (3.9)	488 (12.2)	11.2 (0.19)
United States r	13 (2.0)	538 (10.1)	55 (2.6)	517 (4.9)	32 (2.4)	494 (4.7)	10.8 (0.12)
Australia r	13 (2.4)	569 (15.2)	50 (3.7)	515 (7.7)	37 (3.9)	475 (7.5)	10.4 (0.17)
United Arab Emirates	11 (1.6)	500 (9.0)	62 (2.3)	457 (2.5)	26 (2.2)	430 (4.2)	11.0 (0.09)
Chinese Taipei	11 (2.2)	659 (11.6)	63 (3.7)	612 (4.7)	26 (3.3)	583 (5.7)	11.0 (0.12)
Israel	10 (2.1)	561 (13.3)	61 (3.1)	528 (5.1)	29 (2.6)	485 (9.4)	10.8 (0.12)
Malaysia	9 (2.1)	473 (19.9)	59 (4.1)	447 (6.7)	32 (3.9)	419 (9.1)	10.7 (0.16)
New Zealand	9 (2.1)	520 (17.4)	59 (3.4)	491 (6.8)	32 (3.2)	468 (9.6)	10.7 (0.16)
Indonesia	9 (3.0)	388 (15.9)	55 (4.4)	391 (6.2)	36 (4.3)	377 (8.2)	10.7 (0.17)
Saudi Arabia	8 (2.4)	406 (19.0)	54 (4.1)	406 (5.8)	38 (3.8)	376 (6.6)	10.5 (0.16)
Korea, Rep. of	8 (1.5)	624 (8.2)	56 (3.3)	615 (4.4)	36 (3.1)	605 (4.3)	10.4 (0.13)
Bahrain	5 (0.1)	505 (6.2)	43 (3.1)	428 (4.4)	52 (3.1)	384 (3.1)	9.9 (0.08)
Oman	5 (1.5)	417 (12.7)	54 (3.0)	385 (4.2)	41 (2.7)	334 (4.9)	10.1 (0.12)
Jordan	5 (1.9)	447 (17.2)	50 (4.2)	416 (5.5)	45 (3.8)	390 (6.1)	10.1 (0.14)
Iran, Islamic Rep. of	5 (1.8)	484 (19.9)	47 (3.5)	424 (6.5)	48 (3.4)	399 (5.7)	9.9 (0.15)
Japan	5 (1.9)	599 (14.3)	52 (4.2)	578 (3.9)	43 (4.2)	557 (3.5)	10.0 (0.18)
Ghana	5 (1.7)	367 (23.9)	66 (3.8)	337 (5.8)	29 (3.6)	310 (5.0)	10.7 (0.14)
Turkey	4 (1.4)	586 (37.4)	27 (3.0)	481 (7.5)	69 (3.1)	433 (4.1)	8.7 (0.17)
Romania	4 (1.6)	523 (18.8)	47 (3.7)	473 (6.1)	49 (3.6)	438 (6.4)	10.0 (0.13)
Kazakhstan	4 (1.6)	503 (23.5)	69 (3.8)	484 (5.7)	27 (3.6)	493 (6.7)	10.6 (0.11)
Lebanon	4 (1.5)	496 (8.9)	53 (4.2)	465 (5.9)	43 (4.0)	427 (4.6)	10.1 (0.17)
Macedonia, Rep. of r	4 (1.6)	420 (23.7)	45 (4.1)	435 (10.4)	51 (4.1)	414 (7.1)	9.8 (0.15)
Singapore	4 (1.1)	681 (12.8)	55 (2.6)	625 (5.1)	41 (2.4)	587 (6.2)	10.2 (0.09)
Sweden r	3 (1.4)	517 (13.6)	55 (3.7)	492 (2.9)	42 (3.5)	475 (2.9)	10.2 (0.13)
Norway	3 (1.4)	501 (5.4)	61 (4.4)	482 (2.7)	36 (4.4)	462 (3.3)	10.4 (0.12)
Syrian Arab Republic	3 (1.4)	409 (37.0)	45 (4.5)	386 (5.8)	52 (4.5)	371 (6.7)	9.7 (0.18)
Chile	3 (1.1)	498 (16.1)	30 (3.4)	441 (6.6)	67 (3.5)	403 (3.4)	9.0 (0.15)
Hong Kong SAR	2 (1.4)	~ ~	50 (4.5)	615 (6.6)	47 (4.3)	553 (6.9)	10.0 (0.15)
Morocco	2 (0.9)	~ ~	23 (2.8)	393 (5.9)	76 (3.0)	363 (2.1)	8.5 (0.14)
Lithuania	2 (0.9)	~ ~	60 (3.6)	508 (4.0)	38 (3.6)	493 (4.1)	10.1 (0.11)
Thailand	2 (1.1)	~ ~	44 (3.6)	442 (7.4)	55 (3.7)	415 (5.8)	9.7 (0.13)
Palestinian Nat'l Auth.	1 (1.0)	~ ~	51 (4.3)	406 (4.9)	47 (4.1)	403 (5.8)	9.8 (0.12)
Finland	1 (0.9)	~ ~	51 (3.7)	518 (3.4)	47 (3.8)	510 (2.8)	10.1 (0.11)
Slovenia	1 (0.5)	~ ~	47 (3.0)	510 (2.9)	52 (3.0)	500 (2.7)	9.7 (0.09)
Hungary	1 (0.5)	~ ~	42 (3.7)	529 (5.3)	57 (3.7)	486 (5.1)	9.4 (0.13)
Georgia	0 (0.4)	~ ~	22 (3.5)	447 (10.7)	77 (3.5)	426 (4.2)	8.9 (0.11)
Tunisia	0 (0.2)	~ ~	24 (3.1)	437 (7.9)	76 (3.1)	421 (2.9)	8.7 (0.12)
Armenia	0 (0.0)	~ ~	26 (3.4)	482 (6.7)	74 (3.4)	460 (3.5)	9.0 (0.12)
Russian Federation	0 (0.2)	~ ~	31 (3.4)	568 (6.0)	69 (3.4)	527 (4.2)	9.0 (0.11)
Italy	0 (0.0)	~ ~	36 (3.9)	508 (4.3)	64 (3.9)	494 (3.7)	9.2 (0.12)
Ukraine	0 (0.0)	~ ~	33 (4.1)	505 (6.8)	67 (4.1)	467 (4.7)	9.3 (0.11)
International Avg.	5 (0.3)	506 (3.4)	48 (0.6)	478 (0.9)	47 (0.5)	452 (0.9)	

Centerpoint of scale set at 10.

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

A tilde (~) indicates insufficient data to report achievement.

An "r" indicates data are available for at least 70% but less than 85% of the students.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 6.4: School Emphasis on Academic Success - Teacher Reports (Continued)

TIMSS 2011 Mathematics 8th Grade

Country		Very High Emphasis		High Emphasis		Medium Emphasis		Average Scale Score
		Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Ninth Grade Participants								
Honduras		5 (2.2)	349 (23.7)	42 (5.2)	338 (8.1)	53 (4.9)	336 (5.6)	9.7 (0.19)
South Africa		2 (0.6)	~ ~	39 (3.6)	368 (5.9)	59 (3.6)	339 (3.7)	9.4 (0.14)
Botswana		0 (0.0)	~ ~	32 (3.6)	415 (5.9)	67 (3.6)	387 (2.6)	9.0 (0.14)
Benchmarking Participants								
Alberta, Canada		18 (3.1)	517 (6.4)	64 (3.6)	503 (3.4)	17 (3.2)	498 (5.4)	11.4 (0.15)
Massachusetts, US		17 (3.6)	593 (12.3)	58 (5.7)	557 (8.3)	25 (4.7)	546 (17.0)	11.3 (0.20)
Abu Dhabi, UAE		15 (3.7)	499 (16.3)	54 (4.5)	444 (4.4)	30 (4.2)	434 (7.0)	11.0 (0.20)
California, US	r	13 (5.0)	534 (21.0)	55 (5.8)	504 (9.5)	32 (4.2)	462 (9.4)	10.5 (0.25)
Colorado, US	r	11 (4.4)	555 (16.1)	57 (6.8)	534 (6.9)	31 (6.0)	475 (12.2)	10.9 (0.26)
North Carolina, US	r	11 (4.1)	561 (36.0)	65 (6.3)	549 (8.7)	24 (5.6)	511 (8.9)	11.0 (0.23)
Dubai, UAE		11 (1.8)	533 (11.5)	66 (3.3)	479 (3.8)	23 (2.8)	436 (6.4)	11.2 (0.12)
Minnesota, US		10 (2.3)	584 (26.1)	59 (4.4)	544 (7.1)	32 (4.8)	537 (10.0)	10.8 (0.21)
Connecticut, US		9 (3.6)	539 (13.9)	68 (5.7)	528 (8.5)	22 (4.9)	490 (13.3)	10.9 (0.21)
Ontario, Canada		7 (2.0)	530 (9.7)	62 (3.9)	516 (3.5)	32 (3.9)	502 (3.8)	10.7 (0.16)
Alabama, US	r	5 (2.9)	565 (18.2)	56 (7.8)	468 (9.4)	39 (7.5)	454 (7.4)	10.3 (0.32)
Indiana, US	r	4 (2.1)	561 (18.5)	74 (5.4)	519 (6.0)	22 (5.0)	508 (11.7)	10.6 (0.18)
Quebec, Canada		4 (1.7)	568 (20.4)	46 (4.2)	542 (4.1)	50 (4.1)	520 (3.2)	9.6 (0.17)
Florida, US	r	2 (1.8)	~ ~	48 (7.3)	536 (10.9)	50 (7.6)	504 (11.1)	9.9 (0.33)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

How would you characterize each of the following within your school?

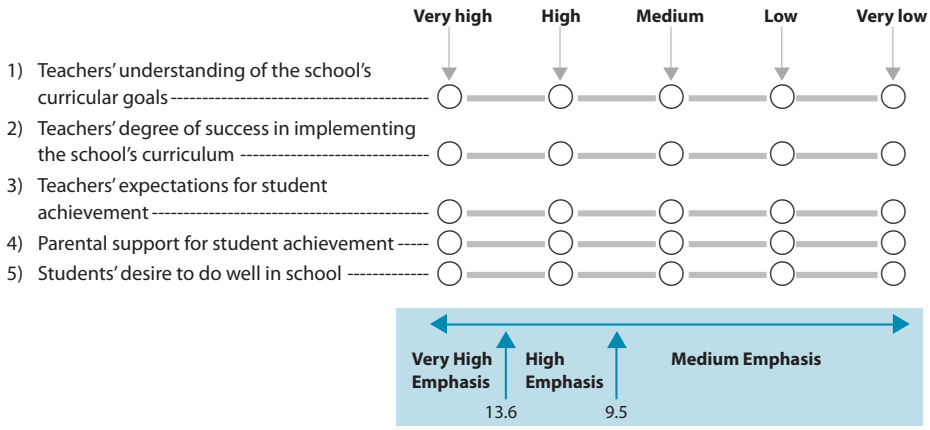


Exhibit 6.5: Principals Spend Time on Leadership Activities

TIMSS 2011 4th Grade Mathematics

Reported by Principals

Country	Percent of Students Whose Principals Spend "A Lot of Time"								
	Promoting the School's Educational Vision or Goals	Developing the School's Curricular and Educational Goals	Monitoring Teachers' Implementation of the School's Educational Goals in Their Teaching	Monitoring Students' Learning Progress to Ensure that the School's Educational Goals Are Reached	Keeping an Orderly Atmosphere in the School	Addressing Disruptive Student Behavior	Advising Teachers Who Have Questions or Problems with Their Teaching	Initiating Educational Projects or Improvements	Participating in Professional Development Activities Specifically for School Principals
Armenia	80 (3.4)	75 (4.0)	60 (4.3)	62 (3.8)	66 (4.3)	32 (4.2)	23 (3.7)	23 (3.4)	31 (4.0)
Australia	60 (4.1)	73 (3.8)	52 (4.6)	68 (3.8)	63 (3.6)	35 (3.8)	27 (3.4)	53 (4.4)	33 (3.7)
Austria	41 (3.9)	13 (3.0)	24 (3.4)	27 (3.2)	73 (4.3)	41 (4.6)	39 (4.5)	22 (3.6)	44 (3.9)
Azerbaijan	50 (4.4)	55 (4.3)	33 (4.0)	40 (4.8)	79 (3.7)	38 (4.4)	29 (3.3)	27 (3.9)	38 (4.0)
Bahrain	70 (4.6)	77 (4.2)	85 (2.9)	85 (3.1)	87 (3.0)	52 (4.8)	72 (4.4)	71 (4.6)	46 (5.8)
Belgium (Flemish)	35 (3.8)	30 (3.7)	24 (3.8)	22 (3.2)	36 (4.4)	31 (3.7)	28 (4.0)	29 (4.4)	34 (4.3)
Chile	59 (4.0)	75 (3.8)	55 (4.3)	63 (4.5)	74 (3.7)	62 (3.4)	39 (4.5)	45 (4.1)	37 (3.9)
Chinese Taipei	72 (3.6)	69 (3.8)	59 (3.9)	54 (3.6)	49 (4.4)	15 (3.0)	44 (4.2)	53 (4.2)	57 (4.4)
Croatia	64 (3.9)	69 (3.9)	39 (4.2)	41 (3.8)	84 (2.9)	50 (4.0)	43 (4.3)	32 (4.0)	70 (3.7)
Czech Republic	69 (3.9)	64 (4.0)	54 (4.3)	66 (3.8)	95 (1.7)	58 (4.2)	40 (4.5)	61 (3.7)	42 (4.1)
Denmark	r 28 (3.9)	r 24 (3.6)	r 6 (2.0)	r 9 (1.9)	r 62 (4.0)	r 26 (2.9)	r 24 (3.5)	r 24 (3.3)	r 17 (2.9)
England	61 (4.0)	62 (5.0)	56 (4.4)	76 (4.5)	53 (4.8)	25 (4.0)	17 (3.3)	37 (4.6)	17 (3.7)
Finland	36 (3.8)	34 (4.4)	18 (3.0)	12 (2.1)	33 (4.6)	26 (4.1)	16 (2.9)	28 (4.1)	23 (3.6)
Georgia	42 (4.8)	36 (4.5)	39 (4.0)	55 (3.7)	72 (3.9)	51 (4.2)	19 (3.5)	20 (3.3)	27 (3.5)
Germany	49 (3.4)	47 (3.3)	15 (2.6)	18 (2.6)	56 (3.6)	49 (3.5)	28 (3.2)	24 (3.2)	17 (2.6)
Hong Kong SAR	52 (4.5)	68 (4.3)	58 (4.4)	62 (4.0)	60 (4.1)	11 (2.6)	16 (3.4)	42 (4.8)	31 (4.3)
Hungary	80 (3.6)	72 (4.0)	59 (4.0)	62 (4.2)	79 (3.2)	59 (4.0)	34 (4.0)	41 (4.4)	35 (4.2)
Iran, Islamic Rep. of	77 (3.1)	88 (2.7)	79 (3.9)	86 (2.5)	89 (2.0)	82 (2.7)	61 (3.6)	44 (3.9)	67 (3.3)
Ireland	40 (4.5)	60 (4.5)	19 (3.2)	34 (4.4)	64 (3.9)	29 (4.0)	10 (2.4)	31 (3.8)	16 (2.8)
Italy	83 (3.6)	62 (3.8)	43 (3.9)	47 (4.2)	49 (3.7)	31 (3.3)	48 (3.7)	61 (3.7)	35 (3.3)
Japan	40 (4.0)	28 (3.8)	47 (4.1)	31 (4.2)	41 (4.0)	15 (3.0)	27 (3.6)	26 (3.9)	17 (3.1)
Kazakhstan	73 (3.0)	77 (3.5)	74 (3.9)	66 (3.9)	69 (3.5)	44 (3.9)	47 (3.6)	58 (4.4)	54 (4.4)
Korea, Rep. of	88 (2.5)	82 (3.5)	81 (3.7)	75 (4.0)	88 (2.9)	77 (3.6)	72 (3.8)	75 (4.0)	80 (2.9)
Kuwait	68 (4.0)	58 (4.1)	82 (3.2)	85 (3.0)	84 (3.2)	73 (3.5)	73 (3.7)	72 (3.6)	67 (4.2)
Lithuania	74 (3.7)	90 (2.4)	60 (3.6)	68 (4.0)	62 (4.5)	42 (3.8)	48 (4.3)	41 (4.3)	44 (3.9)
Malta	58 (0.1)	67 (0.1)	32 (0.1)	40 (0.1)	71 (0.1)	39 (0.1)	39 (0.1)	44 (0.1)	26 (0.1)
Morocco	64 (3.4)	58 (3.6)	63 (3.9)	59 (4.0)	91 (2.1)	66 (3.0)	56 (3.7)	43 (3.8)	42 (3.9)
Netherlands	r 33 (5.2)	r 49 (5.5)	r 48 (4.6)	r 44 (5.9)	r 14 (4.1)	r 15 (4.4)	r 31 (5.5)	r 43 (5.1)	r 23 (5.1)
New Zealand	65 (3.5)	70 (4.0)	45 (3.8)	71 (3.5)	47 (3.6)	21 (3.1)	24 (3.5)	41 (3.6)	18 (3.0)
Northern Ireland	47 (4.5)	73 (3.9)	r 35 (4.6)	61 (4.2)	54 (5.2)	13 (2.9)	r 7 (2.1)	r 35 (4.5)	r 23 (4.5)
Norway	27 (4.4)	19 (3.7)	17 (3.3)	17 (3.2)	56 (4.6)	31 (4.4)	16 (3.5)	23 (4.1)	24 (4.3)
Oman	40 (3.2)	r 18 (2.4)	75 (3.4)	80 (3.1)	82 (2.5)	45 (3.5)	51 (3.5)	36 (3.4)	24 (2.5)
Poland	56 (3.9)	49 (4.2)	59 (4.0)	75 (3.3)	76 (3.8)	40 (3.9)	29 (3.9)	51 (4.1)	54 (4.2)
Portugal	63 (4.4)	50 (5.4)	35 (4.7)	41 (4.9)	49 (4.9)	38 (5.3)	8 (2.6)	28 (5.4)	6 (1.8)
Qatar	70 (2.5)	81 (2.3)	81 (2.4)	81 (2.5)	85 (2.5)	64 (2.7)	69 (2.9)	61 (3.4)	54 (3.2)
Romania	84 (3.3)	84 (3.2)	81 (3.5)	84 (3.0)	87 (2.5)	73 (3.6)	57 (4.3)	63 (3.8)	69 (4.2)
Russian Federation	80 (2.8)	81 (2.6)	81 (2.6)	74 (2.9)	87 (2.1)	64 (3.1)	34 (3.1)	52 (3.6)	64 (4.0)
Saudi Arabia	48 (4.4)	61 (4.1)	77 (3.3)	76 (3.5)	78 (3.5)	57 (3.7)	52 (3.9)	45 (4.4)	40 (4.3)
Serbia	63 (3.3)	72 (3.9)	47 (4.8)	42 (4.6)	64 (3.7)	48 (4.0)	41 (4.1)	47 (4.2)	31 (3.7)
Singapore	76 (0.0)	80 (0.0)	66 (0.0)	77 (0.0)	66 (0.0)	32 (0.0)	33 (0.0)	58 (0.0)	47 (0.0)
Slovak Republic	56 (3.6)	69 (3.6)	45 (3.9)	42 (3.9)	60 (3.7)	55 (3.3)	34 (3.6)	46 (3.7)	46 (3.8)
Slovenia	68 (3.1)	62 (4.1)	61 (3.5)	69 (4.0)	92 (2.2)	59 (3.8)	53 (4.0)	62 (3.9)	73 (3.4)
Spain	58 (4.1)	62 (3.8)	40 (4.4)	47 (4.4)	68 (3.8)	39 (4.2)	19 (3.7)	47 (4.1)	33 (3.6)
Sweden	52 (4.4)	40 (4.8)	17 (3.2)	28 (4.2)	24 (3.7)	19 (3.6)	27 (4.0)	28 (4.1)	16 (3.6)
Thailand	68 (3.9)	74 (3.9)	76 (3.3)	77 (3.6)	94 (2.0)	51 (3.9)	74 (3.4)	68 (4.4)	69 (3.9)
Tunisia	49 (4.4)	52 (4.6)	54 (4.4)	61 (4.9)	86 (2.9)	61 (3.8)	49 (4.0)	26 (3.6)	18 (2.8)
Turkey	63 (3.2)	56 (3.7)	62 (3.6)	54 (3.6)	86 (2.4)	79 (2.8)	55 (3.7)	45 (3.4)	46 (3.2)
United Arab Emirates	69 (2.1)	77 (2.2)	82 (1.8)	85 (1.4)	82 (1.8)	55 (2.1)	62 (2.0)	65 (2.0)	47 (1.9)
United States	72 (2.8)	68 (2.3)	71 (2.4)	76 (2.1)	69 (3.0)	42 (2.8)	42 (2.6)	46 (2.9)	34 (2.7)
Yemen	48 (4.6)	47 (4.2)	71 (4.3)	64 (4.3)	84 (3.2)	64 (4.7)	52 (4.7)	18 (3.5)	28 (4.0)
International Avg.	59 (0.5)	60 (0.5)	53 (0.5)	57 (0.5)	68 (0.5)	44 (0.5)	39 (0.5)	43 (0.6)	39 (0.5)

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.
An "r" indicates data are available for at least 70% but less than 85% of the students.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 6.5: Principals Spend Time on Leadership Activities (Continued)

Country	Percent of Students Whose Principals Spend "A Lot of Time"								
	Promoting the School's Educational Vision or Goals	Developing the School's Curricular and Educational Goals	Monitoring Teachers' Implementation of the School's Educational Goals in Their Teaching	Monitoring Students' Learning Progress to Ensure that the School's Educational Goals Are Reached	Keeping an Orderly Atmosphere in the School	Addressing Disruptive Student Behavior	Advising Teachers Who Have Questions or Problems with Their Teaching	Initiating Educational Projects or Improvements	Participating in Professional Development Activities Specifically for School Principals
Sixth Grade Participants									
Botswana	68 (3.7)	67 (3.9)	83 (2.8)	82 (3.0)	87 (2.5)	62 (4.6)	57 (3.6)	45 (4.2)	52 (4.7)
Honduras	58 (4.5)	63 (4.7)	51 (5.1)	65 (4.4)	90 (2.5)	72 (4.8)	56 (4.6)	63 (4.7)	51 (4.9)
Yemen	49 (4.6)	53 (4.3)	75 (3.9)	66 (4.2)	84 (2.9)	64 (4.6)	56 (4.3)	19 (3.8)	32 (4.3)
Benchmarking Participants									
Alberta, Canada	63 (4.3)	60 (4.7)	44 (4.6)	45 (4.8)	67 (4.1)	30 (4.4)	23 (4.1)	38 (4.4)	30 (4.0)
Ontario, Canada	65 (4.2)	76 (4.0)	53 (4.4)	61 (4.4)	75 (3.8)	52 (4.6)	32 (4.2)	43 (4.3)	44 (4.0)
Quebec, Canada	44 (4.7)	41 (4.3)	18 (3.4)	36 (3.8)	47 (4.3)	47 (4.7)	29 (4.0)	31 (4.0)	19 (3.2)
Abu Dhabi, UAE	78 (3.9)	79 (3.6)	83 (3.3)	87 (2.7)	82 (3.0)	51 (4.4)	66 (4.1)	64 (4.4)	59 (3.7)
Dubai, UAE	72 (0.4)	82 (0.4)	79 (0.4)	80 (0.4)	80 (0.2)	58 (0.5)	55 (0.5)	71 (0.4)	43 (0.3)
Florida, US	r 82 (4.1)	r 79 (5.5)	r 79 (5.0)	r 88 (2.8)	r 77 (6.1)	r 39 (6.5)	r 36 (6.0)	r 38 (5.9)	r 43 (6.3)
North Carolina, US	81 (5.9)	76 (6.8)	88 (4.7)	84 (5.9)	72 (7.2)	29 (7.5)	33 (6.7)	30 (7.9)	41 (7.8)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 6.6: Principals Spend Time on Leadership Activities

TIMSS 2011 8th Grade Mathematics

Reported by Principals

Country	Percent of Students Whose Principals Spend "A Lot of Time"								
	Promoting the School's Educational Vision or Goals	Developing the School's Curricular and Educational Goals	Monitoring Teachers' Implementation of the School's Educational Goals in Their Teaching	Monitoring Students' Learning Progress to Ensure that the School's Educational Goals Are Reached	Keeping an Orderly Atmosphere in the School	Addressing Disruptive Student Behavior	Advising Teachers Who Have Questions or Problems with Their Teaching	Initiating Educational Projects or Improvements	Participating in Professional Development Activities Specifically for School Principals
Armenia	79 (3.5)	75 (3.4)	66 (3.8)	59 (3.9)	69 (4.4)	31 (4.3)	26 (3.6)	23 (3.3)	32 (4.0)
Australia	64 (3.3)	63 (4.1)	34 (3.5)	53 (3.9)	55 (3.5)	35 (3.8)	19 (3.0)	52 (4.1)	30 (3.9)
Bahrain	60 (0.3)	71 (0.3)	78 (0.3)	81 (0.3)	88 (0.2)	70 (0.3)	67 (0.3)	61 (0.3)	46 (0.3)
Chile	65 (4.1)	78 (3.2)	54 (4.4)	58 (4.8)	78 (3.0)	66 (3.7)	37 (4.1)	46 (4.2)	38 (4.2)
Chinese Taipei	62 (3.8)	54 (3.8)	47 (4.0)	54 (4.0)	75 (3.5)	22 (3.6)	25 (3.7)	29 (3.6)	31 (4.2)
England	64 (4.6)	67 (4.4)	55 (4.5)	75 (3.8)	51 (4.6)	29 (4.0)	20 (3.1)	33 (4.7)	9 (2.7)
Finland	34 (4.4)	25 (3.9)	22 (3.8)	28 (4.0)	44 (4.3)	37 (4.1)	17 (3.1)	21 (3.9)	16 (3.2)
Georgia	76 (3.7)	71 (4.3)	72 (3.1)	75 (3.4)	84 (2.9)	68 (4.2)	50 (4.2)	38 (3.9)	52 (3.7)
Ghana	67 (3.9)	48 (4.5)	86 (3.0)	88 (2.8)	89 (2.8)	57 (3.8)	50 (4.4)	25 (3.7)	36 (4.1)
Hong Kong SAR	41 (4.9)	47 (5.1)	48 (4.9)	41 (5.3)	54 (4.9)	11 (2.9)	21 (3.9)	21 (4.4)	24 (4.2)
Hungary	78 (3.7)	71 (3.7)	57 (4.4)	63 (4.0)	78 (3.6)	58 (4.4)	40 (4.1)	39 (4.0)	44 (4.1)
Indonesia	85 (2.8)	85 (3.8)	80 (3.8)	85 (3.4)	95 (2.3)	87 (2.8)	76 (3.5)	38 (4.8)	75 (3.8)
Iran, Islamic Rep. of	84 (2.2)	91 (1.9)	81 (3.0)	92 (2.0)	93 (1.6)	80 (2.9)	48 (3.5)	48 (3.8)	61 (3.7)
Israel	80 (3.4)	71 (3.7)	62 (4.0)	75 (3.6)	85 (3.1)	76 (3.5)	64 (4.1)	67 (4.0)	64 (4.2)
Italy	79 (2.9)	61 (4.0)	40 (4.0)	56 (4.2)	64 (4.0)	49 (4.2)	39 (3.5)	61 (3.7)	29 (3.3)
Japan	31 (3.9)	21 (3.7)	32 (4.0)	19 (3.0)	48 (3.9)	21 (3.2)	18 (3.4)	21 (3.7)	11 (2.7)
Jordan	62 (3.9)	67 (3.8)	88 (2.7)	82 (3.3)	95 (2.0)	84 (2.8)	72 (3.6)	42 (3.6)	41 (3.9)
Kazakhstan	72 (3.8)	79 (3.0)	66 (4.1)	71 (3.7)	64 (4.2)	41 (4.0)	46 (4.1)	58 (4.0)	47 (4.3)
Korea, Rep. of	88 (3.1)	78 (3.7)	77 (3.2)	73 (3.5)	89 (2.5)	70 (3.1)	61 (3.7)	64 (3.7)	75 (3.1)
Lebanon	75 (3.7)	67 (3.8)	76 (4.1)	84 (3.0)	85 (3.2)	73 (3.9)	76 (3.7)	42 (3.9)	45 (4.2)
Lithuania	74 (3.8)	82 (3.4)	42 (4.2)	61 (4.3)	71 (3.9)	41 (4.1)	38 (4.2)	47 (4.5)	42 (4.2)
Macedonia, Rep. of	50 (4.0)	57 (3.7)	46 (3.9)	53 (4.2)	59 (3.7)	42 (3.9)	37 (3.7)	45 (3.7)	43 (3.7)
Malaysia	71 (3.7)	76 (2.9)	74 (3.5)	79 (2.7)	87 (2.4)	75 (3.4)	55 (4.1)	36 (3.5)	42 (3.8)
Morocco	61 (3.3)	48 (2.6)	58 (3.0)	59 (3.7)	92 (1.8)	75 (3.3)	51 (3.0)	55 (3.6)	39 (3.2)
New Zealand	57 (5.1)	59 (5.2)	30 (4.4)	42 (5.6)	54 (5.1)	31 (5.3)	16 (3.3)	37 (3.7)	20 (4.5)
Norway	29 (3.8)	20 (3.6)	20 (3.1)	22 (3.2)	54 (3.7)	45 (4.7)	20 (3.6)	15 (3.1)	16 (3.6)
Oman	52 (3.4)	21 (2.3)	79 (2.5)	77 (2.5)	86 (2.2)	47 (3.3)	56 (3.3)	28 (2.9)	28 (3.4)
Palestinian Nat'l Auth.	60 (4.1)	58 (3.8)	90 (1.5)	92 (2.0)	89 (2.5)	75 (3.3)	58 (3.9)	32 (3.8)	37 (3.8)
Qatar	72 (0.8)	78 (0.5)	79 (1.0)	83 (1.1)	82 (1.1)	69 (1.0)	66 (1.0)	57 (0.9)	54 (0.9)
Romania	87 (2.8)	86 (3.2)	85 (2.9)	84 (3.6)	92 (2.6)	69 (4.1)	55 (4.4)	65 (4.0)	71 (4.2)
Russian Federation	80 (2.7)	82 (2.6)	68 (3.4)	69 (2.8)	78 (2.7)	51 (3.6)	27 (2.8)	54 (3.7)	61 (3.5)
Saudi Arabia	53 (4.3)	59 (3.8)	81 (3.2)	72 (3.2)	88 (2.7)	70 (3.5)	56 (4.5)	37 (3.6)	34 (3.7)
Singapore	68 (0.0)	66 (0.0)	63 (0.0)	72 (0.0)	56 (0.0)	27 (0.0)	21 (0.0)	42 (0.0)	26 (0.0)
Slovenia	58 (3.6)	56 (4.2)	60 (3.9)	62 (3.6)	83 (3.1)	50 (3.9)	48 (4.5)	48 (3.9)	72 (3.5)
Sweden	r 45 (4.8)	r 44 (4.7)	r 20 (3.8)	r 35 (4.3)	r 45 (4.7)	r 29 (3.9)	r 21 (3.6)	r 22 (4.1)	r 24 (3.7)
Syrian Arab Republic	49 (4.3)	49 (4.5)	75 (3.7)	75 (3.6)	86 (3.0)	74 (3.6)	57 (4.5)	23 (3.4)	22 (3.5)
Thailand	72 (3.9)	78 (3.7)	69 (4.1)	68 (4.0)	85 (2.7)	51 (4.0)	61 (4.3)	57 (4.1)	76 (3.4)
Tunisia	39 (3.9)	39 (3.7)	51 (3.8)	59 (3.5)	89 (2.4)	75 (2.9)	44 (4.0)	21 (3.0)	14 (2.6)
Turkey	69 (2.7)	63 (2.9)	65 (3.2)	60 (3.6)	85 (2.4)	81 (2.7)	52 (3.5)	42 (3.1)	48 (3.4)
Ukraine	59 (4.3)	60 (4.0)	84 (3.6)	57 (4.4)	56 (4.1)	36 (4.0)	30 (3.9)	43 (4.2)	22 (3.4)
United Arab Emirates	67 (1.9)	76 (2.0)	83 (1.8)	81 (1.8)	80 (1.8)	56 (2.2)	57 (2.4)	59 (2.1)	48 (2.4)
United States	65 (2.6)	64 (2.2)	64 (2.2)	65 (2.3)	75 (2.2)	46 (2.5)	38 (2.2)	44 (2.5)	36 (2.6)
International Avg.	64 (0.6)	62 (0.5)	62 (0.5)	65 (0.5)	75 (0.5)	54 (0.5)	44 (0.6)	41 (0.6)	40 (0.5)

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.
An "r" indicates data are available for at least 70% but less than 85% of the students.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 6.6: Principals Spend Time on Leadership Activities (Continued)

Country	Percent of Students Whose Principals Spend "A Lot of Time"								
	Promoting the School's Educational Vision or Goals	Developing the School's Curricular and Educational Goals	Monitoring Teachers' Implementation of the School's Educational Goals in Their Teaching	Monitoring Students' Learning Progress to Ensure that the School's Educational Goals Are Reached	Keeping an Orderly Atmosphere in the School	Addressing Disruptive Student Behavior	Advising Teachers Who Have Questions or Problems with Their Teaching	Initiating Educational Projects or Improvements	Participating in Professional Development Activities Specifically for School Principals
Ninth Grade Participants									
Botswana	64 (4.1)	48 (4.4)	56 (4.2)	70 (4.0)	86 (3.2)	71 (4.1)	28 (3.9)	26 (3.7)	33 (4.2)
Honduras	49 (5.2)	53 (4.8)	43 (4.4)	46 (4.6)	86 (3.0)	66 (4.0)	48 (4.5)	35 (4.4)	39 (4.5)
South Africa	60 (3.8)	62 (3.4)	61 (4.0)	69 (3.9)	90 (2.6)	77 (3.5)	51 (3.6)	31 (3.4)	57 (3.9)
Benchmarking Participants									
Alberta, Canada	50 (4.0)	54 (4.0)	33 (3.8)	45 (4.5)	65 (4.4)	40 (4.3)	25 (3.4)	37 (4.9)	27 (4.0)
Ontario, Canada	61 (4.4)	69 (4.0)	49 (4.1)	45 (4.5)	78 (3.6)	44 (4.4)	34 (4.0)	32 (4.0)	38 (3.5)
Quebec, Canada	33 (3.9)	40 (4.0)	22 (2.8)	41 (3.9)	59 (4.3)	66 (4.5)	32 (4.0)	27 (3.5)	12 (2.9)
Abu Dhabi, UAE	70 (3.6)	74 (3.8)	79 (3.5)	78 (4.0)	84 (3.1)	55 (4.3)	62 (4.6)	65 (4.7)	59 (4.0)
Dubai, UAE	68 (0.4)	78 (0.4)	86 (0.2)	80 (0.4)	68 (0.4)	43 (0.5)	40 (0.4)	55 (0.5)	35 (0.5)
Alabama, US	r 53 (9.3)	r 50 (9.1)	r 65 (8.4)	r 73 (6.5)	r 78 (6.2)	r 57 (8.5)	r 25 (6.1)	r 24 (7.6)	r 29 (6.5)
California, US	r 71 (6.0)	r 71 (6.5)	r 76 (6.6)	r 73 (6.7)	r 78 (6.1)	r 52 (7.4)	r 43 (7.5)	r 49 (7.1)	r 45 (6.7)
Colorado, US	72 (7.2)	71 (5.1)	65 (7.1)	59 (6.3)	52 (7.2)	29 (7.7)	41 (6.9)	46 (6.8)	32 (6.9)
Connecticut, US	r 66 (7.7)	65 (6.2)	76 (6.2)	82 (4.8)	77 (5.2)	52 (6.6)	41 (6.6)	47 (8.2)	21 (5.7)
Florida, US	68 (7.0)	67 (7.5)	77 (6.5)	84 (5.6)	85 (5.4)	39 (7.3)	38 (7.6)	52 (7.9)	62 (8.0)
Indiana, US	r 60 (8.6)	r 59 (7.9)	r 61 (8.0)	r 64 (6.9)	r 71 (7.0)	r 33 (7.7)	r 28 (7.1)	r 45 (7.8)	r 22 (5.8)
Massachusetts, US	63 (6.6)	70 (6.5)	68 (7.6)	r 57 (6.5)	52 (7.3)	23 (6.2)	37 (7.5)	40 (7.6)	22 (5.7)
Minnesota, US	66 (7.7)	61 (7.1)	53 (7.8)	57 (7.3)	76 (7.0)	46 (6.0)	25 (6.2)	48 (7.7)	25 (6.1)
North Carolina, US	63 (7.3)	54 (7.9)	60 (7.0)	60 (6.8)	82 (5.6)	46 (7.0)	38 (5.6)	30 (7.0)	39 (6.7)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Principals Spend Time on Leadership Activities

The effectiveness of school leadership has become a central issue, as principals worldwide are held increasingly accountable for their students' achievement outcomes. However, the effects of principal leadership are often indirect and difficult to measure. A meta-analysis of multinational studies conducted between 1986 and 1996 found that "defining and communicating the school's mission" had the largest direct effect on student achievement (Witziers, Bosker, & Kruger, 2003), whereas a different meta-analysis of 27 studies conducted between 1978 and 2006 found strong effects for promoting teacher learning and development, and establishing goals (Robinson, Lloyd, & Rowe, 2008).

TIMSS 2011 used research conducted in the Netherlands (ten Bruggencate, Luyten, Scheerens, & Sleegers, 2012) to develop questions about principals' leadership styles. These questions were included in both the fourth and eighth grade assessments. Exhibit 6.5 presents principals' reports for the fourth grade about the various activities upon which they spend "a lot of time." The pattern of varying reports from country to country held for the fourth grade, the sixth grade, and the benchmarking participants.

The results for the fourth grade were averaged across countries to provide some summary data. The first two questions related to defining and communicating the school's mission, and on average, more than half of the fourth grade students (59% and 60%), were in schools where this occupied "a lot" of the principal's time. The next two questions addressed monitoring whether goals are achieved by teachers and students, with just over half the students (53% and 57%) in schools where principals reported spending "a lot of time" on these activities. The next two categories asked about maintaining discipline: two-thirds of students were in schools where the principal spent "a lot of time" keeping an orderly atmosphere, and 44 percent had principals that needed to spend "a lot of time" addressing disruptive student behavior. The last three areas appear to occupy less time: advising teachers, initiating projects, and participating in professional development activities.

Exhibit 6.6 summarizes principals' reports from the eighth grade assessment about time spent on leadership activities. About two-thirds of the eighth grade students were in schools where the principal reported spending "a lot of time" on defining and communicating the school's mission and in monitoring whether goals were being achieved by teachers and students. Three-fourths of the eighth grade students were in schools where the principal devoted "a lot of time" to keeping an orderly atmosphere, and more than half

had principals that needed to spend "a lot of time" addressing disruptive student behavior. Similar to the fourth grade, the last three areas—advising teachers, initiating projects, and participating in professional development activities—appear to occupy less of the principal's time.

Schools with Discipline and Safety Problems

The sense of security that comes from attending a school with few behavior problems and having little or no concern about student or teacher safety promotes a stable learning environment. There is increasing research showing that a safe school environment is important for students' academic achievement. On the other hand, a general lack of discipline, especially if students and teachers are afraid for their safety, does not facilitate learning. Unfortunately, community and school violence are becoming an increasing problem, especially among urban youth.

Safe and Orderly School

There is growing evidence that students' perceived school safety adversely affects academic performance, even for primary school children (Milam, Furr-Holden, & Leaf, 2010). It seems that safety at school can no longer be taken for granted, even at the fourth grade. To provide information on the extent to which school safety might be affecting mathematics achievement, TIMSS 2011 developed the Safe and Orderly School scale. Teachers in both the fourth and eighth grade assessments were asked the degree to which they agreed or disagreed with five statements:

- This school is located in a safe neighborhood;
- I feel safe at this school;
- This school's security policies and practices are sufficient;
- The students behave in an orderly manner; and
- The students are respectful of the teachers.

Exhibit 6.7 presents the results for the Safe and Orderly School scale for the fourth grade assessment. Students were scored according to their teachers' degree of agreement with the five statements. Students in **Safe and Orderly** schools had teachers that "agreed a lot" with three of the five qualities and "agreed a little" with the other two, on average. There was substantial variation across countries, but internationally, on average, across the fourth grade countries, the majority of students (53%) were attending schools judged by

their teachers to be **Safe and Orderly**. Almost all of the remaining students (43%) were in schools judged to be **Somewhat Safe and Orderly**. In general, only small percentages of students (4% on average) were in schools judged **Not Safe and Orderly** at best, their teachers "disagreed a little" with three of the five statements and "agreed a little" with the other two, on average. Across the fourth grade countries, on average, the safer the school as reported by their teachers, the higher the students' average mathematics achievement.

Exhibit 6.8 presents the corresponding Safe and Orderly School scale results for the eighth grade assessment. Students were assigned to one of the three school orderliness categories using the same criteria as at the fourth grade, and with broadly similar results. Although almost all of the eighth grade students, on average internationally, were in **Safe and Orderly** or **Somewhat Safe and Orderly** schools, the eighth grade mathematics teachers were noticeably less positive in their reports. On average, across the eighth-grade countries, 45 percent of students (compared to 53% at the fourth grade) were attending schools judged by their teachers to be safe and orderly, 49 percent of students (compared to 43%) were in schools judged to be **Somewhat Safe and Orderly**, and 6 percent of students (compared to 4%) were in schools judged **Not Safe and Orderly**. The average mathematics achievement gap between students in the **Safe and Orderly** and **Not Safe and Orderly** schools also was greater at the eighth grade (34 points vs. 28 points).

School Discipline and Safety

Previous TIMSS assessments have asked principals for their perceptions about the degree to which a series of discipline, disorderly, and bullying behaviors are problems in their schools, and found that having fewer problems was related to higher average achievement. Exhibit 6.9 presents the TIMSS 2011 results for the fourth grade School Discipline and Safety scale based on asking principals about the extent of ten different discipline and school safety problems (see the second page of the exhibit for the complete list of problems). Countries are ordered by the percentage of students whose principals reported few student discipline and school safety problems. Principals in schools with **Hardly Any Problems** with discipline or safety reported "not a problem" for five of the ten discipline and safety issues and only "minor problem" for the other five, on average. Principals in schools with **Moderate Problems** reported "moderate problem" for five of the ten issues and "minor problem" for the other five, on average.

More than half of the students (61%), on average, across the fourth grade countries were in the **Hardly Any Problems** category and 29 percent were in the **Minor Problems** category. Only 11 percent, on average, attended schools where principals reported **Moderate Problems** with discipline and school safety. Students whose principals reported **Moderate Problems** in their schools had substantially lower mathematics achievement, by 45 points on average, than students whose principals reported **Hardly Any Problems** (451 vs. 496).The results for the sixth grade and benchmarking participants followed a similar pattern.

Exhibit 6.10 presents the results for the School Discipline and Safety scale for the TIMSS 2011 eighth grade assessment. This scale is based on eleven discipline and school safety problems, ten of which comprised the fourth grade scale plus one additional problem more suited to older students—"Physical injury to teachers or staff" (see the second page of the exhibit for the complete list of problems). Compared to the situation at the fourth grade, relatively speaking, there were fewer eighth grade students in the **Hardly Any Problems** category (38% vs. 61%) and more in the **Minor Problems** category (49% vs. 29%). There were similar percentages of students in schools with **Moderate Problems** at the fourth and eighth grades (11% and 13%). Looking more closely at the problems comprising the scales, the increase from fourth to eighth grade in the percentage of students in schools with discipline and safety problems is largely because eight of these problems (classroom disturbance, cheating, profanity, vandalism, theft, intimidation or verbal abuse among students, students fighting, and intimidation or verbal abuse of teachers) often were "not a problem" for fourth grade principals but more often were a "minor problem" for principals of eighth grade schools.

Exhibit 6.7: Safe and Orderly School

TIMSS 2011 Mathematics 4th Grade

Reported by Teachers

Students were scored according to their teachers' degree of agreement with five statements on the *Safe and Orderly School* scale. Students in **Safe and Orderly** schools had a score on the scale of at least 10.2, which corresponds to their teachers "agreeing a lot" with three of the five qualities of a safe and orderly school and "agreeing a little" with the other two, on average. Students in **Not Safe and Orderly** schools had a score no higher than 6.3, which corresponds to their teachers "disagreeing a little" with three of the five qualities and "agreeing a little" with the other two, on average. All other students attended **Somewhat Safe and Orderly** schools.

Country	Safe and Orderly		Somewhat Safe and Orderly		Not Safe and Orderly		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Northern Ireland r	85 (2.7)	568 (4.0)	15 (2.6)	537 (8.6)	0 (0.4)	~ ~	11.5 (0.14)
Georgia	83 (2.5)	453 (3.9)	16 (2.4)	442 (10.4)	1 (0.7)	~ ~	11.3 (0.12)
Azerbaijan	83 (2.9)	465 (6.5)	16 (2.8)	459 (16.7)	1 (0.7)	~ ~	11.4 (0.13)
Ireland	78 (3.3)	537 (3.0)	20 (3.3)	497 (6.0)	2 (1.0)	~ ~	11.3 (0.15)
Australia r	76 (3.1)	529 (3.7)	20 (3.0)	491 (7.9)	4 (1.4)	460 (12.4)	11.1 (0.16)
United Arab Emirates	76 (2.2)	440 (3.0)	24 (2.2)	418 (5.7)	0 (0.2)	~ ~	10.8 (0.08)
Croatia	73 (3.1)	489 (2.2)	26 (3.0)	495 (4.2)	1 (0.7)	~ ~	10.8 (0.12)
Thailand	72 (3.9)	462 (4.5)	26 (3.8)	462 (10.1)	3 (1.8)	352 (15.0)	11.0 (0.18)
Armenia	72 (2.7)	455 (4.2)	26 (2.6)	447 (6.6)	2 (1.1)	~ ~	10.9 (0.13)
Kuwait	70 (3.1)	346 (3.9)	30 (3.1)	331 (6.3)	0 (0.0)	~ ~	10.4 (0.10)
New Zealand	70 (2.3)	501 (2.9)	29 (2.3)	456 (4.8)	1 (0.5)	~ ~	11.0 (0.10)
Denmark	68 (3.5)	544 (2.7)	32 (3.5)	534 (4.6)	0 (0.0)	~ ~	10.6 (0.12)
Kazakhstan	67 (4.0)	505 (5.8)	33 (4.0)	495 (9.2)	1 (0.4)	~ ~	10.7 (0.15)
England	67 (4.3)	557 (3.8)	31 (4.1)	519 (7.9)	2 (1.3)	~ ~	10.7 (0.18)
United States	66 (2.4)	553 (2.3)	30 (2.3)	526 (3.4)	4 (0.8)	503 (8.4)	10.5 (0.09)
Qatar	65 (3.6)	421 (6.1)	34 (3.7)	393 (8.1)	1 (0.0)	~ ~	10.5 (0.11)
Norway	64 (4.6)	501 (3.5)	36 (4.6)	484 (4.6)	0 (0.0)	~ ~	10.7 (0.17)
Saudi Arabia	62 (4.4)	425 (7.2)	36 (4.4)	389 (7.2)	2 (0.9)	~ ~	10.4 (0.16)
Singapore	61 (2.5)	613 (3.8)	37 (2.5)	595 (5.6)	2 (0.7)	~ ~	10.3 (0.10)
Iran, Islamic Rep. of	60 (3.5)	440 (4.2)	39 (3.4)	419 (6.1)	1 (0.8)	~ ~	10.3 (0.15)
Bahrain	57 (4.2)	446 (4.0)	42 (4.3)	423 (4.9)	1 (0.0)	~ ~	10.3 (0.17)
Austria	57 (3.4)	513 (3.0)	40 (3.5)	504 (3.3)	2 (1.5)	~ ~	10.0 (0.13)
Netherlands r	56 (4.6)	541 (2.6)	43 (4.6)	536 (3.8)	1 (0.8)	~ ~	10.2 (0.18)
Poland	55 (3.4)	478 (2.8)	44 (3.4)	485 (3.3)	1 (0.6)	~ ~	10.0 (0.12)
Hong Kong SAR	55 (4.7)	603 (4.6)	44 (4.8)	602 (6.0)	1 (0.6)	~ ~	10.2 (0.17)
Hungary	52 (3.8)	525 (4.9)	46 (3.6)	506 (5.6)	3 (1.3)	452 (24.4)	9.7 (0.14)
Spain	51 (3.8)	497 (3.2)	45 (3.9)	470 (4.4)	5 (1.8)	449 (14.4)	9.7 (0.16)
Russian Federation	49 (4.0)	546 (5.0)	48 (3.8)	539 (5.4)	2 (1.3)	~ ~	9.9 (0.17)
Malta	49 (0.1)	503 (1.8)	46 (0.1)	488 (2.1)	5 (0.1)	500 (5.9)	9.9 (0.01)
Lithuania	47 (3.2)	538 (3.7)	51 (3.1)	530 (3.2)	2 (0.9)	~ ~	9.7 (0.12)
Germany	47 (3.8)	533 (3.0)	52 (3.7)	525 (3.1)	2 (0.9)	~ ~	9.8 (0.13)
Portugal	46 (5.1)	541 (6.9)	50 (4.9)	527 (4.6)	4 (1.3)	507 (12.7)	9.6 (0.20)
Belgium (Flemish)	46 (3.0)	555 (2.6)	52 (2.9)	545 (2.3)	1 (0.8)	~ ~	9.7 (0.11)
Oman	46 (2.6)	400 (3.7)	52 (2.7)	374 (4.1)	2 (0.9)	~ ~	9.8 (0.09)
Yemen	46 (4.4)	257 (8.4)	52 (4.5)	235 (7.9)	2 (0.9)	~ ~	9.9 (0.15)
Czech Republic	45 (3.8)	512 (3.7)	53 (3.6)	510 (3.5)	2 (0.9)	~ ~	9.6 (0.12)
Sweden r	41 (4.8)	516 (3.4)	54 (4.9)	501 (3.2)	5 (1.3)	453 (3.6)	9.6 (0.16)
Chile	41 (3.7)	484 (4.6)	46 (3.7)	451 (4.2)	13 (3.1)	430 (13.1)	9.2 (0.19)
Slovak Republic	40 (3.6)	509 (5.9)	58 (3.6)	506 (4.8)	1 (0.7)	~ ~	9.4 (0.09)
Serbia	40 (4.2)	515 (4.8)	55 (4.1)	520 (3.9)	5 (1.6)	478 (20.5)	9.4 (0.16)
Romania	40 (3.6)	480 (9.7)	55 (3.7)	483 (7.4)	5 (1.6)	459 (17.9)	9.5 (0.14)
Tunisia	40 (3.9)	367 (6.9)	51 (3.8)	355 (4.8)	10 (2.6)	347 (17.0)	9.3 (0.16)
Turkey	37 (3.3)	495 (4.8)	45 (3.1)	461 (6.8)	18 (2.7)	438 (15.9)	8.9 (0.17)
Finland	36 (3.5)	554 (3.5)	59 (4.0)	544 (2.7)	6 (1.7)	519 (8.8)	9.4 (0.12)
Chinese Taipei	31 (3.8)	590 (2.4)	62 (3.7)	594 (2.7)	7 (2.0)	575 (5.2)	9.0 (0.15)
Morocco	29 (3.7)	363 (8.8)	53 (4.4)	331 (7.0)	17 (3.0)	321 (11.7)	8.8 (0.18)
Slovenia	27 (3.1)	511 (3.6)	67 (3.2)	515 (2.8)	6 (1.6)	498 (9.0)	8.9 (0.11)
Korea, Rep. of	24 (3.7)	615 (5.0)	69 (3.8)	603 (2.2)	7 (2.2)	593 (4.5)	8.7 (0.18)
Italy	18 (2.6)	508 (5.6)	75 (2.8)	511 (3.4)	6 (2.0)	487 (12.1)	8.6 (0.12)
Japan	5 (1.7)	589 (5.7)	83 (3.1)	587 (1.9)	12 (2.6)	574 (5.6)	7.9 (0.09)
International Avg.	53 (0.5)	498 (0.7)	43 (0.5)	483 (0.8)	4 (0.2)	470 (2.9)	

Centerpoint of scale set at 10.

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

A tilde (~) indicates insufficient data to report achievement.

An "r" indicates data are available for at least 70% but less than 85% of the students.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 6.7: Safe and Orderly School (Continued)

Country		Safe and Orderly		Somewhat Safe and Orderly		Not Safe and Orderly		Average Scale Score
		Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Sixth Grade Participants								
Honduras		62 (4.4)	392 (7.8)	33 (4.2)	404 (8.3)	5 (1.7)	393 (22.7)	10.5 (0.19)
Yemen		48 (3.6)	346 (8.5)	49 (3.6)	348 (7.7)	3 (1.4)	360 (25.0)	9.7 (0.13)
Botswana		22 (3.8)	455 (10.3)	56 (4.4)	412 (5.9)	22 (3.5)	405 (6.1)	8.2 (0.18)
Benchmarking Participants								
Dubai, UAE	r	84 (1.9)	474 (3.1)	15 (1.9)	453 (10.1)	0 (0.3)	~ ~	11.4 (0.09)
Alberta, Canada	r	80 (3.5)	510 (3.1)	19 (3.6)	497 (5.9)	1 (0.8)	~ ~	11.3 (0.16)
Abu Dhabi, UAE		78 (4.0)	422 (5.4)	21 (3.9)	404 (11.7)	0 (0.5)	~ ~	10.8 (0.14)
Florida, US	r	65 (4.4)	553 (4.3)	28 (4.3)	527 (4.6)	7 (2.7)	523 (23.8)	10.4 (0.25)
Ontario, Canada		62 (3.9)	526 (3.0)	35 (3.9)	506 (5.6)	3 (0.9)	513 (14.1)	10.5 (0.16)
North Carolina, US		59 (6.5)	564 (4.6)	34 (5.7)	537 (7.8)	7 (3.5)	530 (21.3)	10.2 (0.28)
Quebec, Canada		45 (4.5)	533 (2.9)	50 (4.4)	533 (3.5)	5 (1.9)	519 (9.3)	9.8 (0.17)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Thinking about your current school, indicate the extent to which you agree or disagree with each of the following statements.

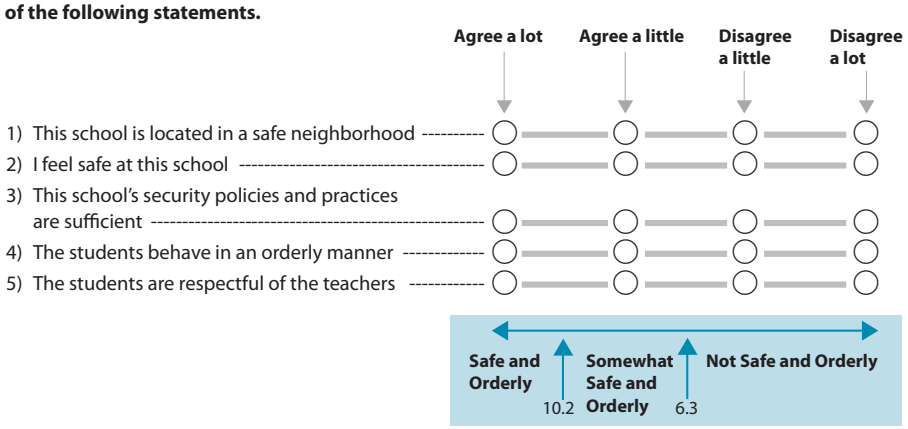


Exhibit 6.8 Safe and Orderly School

TIMSS 2011 8th Grade Mathematics

Reported by Teachers

Students were scored according to their teachers' degree of agreement with five statements on the *Safe and Orderly School* scale. Students in **Safe and Orderly** schools had a score on the scale of at least 10.7, which corresponds to their teachers "agreeing a lot" with three of the five qualities of a safe and orderly school and "agreeing a little" with the other two, on average. Students in **Not Safe and Orderly** schools had a score no higher than 6.8, which corresponds to their teachers "disagreeing a little" with three of the five qualities and "agreeing a little" with the other two, on average. All other students attended **Somewhat Safe and Orderly** schools.

Country	Safe and Orderly		Somewhat Safe and Orderly		Not Safe and Orderly		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Georgia	73 (3.2)	435 (4.6)	26 (3.1)	417 (7.8)	1 (0.6)	~ ~	11.2 (0.12)
Qatar	68 (3.2)	421 (5.9)	29 (3.0)	384 (9.8)	3 (1.1)	396 (25.6)	11.0 (0.11)
United Arab Emirates	68 (2.2)	465 (3.0)	31 (2.2)	435 (4.4)	1 (0.2)	~ ~	10.9 (0.07)
Ukraine	66 (3.9)	477 (5.0)	34 (3.9)	484 (7.1)	0 (0.0)	~ ~	10.7 (0.10)
Kazakhstan	65 (4.1)	489 (5.1)	34 (4.1)	483 (7.6)	1 (0.7)	~ ~	11.0 (0.14)
Israel	64 (2.9)	532 (5.5)	32 (2.9)	496 (8.6)	3 (1.4)	488 (31.6)	10.8 (0.13)
Armenia	63 (3.7)	471 (3.9)	35 (3.5)	457 (4.9)	2 (0.8)	~ ~	10.9 (0.14)
Norway	62 (4.4)	479 (3.2)	38 (4.4)	470 (3.8)	0 (0.0)	~ ~	10.9 (0.14)
Syrian Arab Republic	60 (4.3)	386 (5.5)	38 (4.4)	366 (7.6)	2 (1.3)	~ ~	10.6 (0.15)
Singapore	58 (2.4)	623 (5.1)	39 (2.4)	596 (5.8)	2 (0.7)	~ ~	10.7 (0.10)
Iran, Islamic Rep. of	55 (3.5)	424 (6.8)	42 (3.5)	406 (5.7)	3 (1.0)	377 (14.0)	10.6 (0.12)
Australia r	55 (4.2)	530 (8.3)	36 (3.9)	482 (7.0)	9 (2.3)	465 (17.0)	10.5 (0.20)
New Zealand	55 (3.3)	495 (6.9)	40 (3.5)	475 (10.2)	5 (1.8)	486 (16.8)	10.5 (0.15)
Thailand	54 (3.6)	436 (6.7)	41 (3.8)	415 (8.3)	4 (1.7)	432 (16.0)	10.4 (0.14)
United States r	54 (2.5)	526 (4.3)	38 (2.1)	494 (4.6)	8 (1.7)	500 (13.2)	10.4 (0.13)
Romania	54 (4.2)	463 (6.8)	45 (4.1)	455 (6.8)	1 (0.6)	~ ~	10.5 (0.15)
Hong Kong SAR	54 (4.7)	599 (6.8)	45 (4.7)	564 (8.2)	1 (0.0)	~ ~	10.5 (0.16)
Macedonia, Rep. of r	53 (3.7)	441 (8.0)	44 (3.9)	402 (8.7)	3 (1.4)	436 (33.5)	10.5 (0.16)
England	53 (4.5)	521 (7.2)	42 (4.2)	487 (10.3)	6 (1.9)	505 (19.1)	10.6 (0.19)
Saudi Arabia	51 (3.8)	405 (6.4)	46 (4.0)	386 (6.3)	2 (1.2)	~ ~	10.3 (0.14)
Bahrain	49 (3.1)	429 (4.4)	47 (3.3)	396 (4.3)	4 (1.2)	345 (4.8)	10.2 (0.11)
Hungary	48 (3.5)	515 (4.7)	47 (3.6)	501 (5.8)	5 (1.7)	439 (18.8)	9.9 (0.11)
Malaysia	44 (4.3)	459 (8.6)	53 (3.9)	425 (6.3)	3 (1.4)	429 (19.2)	10.2 (0.17)
Russian Federation	42 (3.6)	547 (5.1)	56 (3.6)	533 (4.6)	2 (0.9)	~ ~	10.0 (0.14)
Lithuania	40 (3.7)	504 (6.0)	59 (3.7)	501 (3.5)	1 (0.4)	~ ~	9.9 (0.10)
Lebanon	39 (4.1)	466 (6.1)	53 (4.3)	443 (5.2)	8 (2.6)	411 (12.7)	9.8 (0.19)
Turkey	38 (3.2)	483 (8.3)	49 (3.3)	441 (5.8)	13 (2.1)	407 (7.6)	9.3 (0.12)
Indonesia	37 (4.1)	387 (6.3)	61 (4.2)	386 (6.2)	2 (1.0)	~ ~	10.0 (0.16)
Oman	37 (3.0)	384 (4.9)	61 (3.0)	357 (4.1)	2 (1.1)	~ ~	9.9 (0.12)
Ghana	36 (3.6)	355 (8.6)	55 (3.8)	316 (5.3)	9 (2.0)	320 (12.7)	9.6 (0.17)
Palestinian Nat'l Auth.	36 (4.4)	403 (6.1)	54 (4.2)	407 (5.3)	10 (2.6)	385 (14.5)	9.5 (0.18)
Jordan	36 (3.6)	418 (5.6)	59 (3.7)	403 (5.7)	5 (1.5)	355 (21.9)	9.6 (0.13)
Chile	34 (3.4)	447 (6.5)	51 (4.1)	408 (4.0)	15 (3.1)	376 (6.9)	9.4 (0.18)
Finland	31 (3.4)	519 (4.4)	63 (3.6)	512 (2.6)	6 (1.6)	508 (9.3)	9.4 (0.11)
Chinese Taipei	31 (3.7)	627 (6.7)	57 (3.8)	603 (5.0)	12 (2.7)	593 (10.9)	9.1 (0.15)
Sweden r	31 (3.3)	495 (4.1)	67 (3.2)	483 (2.7)	3 (0.8)	446 (13.9)	9.5 (0.12)
Morocco	26 (2.3)	399 (6.0)	59 (3.3)	364 (2.8)	16 (2.4)	355 (5.5)	9.0 (0.10)
Tunisia	22 (3.1)	419 (6.4)	61 (3.4)	427 (4.5)	17 (2.8)	424 (6.1)	8.8 (0.15)
Slovenia	19 (2.4)	511 (4.9)	75 (2.5)	503 (2.6)	7 (1.6)	502 (9.3)	9.0 (0.10)
Italy	17 (2.9)	509 (4.2)	76 (3.1)	499 (3.3)	8 (2.1)	474 (11.1)	8.9 (0.12)
Japan	14 (3.0)	593 (10.5)	71 (3.6)	567 (3.1)	15 (2.6)	560 (5.1)	8.5 (0.13)
Korea, Rep. of	13 (2.4)	624 (8.3)	74 (2.9)	611 (3.6)	13 (2.3)	607 (8.5)	8.5 (0.11)
International Avg.	45 (0.5)	479 (1.0)	49 (0.6)	458 (0.9)	6 (0.3)	445 (3.1)	

Centerpoint of scale set at 10.

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

A tilde (~) indicates insufficient data to report achievement.

An "r" indicates data are available for at least 70% but less than 85% of the students.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 6.8: Safe and Orderly School (Continued)

Country	Safe and Orderly		Somewhat Safe and Orderly		Not Safe and Orderly		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Ninth Grade Participants							
Honduras	36 (4.2)	346 (9.0)	51 (4.5)	331 (5.1)	13 (2.6)	339 (8.9)	9.6 (0.20)
South Africa	21 (2.8)	379 (11.0)	55 (3.7)	347 (4.2)	24 (3.0)	341 (5.3)	8.5 (0.15)
Botswana	12 (3.0)	414 (10.4)	55 (4.2)	398 (3.5)	33 (3.9)	390 (3.8)	8.0 (0.17)
Benchmarking Participants							
Dubai, UAE	80 (2.3)	483 (3.3)	18 (2.2)	453 (9.4)	2 (0.4)	~ ~	11.3 (0.09)
Minnesota, US	69 (5.2)	552 (6.0)	31 (5.2)	529 (11.2)	0 (0.0)	~ ~	11.4 (0.20)
Alberta, Canada	68 (3.9)	508 (3.5)	27 (3.6)	496 (4.1)	5 (1.8)	501 (9.7)	11.1 (0.18)
Massachusetts, US	65 (4.7)	573 (6.9)	27 (3.9)	546 (13.3)	7 (2.8)	490 (13.9)	11.0 (0.24)
Colorado, US r	62 (5.6)	538 (5.0)	33 (5.6)	490 (12.3)	4 (1.7)	460 (37.2)	11.0 (0.22)
Abu Dhabi, UAE	62 (3.7)	457 (5.9)	38 (3.7)	439 (5.7)	0 (0.3)	~ ~	10.7 (0.13)
North Carolina, US r	58 (7.0)	539 (8.4)	36 (6.5)	535 (9.4)	6 (3.2)	605 (28.6)	10.6 (0.29)
Indiana, US r	55 (7.5)	526 (7.6)	43 (7.6)	511 (9.3)	2 (1.4)	~ ~	10.6 (0.25)
Connecticut, US	54 (5.2)	549 (7.3)	39 (5.6)	483 (9.8)	7 (3.7)	504 (54.7)	10.5 (0.25)
Ontario, Canada	52 (4.4)	524 (3.2)	44 (4.5)	501 (3.8)	4 (1.5)	492 (5.7)	10.6 (0.19)
California, US r	49 (5.7)	512 (6.8)	37 (5.5)	476 (12.9)	14 (4.6)	466 (19.6)	10.1 (0.30)
Alabama, US r	44 (7.3)	492 (11.3)	47 (7.4)	451 (8.3)	9 (3.6)	435 (31.4)	9.8 (0.29)
Quebec, Canada	38 (3.8)	545 (5.1)	59 (3.7)	525 (3.1)	3 (1.2)	526 (17.6)	9.9 (0.15)
Florida, US r	37 (6.4)	533 (10.4)	52 (6.2)	511 (9.7)	11 (4.4)	531 (37.5)	9.5 (0.24)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Thinking about your current school, indicate the extent to which you agree or disagree with each of the following statements.

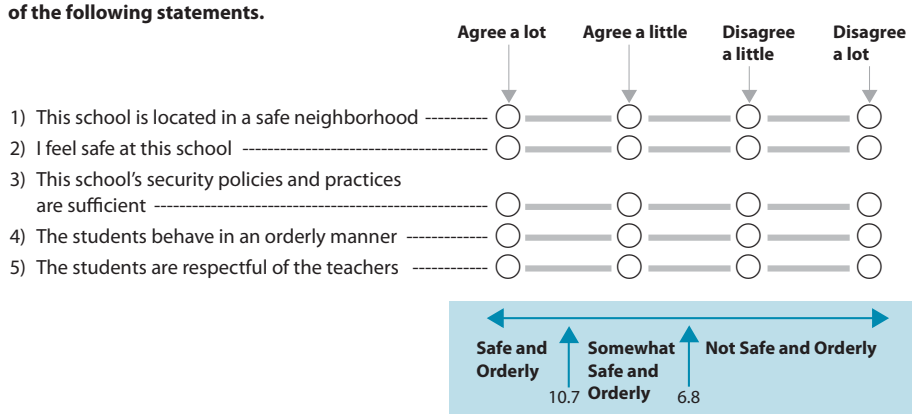


Exhibit 6.9: School Discipline and Safety

Reported by Principals

Students were scored according to their principals' responses concerning ten potential school problems on the *School Discipline and Safety* scale. Students in schools with **Hardly Any Problems** had a score on the scale of at least 9.7, which corresponds to their principals reporting "not a problem" for five of the ten discipline and safety issues and "minor problem" for the other five, on average. Students in schools with **Moderate Problems** had a score no higher than 7.6, which corresponds to their principals reporting "moderate problem" for five of the ten issues and "minor problem" for the other five, on average. All other students attended schools with **Minor Problems**.

Country	Hardly Any Problems		Minor Problems		Moderate Problems		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Kazakhstan	91 (2.2)	505 (5.0)	9 (2.4)	465 (13.3)	1 (0.6)	~ ~	11.1 (0.10)
Armenia	87 (2.7)	450 (3.8)	8 (2.3)	460 (11.8)	4 (1.7)	479 (20.6)	11.1 (0.12)
Northern Ireland	85 (3.7)	566 (3.8)	15 (3.7)	542 (7.7)	0 (0.0)	~ ~	11.0 (0.13)
Netherlands r	85 (3.6)	544 (2.2)	15 (3.6)	524 (6.9)	0 (0.0)	~ ~	11.3 (0.16)
Hong Kong SAR	84 (2.9)	606 (3.0)	15 (2.8)	574 (16.0)	1 (0.0)	~ ~	11.2 (0.12)
Ireland	83 (3.1)	532 (2.9)	16 (3.0)	512 (9.9)	1 (1.0)	~ ~	11.1 (0.13)
Georgia	81 (2.8)	449 (4.7)	13 (2.4)	447 (9.8)	6 (1.4)	471 (14.3)	10.7 (0.15)
Spain	80 (3.3)	487 (2.7)	12 (2.8)	459 (10.1)	8 (2.3)	481 (14.2)	10.7 (0.17)
Chinese Taipei	77 (3.3)	591 (2.5)	23 (3.3)	591 (4.2)	0 (0.0)	~ ~	11.4 (0.13)
England	77 (4.1)	551 (4.2)	20 (4.2)	515 (11.0)	3 (1.6)	495 (10.9)	10.6 (0.11)
Korea, Rep. of	76 (3.6)	606 (2.3)	18 (3.4)	599 (3.9)	6 (2.0)	596 (7.5)	10.9 (0.15)
Lithuania	75 (3.5)	538 (2.8)	25 (3.5)	523 (5.8)	0 (0.0)	~ ~	10.5 (0.11)
Iran, Islamic Rep. of	74 (3.9)	437 (4.6)	25 (3.9)	417 (7.8)	0 (0.0)	~ ~	10.7 (0.11)
Japan	72 (3.2)	585 (1.9)	24 (3.3)	587 (4.8)	4 (1.6)	582 (10.4)	10.5 (0.12)
New Zealand	69 (3.4)	502 (3.3)	28 (3.2)	458 (5.5)	3 (1.3)	419 (15.2)	10.7 (0.12)
Czech Republic	68 (3.6)	512 (3.0)	29 (3.5)	506 (5.1)	2 (1.0)	~ ~	10.2 (0.11)
Belgium (Flemish)	67 (4.4)	553 (2.2)	32 (4.3)	545 (3.9)	1 (0.0)	~ ~	10.4 (0.13)
Singapore	67 (0.0)	606 (3.9)	33 (0.0)	603 (6.0)	0 (0.0)	~ ~	10.7 (0.00)
Croatia	66 (4.0)	492 (2.6)	31 (4.0)	484 (3.8)	2 (1.2)	~ ~	10.4 (0.12)
Portugal	66 (5.4)	536 (4.1)	30 (5.5)	525 (7.9)	5 (1.7)	529 (18.7)	10.3 (0.17)
Russian Federation	65 (3.9)	545 (4.5)	35 (3.8)	536 (5.4)	0 (0.5)	~ ~	10.1 (0.09)
United States	64 (2.7)	551 (3.0)	34 (2.6)	531 (3.3)	2 (0.7)	~ ~	10.3 (0.09)
Australia	64 (3.9)	523 (4.1)	34 (3.8)	511 (5.3)	2 (1.0)	~ ~	10.4 (0.12)
Finland	64 (4.5)	549 (2.5)	34 (4.4)	540 (4.8)	2 (1.2)	~ ~	10.2 (0.12)
Romania	64 (4.1)	495 (5.6)	23 (3.4)	478 (12.3)	13 (2.9)	430 (27.6)	10.2 (0.17)
Malta	64 (0.1)	503 (1.8)	30 (0.1)	486 (2.4)	6 (0.1)	473 (4.9)	10.1 (0.00)
Bahrain	63 (4.2)	438 (4.8)	25 (4.1)	430 (9.2)	12 (4.7)	437 (7.4)	10.1 (0.30)
Qatar	63 (3.2)	430 (5.1)	23 (2.6)	391 (10.1)	14 (2.3)	373 (10.2)	9.9 (0.14)
Azerbaijan	62 (4.2)	461 (7.6)	8 (2.3)	462 (13.8)	30 (3.9)	466 (9.3)	9.5 (0.26)
United Arab Emirates	61 (2.3)	444 (2.9)	24 (2.0)	411 (4.6)	15 (1.7)	415 (6.8)	9.9 (0.11)
Denmark r	60 (4.0)	543 (3.4)	40 (4.0)	535 (4.1)	1 (0.0)	~ ~	10.0 (0.09)
Norway	58 (4.4)	495 (3.7)	39 (4.2)	492 (4.0)	3 (1.6)	485 (10.1)	9.9 (0.13)
Thailand	58 (4.6)	469 (4.8)	36 (4.4)	444 (9.0)	6 (2.3)	442 (21.5)	10.1 (0.16)
Slovak Republic	57 (3.6)	513 (3.7)	35 (3.4)	503 (7.5)	9 (2.0)	477 (16.9)	9.9 (0.12)
Italy	56 (3.9)	509 (3.8)	25 (3.8)	509 (5.9)	19 (2.9)	505 (6.3)	9.5 (0.14)
Serbia	55 (4.7)	514 (4.8)	30 (4.2)	524 (5.8)	15 (3.2)	506 (6.9)	9.7 (0.18)
Slovenia	53 (3.7)	512 (3.4)	42 (3.6)	516 (3.6)	4 (1.4)	500 (5.6)	10.0 (0.12)
Poland	51 (3.9)	481 (3.0)	46 (4.2)	481 (3.2)	3 (1.4)	493 (14.4)	9.7 (0.09)
Hungary	50 (4.2)	530 (4.8)	45 (4.2)	509 (6.0)	5 (1.5)	433 (24.6)	9.7 (0.13)
Sweden	49 (4.7)	514 (2.8)	45 (4.7)	495 (3.7)	6 (1.2)	479 (12.7)	9.7 (0.13)
Austria	46 (4.3)	513 (3.4)	42 (4.1)	508 (3.7)	12 (3.3)	492 (9.1)	9.4 (0.14)
Saudi Arabia	45 (3.9)	417 (6.2)	25 (3.8)	395 (13.8)	30 (3.8)	414 (9.8)	9.1 (0.18)
Germany	41 (3.3)	539 (3.1)	53 (3.5)	526 (3.0)	6 (1.5)	487 (7.8)	9.5 (0.08)
Chile	39 (3.4)	481 (5.0)	43 (4.1)	459 (4.6)	18 (2.9)	439 (6.4)	9.2 (0.14)
Turkey	38 (2.9)	491 (6.8)	35 (3.4)	464 (7.2)	26 (3.4)	445 (12.0)	8.9 (0.14)
Oman	28 (2.9)	385 (4.8)	37 (3.1)	374 (4.6)	35 (3.0)	380 (6.2)	8.4 (0.15)
Tunisia	26 (3.3)	362 (7.1)	27 (3.2)	357 (7.9)	46 (4.0)	359 (6.2)	8.0 (0.19)
Kuwait	24 (3.5)	348 (6.8)	48 (4.2)	345 (5.0)	29 (3.6)	332 (7.3)	8.4 (0.15)
Morocco	14 (2.4)	340 (9.1)	24 (3.1)	317 (7.6)	62 (3.9)	342 (6.1)	7.2 (0.15)
Yemen	13 (2.8)	263 (12.4)	33 (4.1)	259 (10.5)	54 (4.0)	238 (9.7)	7.5 (0.16)
International Avg.	61 (0.5)	496 (0.7)	29 (0.5)	482 (1.1)	11 (0.3)	451 (2.2)	

Centerpoint of scale set at 10.

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

A tilde (~) indicates insufficient data to report achievement.

An "r" indicates data are available for at least 70% but less than 85% of the students.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 6.9: School Discipline and Safety (Continued)

Country	Hardly Any Problems		Minor Problems		Moderate Problems		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Sixth Grade Participants							
Honduras	44 (4.5)	403 (9.9)	37 (4.9)	395 (10.2)	19 (3.3)	382 (8.1)	9.1 (0.17)
Botswana	27 (3.9)	443 (10.9)	58 (4.2)	416 (4.4)	14 (2.9)	385 (8.2)	9.0 (0.12)
Yemen	13 (3.0)	372 (14.2)	34 (4.3)	345 (7.7)	53 (4.0)	341 (9.1)	7.5 (0.15)
Benchmarking Participants							
Dubai, UAE	74 (0.4)	481 (1.9)	17 (0.4)	420 (5.1)	10 (0.1)	443 (4.1)	10.6 (0.01)
Alberta, Canada	68 (4.3)	511 (3.2)	32 (4.3)	500 (3.7)	0 (0.0)	~ ~	10.5 (0.13)
Ontario, Canada	66 (4.5)	522 (3.5)	33 (4.6)	512 (5.1)	1 (0.9)	~ ~	10.4 (0.13)
Abu Dhabi, UAE	63 (4.2)	427 (5.9)	25 (4.0)	392 (8.0)	12 (2.8)	386 (10.7)	9.9 (0.18)
Florida, US r	60 (6.5)	552 (5.7)	40 (6.5)	533 (4.0)	0 (0.0)	~ ~	10.3 (0.21)
North Carolina, US	59 (7.5)	564 (5.6)	41 (7.5)	544 (9.1)	0 (0.0)	~ ~	10.1 (0.23)
Quebec, Canada	56 (4.3)	538 (3.3)	40 (4.1)	528 (3.7)	4 (1.9)	509 (12.1)	9.9 (0.12)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

To what degree is each of the following a problem among fourth grade students in your school?

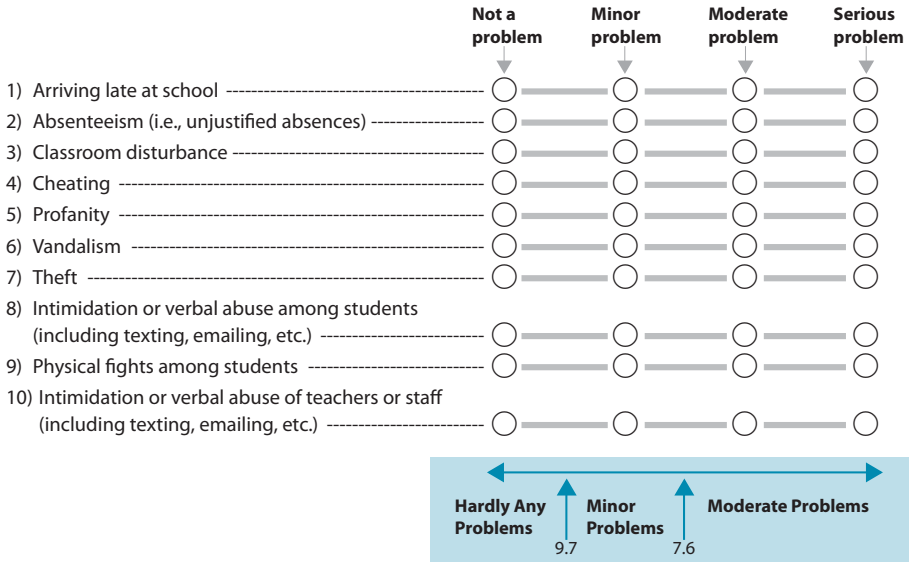


Exhibit 6.10: School Discipline and Safety

TIMSS 2011 8th Grade Mathematics

Reported by Principals

Students were scored according to their principals' responses concerning eleven potential school problems on the *School Discipline and Safety* scale. Students in schools with **Hardly Any Problems** had a score on the scale of at least 10.7, which corresponds to their principals reporting "not a problem" for six of the eleven discipline and safety issues and "minor problem" for the other five, on average. Students in schools with **Moderate Problems** had a score no higher than 8.0, which corresponds to their principals reporting "moderate problem" for six of the eleven issues and "minor problem" for the other five, on average. All other students attended schools with **Minor Problems.**

Country	Hardly Any Problems		Minor Problems		Moderate Problems		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Kazakhstan	81 (3.5)	487 (5.0)	19 (3.5)	486 (7.4)	0 (0.0)	~ ~	11.8 (0.11)
Iran, Islamic Rep. of	66 (3.3)	423 (6.0)	33 (3.3)	400 (5.4)	1 (0.6)	~ ~	11.4 (0.11)
Armenia	66 (4.0)	470 (4.1)	29 (3.9)	460 (7.8)	6 (1.9)	469 (10.8)	11.0 (0.13)
Chinese Taipei	64 (4.1)	611 (4.7)	35 (4.2)	606 (6.8)	1 (0.8)	~ ~	11.4 (0.15)
Georgia	61 (3.2)	425 (5.1)	35 (3.3)	442 (6.9)	3 (1.4)	445 (27.7)	10.8 (0.10)
Ukraine	59 (4.5)	484 (5.5)	33 (4.3)	474 (5.9)	7 (2.4)	465 (15.9)	10.7 (0.16)
Qatar	52 (0.6)	415 (5.2)	36 (0.3)	403 (2.9)	12 (0.4)	391 (7.5)	10.6 (0.04)
Singapore	51 (0.0)	623 (4.6)	49 (0.0)	597 (5.7)	0 (0.0)	~ ~	10.9 (0.00)
Hong Kong SAR	51 (4.6)	615 (7.0)	49 (4.7)	554 (7.6)	1 (0.0)	~ ~	10.9 (0.15)
Russian Federation	50 (3.4)	550 (5.1)	50 (3.5)	528 (5.4)	0 (0.4)	~ ~	10.5 (0.07)
Romania	50 (3.9)	464 (5.5)	41 (3.7)	461 (7.9)	9 (2.6)	411 (14.8)	10.5 (0.17)
Indonesia	47 (4.0)	396 (5.8)	39 (4.7)	379 (7.9)	14 (3.1)	371 (11.0)	10.3 (0.13)
United Arab Emirates	47 (2.0)	469 (3.1)	36 (2.4)	442 (4.7)	17 (1.4)	446 (4.8)	10.2 (0.08)
Lebanon	47 (4.4)	460 (5.6)	39 (4.4)	443 (7.1)	14 (2.9)	428 (7.4)	10.2 (0.20)
Saudi Arabia	46 (4.6)	393 (5.6)	26 (3.8)	397 (10.6)	29 (3.9)	395 (9.4)	9.7 (0.22)
Japan	45 (4.1)	579 (4.8)	35 (4.1)	568 (4.8)	20 (3.3)	551 (4.3)	10.0 (0.18)
Oman	43 (3.3)	388 (3.9)	33 (3.2)	343 (6.1)	25 (2.9)	357 (6.7)	9.8 (0.19)
England	41 (4.6)	531 (9.1)	58 (4.7)	492 (8.6)	1 (0.0)	~ ~	10.6 (0.14)
Macedonia, Rep. of	38 (3.9)	436 (8.5)	49 (4.0)	428 (7.5)	13 (2.1)	403 (16.9)	10.0 (0.15)
Korea, Rep. of	38 (3.7)	617 (3.8)	50 (4.2)	614 (3.6)	13 (3.0)	595 (7.9)	10.1 (0.17)
Bahrain	37 (0.3)	412 (3.0)	49 (0.3)	407 (2.6)	14 (0.2)	411 (7.2)	10.0 (0.01)
Thailand	34 (4.1)	428 (9.1)	61 (4.3)	425 (6.4)	5 (1.8)	448 (22.3)	10.1 (0.13)
Australia	33 (3.8)	538 (10.7)	62 (3.9)	496 (5.6)	5 (1.5)	458 (18.4)	10.1 (0.10)
Ghana	33 (4.3)	353 (7.5)	62 (4.2)	322 (5.7)	6 (1.9)	296 (10.4)	10.0 (0.13)
Slovenia	32 (3.5)	500 (3.6)	61 (4.0)	510 (2.5)	7 (2.3)	485 (6.7)	9.9 (0.11)
Norway	32 (4.7)	482 (3.8)	64 (4.7)	472 (3.0)	4 (1.7)	459 (10.8)	10.1 (0.13)
United States	30 (2.3)	518 (4.8)	66 (2.3)	509 (3.7)	4 (0.8)	474 (19.5)	10.1 (0.07)
Italy	30 (3.3)	508 (4.1)	48 (3.3)	500 (4.7)	23 (2.7)	484 (5.0)	9.4 (0.13)
Chile	29 (3.9)	448 (7.5)	54 (4.5)	412 (4.2)	16 (3.4)	380 (5.2)	9.6 (0.15)
Palestinian Nat'l Auth.	27 (3.7)	412 (6.4)	44 (3.7)	402 (6.5)	29 (3.4)	400 (7.9)	9.1 (0.20)
Finland	27 (4.1)	526 (3.9)	70 (4.1)	509 (2.6)	3 (1.5)	498 (7.7)	9.9 (0.11)
Turkey	26 (3.1)	485 (9.8)	49 (3.4)	444 (5.2)	25 (2.7)	434 (8.1)	9.2 (0.14)
Lithuania	26 (3.5)	490 (7.0)	72 (3.7)	507 (3.0)	2 (1.1)	~ ~	10.1 (0.11)
Israel	26 (3.9)	534 (7.6)	58 (4.2)	526 (5.6)	16 (2.6)	465 (14.8)	9.4 (0.16)
Malaysia	25 (3.8)	469 (10.1)	72 (4.0)	433 (6.3)	3 (1.2)	349 (15.1)	9.9 (0.10)
New Zealand	23 (3.5)	506 (11.0)	74 (3.9)	483 (6.0)	3 (1.7)	482 (30.1)	9.7 (0.09)
Jordan	22 (3.0)	416 (8.4)	51 (4.0)	406 (4.9)	27 (3.6)	397 (7.5)	9.1 (0.14)
Hungary	22 (3.5)	525 (6.4)	68 (3.9)	506 (4.2)	10 (2.5)	450 (13.7)	9.6 (0.11)
Sweden r	18 (4.1)	498 (5.6)	80 (4.4)	482 (2.5)	3 (1.5)	463 (10.2)	9.5 (0.10)
Morocco	13 (2.0)	389 (7.3)	38 (3.6)	360 (4.7)	49 (3.4)	376 (3.4)	8.2 (0.13)
Syrian Arab Republic	11 (2.5)	386 (13.7)	21 (3.8)	391 (10.1)	68 (4.1)	376 (5.4)	7.4 (0.19)
Tunisia	9 (1.8)	421 (7.2)	44 (3.9)	423 (3.9)	47 (3.9)	428 (4.4)	8.1 (0.12)
International Avg.	38 (0.5)	478 (1.0)	49 (0.6)	463 (0.9)	13 (0.4)	434 (2.2)	

Centerpoint of scale set at 10.

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

A tilde (~) indicates insufficient data to report achievement.

An "r" indicates data are available for at least 70% but less than 85% of the students.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

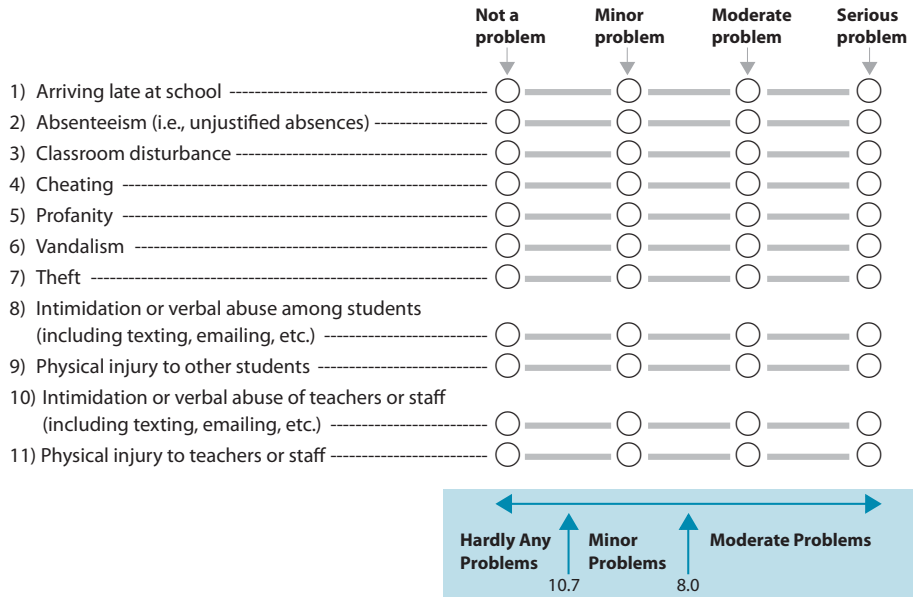
Exhibit 6.10: School Discipline and Safety (Continued)

TIMSS 2011 Mathematics 8th Grade

Country	Hardly Any Problems		Minor Problems		Moderate Problems		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Ninth Grade Participants							
Honduras	35 (4.2)	349 (8.8)	51 (4.8)	332 (4.6)	14 (3.2)	325 (7.6)	9.8 (0.15)
South Africa	8 (2.1)	367 (19.4)	63 (3.6)	355 (3.9)	29 (3.3)	338 (5.3)	8.8 (0.10)
Botswana	5 (1.8)	403 (14.7)	70 (3.7)	399 (3.1)	25 (3.6)	384 (5.6)	8.8 (0.09)
Benchmarking Participants							
Dubai, UAE	62 (0.4)	486 (3.1)	25 (0.4)	471 (4.0)	13 (0.1)	447 (3.3)	10.9 (0.01)
Massachusetts, US	49 (7.2)	582 (9.3)	41 (7.0)	538 (6.6)	9 (4.4)	530 (15.7)	10.6 (0.20)
Abu Dhabi, UAE	41 (3.9)	468 (6.0)	43 (4.5)	435 (9.1)	16 (3.2)	437 (7.8)	10.1 (0.18)
Minnesota, US	37 (6.3)	553 (11.7)	59 (6.9)	542 (5.2)	4 (3.5)	544 (20.5)	10.2 (0.22)
Ontario, Canada	36 (4.3)	515 (3.5)	58 (4.3)	512 (3.6)	6 (2.2)	488 (14.0)	10.2 (0.14)
Alberta, Canada	32 (3.9)	515 (4.7)	68 (3.9)	501 (3.1)	0 (0.0)	~ ~	10.3 (0.12)
Quebec, Canada	31 (3.7)	544 (4.1)	63 (4.2)	527 (3.9)	6 (2.0)	523 (12.8)	10.0 (0.11)
Florida, US	27 (7.4)	511 (19.0)	69 (7.7)	514 (8.2)	4 (2.5)	496 (28.5)	9.8 (0.22)
Indiana, US r	27 (7.1)	541 (10.9)	71 (7.5)	520 (6.6)	2 (0.1)	~ ~	10.2 (0.18)
Colorado, US	25 (6.3)	538 (8.1)	72 (7.0)	511 (7.7)	3 (0.2)	497 (3.5)	9.8 (0.18)
Connecticut, US r	24 (4.7)	558 (11.3)	76 (4.7)	507 (7.8)	0 (0.0)	~ ~	10.1 (0.11)
Alabama, US r	23 (7.5)	488 (15.0)	72 (7.4)	463 (8.6)	5 (2.9)	420 (11.9)	9.9 (0.23)
North Carolina, US	17 (5.2)	547 (21.6)	82 (5.4)	533 (8.7)	1 (0.1)	~ ~	9.7 (0.19)
California, US r	14 (6.1)	497 (6.1)	77 (6.2)	500 (6.5)	8 (2.6)	407 (16.5)	9.6 (0.20)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

To what degree is each of the following a problem among eighth grade students in your school?



Students Bullied at School

In general, bullying involves aggression or negative behavior intended to harm or bother less physically or psychologically powerful persons, although a New Zealand review of the literature found a range of definitions and terminology relating bullying to violence and abuse (Carroll-Lind, 2009). There is growing evidence that bullying in schools is on the rise, especially with the emergence of cyber-bullying, and that bullying does have a negative impact on students' educational achievement. To provide data about bullying in the participating countries, TIMSS 2011 created the Students Bullied at School scale, based on how often students experienced six bullying behaviors:

- I was made fun of or called names;
- I was left out of games or activities by other students;
- Someone spread lies about me;
- Something was stolen from me;
- I was hit or hurt by other student(s); and
- I was made to do things I didn't want to do by other students.

Exhibit 6.11 provides the results for the Students Bullied at School scale for the TIMSS 2011 fourth grade assessment. Students were scored according to their responses to how often they experienced six bullying behaviors (detailed on the second page of the exhibit). Students bullied **Almost Never** reported never experiencing three of six bullying behaviors and each of the other three behaviors "a few times a year," on average. Internationally, across the fourth-grade countries, 48 percent of the students, on average, **Almost Never** experienced these bullying behaviors. However, the percentages ranged from 17 to 80 percent.

The majority of fourth grade students reported being bullied either **About Monthly** or **About Weekly**. Internationally, on average across the fourth grade countries, 32 percent of the students were reportedly bullied **About Monthly** and 20 percent were bullied **About Weekly**. Students bullied **About Weekly** reported experiencing each of three of the six behaviors "once or twice a month" (bullied 3–6 times a month) and, in addition, each of the other three "a few times a year," on average.

The fourth grade students' reports about being bullied were related to their average mathematics achievement on TIMSS 2011. Each successive category of increased bullying was related to a decrease in average mathematics

achievement to the extent that there was a 32-point difference in achievement between **Almost Never** being bullied and being bullied **About Weekly** (501 vs. 469).

Exhibit 6.12 provides the results for the TIMSS 2011 eighth grade assessment for the Students Bullied at School scale, which was based on the same six bullying behaviors (detailed on the second page of the exhibit) as the fourth grade scale. In contrast to the previous section, where principals reported more school discipline and safety problems at the eighth than at the fourth grade, the eighth grade students reported experiencing somewhat less bullying behavior than the fourth grade students. On average internationally, the majority of eighth grade students (59%) **Almost Never** experienced these bullying behaviors, compared to 48 percent at the fourth grade, whereas just 12 percent of the eighth grade students reported being bullied **About Weekly**, compared to 20 percent at the fourth grade. Similar to the fourth grade, there was a negative relationship between eighth grade students' reports about being bullied and average mathematics achievement, with students who were **Almost Never** bullied having achievement 32 points higher than students who reported being bullied **About Weekly** (473 vs. 441).

Exhibit 6.11: Students Bullied at School

TIMSS 2011 4th Grade Mathematics

Reported by Students

Students were scored according to their responses to how often they experienced six bullying behaviors on the *Students Bullied at School* scale. Students bullied **Almost Never** had a score on the scale of at least 10.1, which corresponds to "never" experiencing three of the six bullying behaviors and each of the other three behaviors "a few times a year," on average. Students bullied **About Weekly** had a score no higher than 8.3, which corresponds to their experiencing each of three of the six behaviors "once or twice a month" and each of the other three "a few times a year," on average. All other students were bullied **About Monthly.**

Country	Almost Never		About Monthly		About Weekly		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Armenia	80 (0.8)	459 (3.5)	13 (0.7)	447 (5.6)	7 (0.5)	408 (6.4)	11.5 (0.05)
Azerbaijan	75 (1.5)	483 (6.0)	16 (1.0)	454 (6.5)	9 (0.7)	418 (7.8)	11.4 (0.08)
Sweden	68 (1.0)	509 (2.1)	25 (1.0)	498 (3.4)	7 (0.5)	483 (5.8)	10.9 (0.04)
Georgia	66 (1.2)	464 (3.0)	23 (0.8)	451 (5.6)	11 (0.8)	407 (8.8)	10.9 (0.06)
Kazakhstan	64 (1.7)	503 (4.2)	23 (1.2)	512 (6.4)	13 (0.9)	489 (8.0)	10.8 (0.08)
Ireland	64 (1.3)	539 (2.7)	25 (1.0)	522 (3.4)	12 (0.9)	486 (5.0)	10.7 (0.06)
Croatia	61 (1.1)	497 (2.3)	28 (0.9)	487 (3.0)	11 (0.6)	462 (4.8)	10.6 (0.05)
Finland	61 (1.2)	549 (2.5)	30 (0.9)	546 (3.4)	9 (0.6)	523 (5.0)	10.5 (0.04)
Poland	61 (0.9)	487 (2.4)	26 (0.7)	481 (2.9)	13 (0.6)	462 (4.1)	10.6 (0.04)
Denmark	60 (1.1)	544 (2.4)	31 (0.8)	535 (3.2)	9 (0.7)	513 (5.7)	10.5 (0.04)
Serbia	57 (1.2)	523 (3.4)	30 (0.9)	520 (4.1)	13 (0.7)	484 (7.1)	10.5 (0.06)
Northern Ireland	57 (1.3)	571 (3.4)	29 (1.0)	565 (4.1)	14 (1.0)	528 (7.3)	10.4 (0.06)
Austria	53 (1.3)	513 (3.0)	30 (0.9)	510 (3.5)	17 (0.9)	493 (3.4)	10.2 (0.05)
Norway	53 (1.8)	502 (3.1)	33 (1.1)	493 (3.6)	14 (0.9)	473 (7.0)	10.2 (0.06)
Korea, Rep. of	53 (1.2)	608 (2.2)	32 (0.8)	608 (2.3)	15 (0.6)	592 (3.9)	10.3 (0.05)
Chinese Taipei	53 (1.3)	597 (2.1)	30 (0.8)	592 (2.7)	17 (0.8)	573 (3.6)	10.2 (0.05)
United States	51 (0.7)	549 (2.1)	29 (0.5)	544 (2.0)	20 (0.6)	520 (3.2)	10.1 (0.03)
Italy	51 (1.2)	514 (3.1)	33 (1.0)	509 (3.0)	16 (0.7)	491 (3.9)	10.2 (0.05)
Slovenia	50 (1.3)	520 (2.5)	32 (0.8)	517 (2.7)	18 (1.0)	488 (3.5)	10.0 (0.05)
Japan	50 (1.2)	588 (2.1)	33 (0.8)	589 (2.8)	17 (0.8)	574 (3.3)	10.1 (0.05)
Hong Kong SAR	50 (1.2)	608 (3.1)	33 (0.9)	604 (3.5)	17 (0.7)	582 (7.1)	10.1 (0.04)
Portugal	49 (1.4)	536 (4.0)	35 (1.2)	535 (3.8)	17 (0.9)	515 (4.8)	10.1 (0.06)
Germany	48 (1.2)	537 (2.7)	36 (0.9)	530 (2.4)	16 (0.8)	511 (4.0)	10.1 (0.05)
Lithuania	48 (1.3)	543 (3.0)	36 (0.9)	534 (2.6)	17 (0.8)	508 (4.1)	10.0 (0.05)
Romania	47 (1.8)	504 (5.9)	32 (1.5)	475 (6.4)	21 (1.1)	450 (8.7)	9.9 (0.07)
Slovak Republic	46 (1.1)	517 (3.4)	34 (0.8)	505 (4.6)	20 (0.9)	488 (4.8)	9.9 (0.05)
Czech Republic	46 (1.2)	519 (2.8)	34 (1.0)	514 (3.1)	20 (0.8)	488 (4.5)	10.0 (0.05)
Netherlands	46 (1.2)	543 (1.7)	37 (1.1)	543 (2.3)	17 (0.9)	526 (3.1)	9.9 (0.05)
Russian Federation	45 (1.4)	549 (4.3)	35 (1.0)	540 (4.0)	19 (1.0)	530 (4.9)	10.0 (0.06)
England	45 (1.3)	549 (4.2)	36 (1.0)	548 (4.5)	20 (0.8)	519 (5.3)	9.8 (0.05)
Spain	44 (1.3)	488 (3.1)	34 (0.9)	487 (3.3)	23 (1.0)	469 (3.8)	9.8 (0.05)
Yemen	42 (2.1)	260 (6.8)	31 (1.4)	256 (6.9)	27 (1.8)	233 (8.7)	9.7 (0.11)
Malta	42 (0.7)	507 (1.7)	36 (0.7)	499 (2.5)	22 (0.6)	471 (2.6)	9.7 (0.03)
Iran, Islamic Rep. of	41 (1.7)	431 (5.0)	35 (1.2)	434 (4.0)	23 (1.3)	428 (5.0)	9.8 (0.07)
Hungary	40 (1.1)	521 (5.6)	36 (0.8)	525 (3.3)	24 (0.8)	497 (4.4)	9.7 (0.04)
Singapore	39 (0.9)	618 (3.3)	38 (0.6)	610 (3.3)	23 (0.8)	582 (4.2)	9.7 (0.03)
Saudi Arabia	39 (1.7)	422 (5.6)	33 (1.2)	419 (6.2)	27 (1.2)	386 (6.7)	9.6 (0.08)
Tunisia	39 (1.4)	377 (4.8)	37 (1.1)	362 (4.1)	24 (1.2)	333 (5.4)	9.7 (0.06)
Belgium (Flemish)	39 (1.1)	556 (2.6)	41 (0.9)	552 (2.2)	20 (0.8)	533 (2.7)	9.7 (0.04)
Chile	38 (1.1)	478 (2.4)	31 (0.9)	467 (2.6)	31 (1.0)	441 (3.2)	9.5 (0.05)
Australia	38 (1.1)	525 (2.9)	38 (1.0)	521 (3.7)	25 (0.7)	498 (4.2)	9.5 (0.04)
Turkey	37 (0.9)	494 (3.8)	33 (0.7)	477 (4.6)	30 (0.9)	442 (5.7)	9.5 (0.04)
Kuwait	37 (1.5)	362 (3.4)	33 (1.0)	358 (4.2)	30 (1.3)	319 (5.5)	9.5 (0.07)
Morocco	35 (1.9)	354 (6.7)	33 (1.1)	338 (4.0)	32 (1.6)	317 (4.8)	9.4 (0.08)
United Arab Emirates	34 (0.8)	454 (2.8)	35 (0.5)	439 (2.6)	31 (0.8)	412 (2.9)	9.4 (0.04)
New Zealand	32 (1.0)	499 (3.4)	37 (1.0)	494 (2.9)	31 (0.9)	468 (4.1)	9.3 (0.04)
Bahrain	31 (1.1)	460 (3.9)	33 (1.1)	442 (4.0)	36 (1.3)	421 (3.9)	9.2 (0.06)
Oman	31 (1.2)	399 (3.3)	37 (0.9)	387 (3.9)	31 (1.0)	372 (4.0)	9.3 (0.05)
Qatar	30 (1.1)	441 (5.4)	32 (1.0)	425 (4.5)	38 (1.0)	392 (3.9)	9.1 (0.05)
Thailand	17 (1.2)	476 (5.9)	35 (1.2)	461 (5.0)	48 (1.6)	451 (5.4)	8.6 (0.06)
International Avg.	48 (0.2)	501 (0.5)	32 (0.1)	493 (0.6)	20 (0.1)	469 (0.7)	

Centerpoint of scale set at 10.

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 6.11: Students Bullied at School (Continued)

Country	Almost Never		About Monthly		About Weekly		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Sixth Grade Participants							
Yemen	43 (1.9)	355 (8.0)	34 (1.3)	358 (6.1)	23 (1.3)	330 (7.1)	9.8 (0.08)
Honduras	38 (1.2)	405 (6.3)	32 (0.9)	404 (5.9)	30 (1.2)	384 (5.9)	9.5 (0.06)
Botswana	12 (0.7)	449 (7.5)	41 (0.9)	427 (4.4)	47 (1.1)	410 (3.9)	8.6 (0.03)
Benchmarking Participants							
Florida, US	50 (1.4)	552 (3.6)	29 (0.9)	550 (3.7)	21 (1.1)	526 (4.0)	10.1 (0.06)
North Carolina, US	49 (1.5)	563 (4.6)	32 (1.2)	556 (4.7)	19 (1.1)	534 (5.9)	10.0 (0.06)
Quebec, Canada	44 (1.4)	540 (2.4)	37 (1.1)	534 (3.3)	19 (1.1)	515 (3.5)	9.8 (0.05)
Alberta, Canada	42 (1.3)	514 (3.1)	35 (0.9)	509 (3.2)	22 (1.0)	489 (3.2)	9.7 (0.05)
Ontario, Canada	42 (1.1)	523 (3.2)	36 (0.9)	525 (3.5)	22 (1.0)	501 (4.2)	9.7 (0.04)
Dubai, UAE	37 (1.6)	489 (2.6)	35 (0.9)	476 (3.4)	28 (1.2)	440 (3.8)	9.5 (0.06)
Abu Dhabi, UAE	33 (1.4)	436 (6.1)	36 (0.8)	422 (5.3)	31 (1.4)	398 (5.3)	9.4 (0.07)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

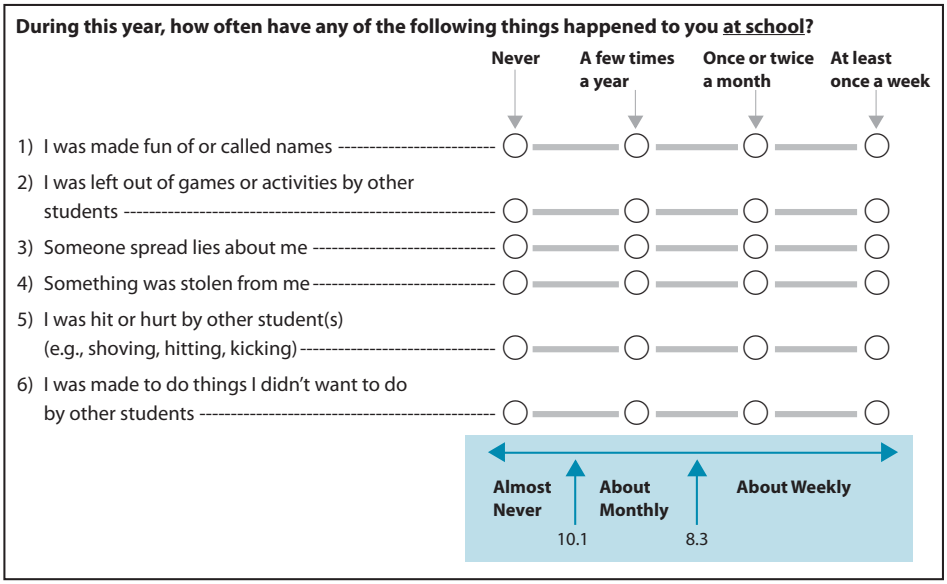


Exhibit 6.12: Students Bullied at School

TIMSS 2011 8th Grade Mathematics

Reported by Students

Students were scored according to their responses to how often they experienced six bullying behaviors on the *Students Bullied at School* scale. Students bullied **Almost Never** had a score on the scale of at least 9.6, which corresponds to "never" experiencing three of the six bullying behaviors and each of the other three behaviors "a few times a year," on average. Students bullied **About Weekly** had a score no higher than 7.7, which corresponds to their experiencing each of three of the six behaviors "once or twice a month" and each of the other three "a few times a year," on average. All other students were bullied **About Monthly**.

Country	Almost Never		About Monthly		About Weekly		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Armenia	87 (0.7)	472 (2.7)	11 (0.6)	462 (5.7)	3 (0.3)	418 (9.7)	11.5 (0.04)
Sweden	79 (0.6)	487 (2.0)	18 (0.5)	482 (2.8)	3 (0.3)	454 (7.6)	10.9 (0.03)
Georgia	79 (0.9)	443 (4.0)	17 (0.8)	425 (5.0)	4 (0.4)	374 (10.0)	11.2 (0.05)
Norway	77 (0.8)	477 (2.6)	19 (0.7)	473 (4.2)	4 (0.3)	446 (10.3)	10.8 (0.04)
Italy	76 (1.1)	502 (2.4)	19 (0.9)	492 (3.7)	5 (0.4)	468 (7.8)	10.7 (0.05)
Kazakhstan	73 (1.1)	487 (4.0)	21 (1.0)	496 (5.2)	5 (0.5)	472 (7.6)	11.0 (0.06)
Finland	71 (0.9)	517 (2.5)	24 (0.8)	509 (3.5)	5 (0.4)	502 (5.0)	10.5 (0.04)
Ukraine	70 (1.2)	486 (4.3)	24 (1.1)	478 (4.4)	6 (0.5)	442 (7.6)	10.4 (0.05)
Russian Federation	69 (0.9)	542 (3.5)	25 (0.7)	538 (4.0)	6 (0.4)	522 (8.5)	10.4 (0.04)
England	68 (1.1)	509 (5.6)	24 (0.7)	511 (6.0)	7 (0.6)	486 (11.1)	10.4 (0.05)
Macedonia, Rep. of	68 (0.9)	445 (5.4)	22 (0.7)	422 (5.6)	10 (0.6)	377 (8.5)	10.3 (0.05)
Chinese Taipei	67 (1.0)	612 (3.7)	26 (0.8)	611 (3.8)	7 (0.4)	580 (5.7)	10.4 (0.05)
Lithuania	65 (1.1)	507 (2.5)	28 (1.0)	504 (3.3)	7 (0.5)	465 (5.1)	10.2 (0.05)
Korea, Rep. of	65 (1.1)	613 (3.1)	28 (0.9)	616 (3.7)	7 (0.5)	603 (5.7)	10.3 (0.05)
Japan	63 (1.2)	566 (3.2)	28 (0.8)	576 (3.4)	9 (0.6)	562 (6.0)	10.3 (0.05)
United States	63 (0.7)	513 (2.7)	28 (0.6)	510 (3.5)	9 (0.3)	496 (3.3)	10.1 (0.02)
Chile	62 (0.9)	423 (2.8)	30 (0.8)	410 (3.2)	9 (0.5)	394 (4.9)	9.9 (0.03)
Hungary	61 (1.2)	508 (3.8)	31 (0.9)	505 (4.3)	8 (0.5)	487 (6.1)	10.0 (0.05)
Saudi Arabia	60 (1.2)	400 (5.1)	30 (1.0)	393 (4.9)	10 (0.6)	372 (6.5)	10.1 (0.06)
Slovenia	59 (1.0)	504 (2.5)	32 (1.0)	509 (2.9)	8 (0.5)	499 (5.1)	9.9 (0.04)
Australia	58 (1.1)	511 (5.3)	31 (1.0)	504 (5.3)	11 (0.7)	480 (7.3)	9.9 (0.05)
Tunisia	58 (1.0)	426 (2.7)	31 (0.7)	426 (3.2)	11 (0.7)	419 (5.2)	9.9 (0.04)
Iran, Islamic Rep. of	56 (1.1)	420 (4.8)	33 (0.8)	415 (4.8)	12 (0.6)	395 (5.8)	9.9 (0.05)
New Zealand	55 (0.9)	495 (5.3)	33 (0.7)	489 (5.9)	12 (0.5)	471 (6.3)	9.8 (0.04)
Bahrain	55 (1.1)	422 (2.7)	29 (1.0)	411 (3.1)	16 (0.6)	370 (5.6)	9.8 (0.04)
Syrian Arab Republic	54 (1.4)	392 (5.0)	31 (1.0)	375 (4.6)	14 (0.8)	361 (5.9)	9.8 (0.06)
Hong Kong SAR	54 (1.3)	585 (4.2)	36 (1.0)	589 (3.8)	10 (0.7)	582 (8.4)	9.7 (0.05)
Lebanon	53 (1.9)	464 (4.4)	30 (1.1)	444 (4.7)	17 (1.3)	418 (3.6)	9.7 (0.08)
Romania	53 (1.2)	476 (4.6)	34 (0.9)	457 (4.2)	13 (0.7)	411 (5.8)	9.7 (0.05)
Turkey	52 (1.1)	466 (4.7)	33 (0.8)	454 (4.2)	15 (0.7)	413 (5.5)	9.7 (0.05)
Singapore	52 (0.8)	618 (3.9)	36 (0.6)	609 (4.0)	12 (0.5)	589 (5.4)	9.7 (0.03)
United Arab Emirates	51 (0.9)	468 (2.2)	33 (0.6)	456 (2.4)	16 (0.5)	420 (3.3)	9.6 (0.04)
Qatar	51 (1.6)	426 (4.5)	31 (1.2)	409 (5.2)	18 (0.8)	374 (6.1)	9.6 (0.06)
Malaysia	49 (1.2)	444 (5.2)	39 (0.9)	442 (5.8)	12 (0.8)	418 (9.3)	9.6 (0.05)
Morocco	49 (1.1)	375 (2.3)	36 (0.8)	377 (2.5)	15 (0.7)	359 (3.6)	9.6 (0.04)
Jordan	48 (1.2)	426 (3.4)	33 (1.0)	412 (4.2)	19 (0.7)	362 (5.7)	9.5 (0.05)
Palestinian Nat'l Auth.	46 (1.2)	426 (3.2)	38 (0.9)	399 (4.0)	16 (0.8)	363 (7.0)	9.5 (0.05)
Indonesia	45 (1.4)	382 (4.1)	34 (0.9)	392 (5.0)	21 (0.9)	387 (6.7)	9.5 (0.07)
Oman	41 (0.9)	389 (3.0)	37 (0.7)	370 (3.4)	21 (0.7)	330 (4.2)	9.2 (0.03)
Thailand	30 (0.8)	426 (4.7)	43 (0.7)	431 (4.8)	27 (0.8)	424 (4.5)	8.8 (0.04)
Ghana	22 (1.0)	349 (6.0)	38 (1.0)	342 (4.6)	40 (1.2)	317 (4.7)	8.4 (0.05)
Israel	– –	– –	– –	– –	– –	– –	– –
International Avg.	59 (0.2)	473 (0.6)	29 (0.1)	467 (0.7)	12 (0.1)	441 (1.0)	

Centerpoint of scale set at 10.

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

A dash (–) indicates comparable data are not available.

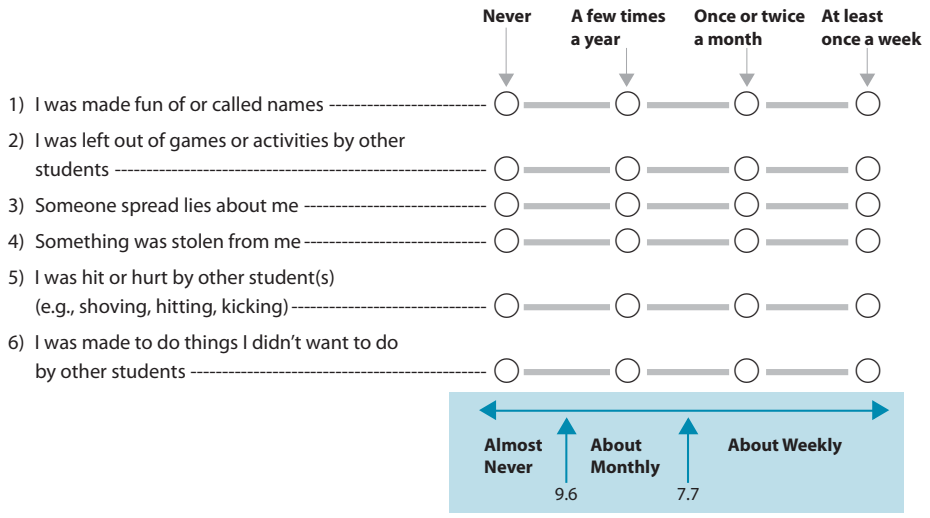
SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 6.12: Students Bullied at School (Continued)

Country	Almost Never		About Monthly		About Weekly		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Ninth Grade Participants							
Honduras	49 (1.1)	340 (4.5)	36 (0.9)	346 (4.3)	15 (0.6)	325 (4.8)	9.6 (0.04)
South Africa	25 (0.7)	393 (3.9)	42 (0.8)	362 (2.3)	33 (1.0)	322 (3.0)	8.5 (0.04)
Botswana	19 (0.7)	424 (3.6)	48 (0.7)	403 (2.7)	33 (0.7)	379 (3.2)	8.4 (0.02)
Benchmarking Participants							
Quebec, Canada	73 (0.9)	533 (2.4)	22 (0.7)	531 (3.0)	5 (0.4)	521 (5.9)	10.5 (0.04)
Massachusetts, US	71 (1.0)	563 (5.7)	23 (1.0)	562 (6.1)	6 (0.6)	533 (7.1)	10.5 (0.05)
California, US	67 (1.7)	496 (5.2)	24 (1.3)	493 (5.6)	9 (0.6)	477 (7.3)	10.3 (0.07)
Florida, US	64 (1.5)	517 (6.7)	27 (1.4)	519 (7.3)	9 (0.9)	488 (8.8)	10.1 (0.07)
North Carolina, US	64 (1.0)	540 (6.6)	28 (1.0)	537 (6.7)	8 (0.8)	515 (17.6)	10.1 (0.06)
Connecticut, US	63 (1.4)	520 (4.9)	28 (1.0)	523 (5.9)	9 (0.8)	511 (8.2)	10.1 (0.06)
Minnesota, US	61 (1.6)	550 (5.2)	30 (1.4)	539 (5.2)	9 (0.7)	532 (5.2)	10.0 (0.06)
Indiana, US	59 (1.5)	523 (5.4)	30 (1.3)	523 (5.5)	11 (0.9)	517 (7.2)	9.9 (0.07)
Colorado, US	58 (1.8)	520 (5.1)	31 (1.5)	519 (5.9)	11 (1.0)	506 (8.0)	9.9 (0.07)
Ontario, Canada	58 (1.2)	515 (3.1)	31 (0.9)	508 (3.4)	12 (0.8)	496 (3.4)	9.9 (0.05)
Alabama, US	57 (1.9)	469 (6.7)	32 (1.5)	467 (5.9)	11 (0.8)	455 (7.8)	9.9 (0.07)
Dubai, UAE	54 (2.1)	491 (2.7)	32 (1.3)	475 (4.1)	14 (1.1)	439 (5.2)	9.7 (0.09)
Alberta, Canada	52 (1.1)	511 (3.0)	35 (0.8)	505 (2.7)	14 (0.8)	485 (3.3)	9.6 (0.05)
Abu Dhabi, UAE	50 (1.4)	456 (4.2)	33 (0.9)	455 (3.8)	17 (1.0)	418 (5.6)	9.6 (0.06)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

During this year, how often have any of the following things happened to you at school?



OECD Program for International Student Assessment 2012

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

USA

Date of Test (Main Survey PISA 2012)		
Month	Day	2012

Student Questionnaire - Form A

Place Label Here

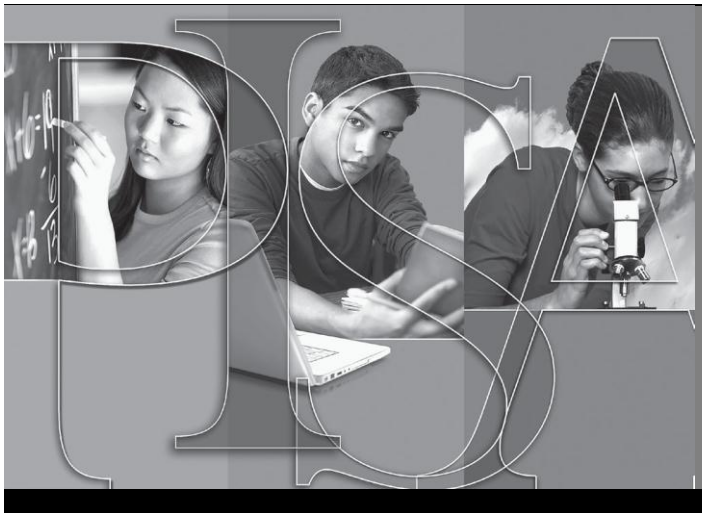
Student ID

Participation Status

English 313

U.S. participation in this study is sponsored by the National Center for Education Statistics (NCES), U.S. Department of Education. Your responses are protected by federal statute (20 U.S.C., § 9573). Your answers may be used only for statistical purposes and may not be disclosed, or used, in identifiable form for any other purpose except as required by law. By law, everyone working on this NCES survey is subject to a jail term of up to 5 years, a fine of up to $250,000, or both if he or she willfully discloses ANY identifiable information about you.

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. The valid OMB control number for this voluntary information collection is 1850-0755. The time required to complete this information collection is estimated to average 30 minutes per response, including the time to review instructions, search existing data resources, gather the data needed, and complete and review the information collection. **If you have any comments concerning the accuracy of the time estimate(s) or suggestions for improving the form, please write to**: U.S. Department of Education, Washington, D.C. 20202-4537. **If you have comments or concerns regarding the status of your individual submission of this form, write directly to:** Program for International Student Assessment (PISA), National Center for Education Statistics, U.S. Department of Education, 1990 K Street, N.W., Washington, D.C. 20006. O.M.B. No. 1850-0755, Approval Expires 11/30/2013



Project Consortium

- **Australian Council for Educational Research (ACER)**
- **cApStAn Linguistic Quality Control (Belgium)**
- **Deutsches Institut für Internationale Pädagogische Forschung (DIPF, Germany)**
- **Educational Testing Service (ETS, USA)**
- **Institutt for Lærerutdanning og Skoleutvikling (ILS, Norway)**
- **Leibniz - Institute for Science and Mathematics Education (IPN, Germany)**
- **National Institute for Educational Policy Research (NIER, Japan)**
- **The Tao Initiative: CRP - Henri Tudor and Université de Luxembourg EMACS (Luxembourg)**
- **Unité d'analyse des systèmes et des pratiques d'enseignement (aSPe, Belgium)**
- **Westat (USA)**

In this booklet you will find questions about:

- You (Section A)
- Your Family and Home (Section B)
- Learning Mathematics (Section C and Section E)
- Your Problem Solving Experiences (Section D)

Please read each question carefully and answer as accurately as you can. In the test, you usually circled your answers. For this questionnaire, you will normally answer by darkening a circle. For a few questions you will need to write a short answer.

If you make a mistake when darkening a circle, erase your mistake and darken the correct circle. If you make a mistake when writing an answer, simply cross it out and write the correct answer next to it.

In this questionnaire, there are no right or wrong answers. Your answers should be the ones that are right for you.

You may ask for help if you do not understand something or are not sure how to answer a question.

Your answers will be combined with answers from other students to calculate totals and averages. All information (or responses) you provide may only be used for statistical purposes and may not be disclosed, or used, in identifiable form for any other purpose except as required by law [Education Science Reform Act of 2002 (ESRA 2002), 20 U.S. Code, Section 9573].

SECTION A: ABOUT YOU

ST01

Q1 What grade are you in?

_____

grade

ST03

Q2 When were you born?

(Please write the month, day and year you were born.)

_____ _____ *19___*

Month *Day* *Year*

ST04

Q3 Are you female or male?

Female ○1 *Male* ○2

USA_ST116A01

Q4 Which best describes you?

(Please darken only one circle.)

I am Hispanic or Latino. ○1

I am <u>not</u> Hispanic or Latino. ○2

USA_ST117A01

Q5 Which of these categories best describes your race?

(Please darken one or more circles.)

White ○1

Black or African American ○1

Asian ○1

American Indian or Alaska Native ○1

Native Hawaiian or Other Pacific Islander ○1

USA_ST05N01

Q6 Did you attend pre-school?

No ○1

Yes, for one year or less ○2

Yes, for more than one year ○3

USA_ST05N02

Q7 Did you attend kindergarten?

No ○1

Yes ○2

ST06

Q8 How old were you when you started first grade?

_____ *years old*

USA_ST07N01

Q9 Have you ever repeated a grade?

(Please darken only one circle in each row.)

	No, never	*Yes, once*	*Yes, twice or more*
a) In kindergarten	○1	○2	○3
b) In grades 1-6	○1	○2	○3
c) In grades 7-9	○1	○2	○3
d) In grades 10-12	○1	○2	○3

ST08

Q10 In the last two full weeks of school, how many times did you arrive late for school?

(Please darken only one circle.)

None ○1

One or two times ○2

Three or four times ○3

Five or more times ○4

ST09

Q11 In the last two full weeks of school, how many times did you skip a whole school day?

(Please darken only one circle.)

None ○1

One or two times ○2

Three or four times ○3

Five or more times ○4

ST115

Q12 In the last two full weeks of school, how many times did you skip some classes?

(Please darken only one circle.)

None ○1

One or two times ○2

Three or four times ○3

Five or more times ○4

Q13 What is the highest grade or level of school you expect to complete?

(Please darken only one circle.)

Less than high school	○ 1
High school (high school diploma or GED)	○ 2
Vocational or technical certificate (such as cosmetology or auto mechanics)	○ 3
Associate's degree (2-year degree from a community college)	○ 4
Bachelor's degree (4-year college degree)	○ 5
Master's degree	○ 6
Doctoral or professional degree such as medicine or law	○ 7

SECTION B: ABOUT YOUR FAMILY AND HOME

In this section you will be asked some questions about your family and your home.

Some of the following questions are about your mother and father or those persons who are like a mother or father to you — for example, guardians, step-parents, foster parents, etc.

If you share your time with more than one set of parents or guardians, please answer the following questions for those parents/guardians you spend the most time with.

ST11

Q14 Who usually lives at home with you?

(Please darken only one circle in each row.)

	Yes	*No*
a) Mother (including stepmother or foster mother)	○1	○2
b) Father (including stepfather or foster father)	○1	○2
c) Brother(s) (including stepbrothers)	○1	○2
d) Sister(s) (including stepsisters)	○1	○2
e) Grandparent(s)	○1	○2
f) Others (e.g., cousin)	○1	○2

Q15 What is your mother's main job? (e.g., school teacher, cook, sales manager)

(If she is not working now, please tell us her last main job.)

Please write in the job title. _____

Q16 What does your mother do in her main job? (e.g., teaches high school students, helps prepare meals in a restaurant, manages a sales team)

Please use a sentence to describe the kind of work she does or did in that job.

_____

Q17 What is the highest level of schooling (not including college) completed by your mother?

If you are not sure which circle to choose, please ask the test administrator for help.

(Please darken only one circle.)

She completed grade 12 (high school diploma or GED)	○ 1
She completed grade 9	○ 2
She completed grade 6	○ 3
She did not complete grade 6	○ 4

Q18 Does your mother have any of the following degrees, certificates, or diplomas?

If you are not sure how to answer this question, please ask the test administrator for help.

(Please darken only one circle in each row.)

		Yes	*No*
a)	Master's, doctoral, or professional degree such as medicine or law	○ 1	○ 2
b)	Bachelor's degree (4-year college degree)	○ 1	○ 2
c)	Associate's degree (2-year degree from a community college)	○ 1	○ 2
d)	Vocational or technical certificate/diploma after high school (such as cosmetology or auto mechanics)	○ 1	○ 2

Q19 What is your mother currently doing?

(Please darken only one circle.)

Working full-time for pay ○ 1

Working part-time for pay ○ 2

Not working, but looking for a job ○ 3

Other (e.g., home duties, retired) ○ 4

ST16

Q20 What is your father's main job? (e.g., school teacher, cook, sales manager)

(If he is not working now, please tell us his last main job.)

Please write in the job title. _____

Q21 What does your father do in his main job? (e.g., teaches high school students, helps prepare meals in a restaurant, manages a sales team)

Please use a sentence to describe the kind of work he does or did in that job.

_____

ST17

Q22 What is the highest level of schooling (not including college) completed by your father?

If you are not sure how to answer this question, please ask the test administrator for help.

(Please darken only one circle.)

He completed grade 12 (high school diploma or GED)	○ 1
He completed grade 9	○ 2
He completed grade 6	○ 3
He did not complete grade 6	○ 4

Q23 Does your father have any of the following degrees, certificates, or diplomas?

If you are not sure which circle to choose, please ask the test administrator for help.

(Please darken one circle in each row.)

	Yes	*No*
a) Master's, doctoral, or professional degree such as medicine or law	○1	○2
b) Bachelor's degree (4-year college degree)	○1	○2
c) Associate's degree (2-year degree from a community college)	○1	○2
d) Vocational or technical certificate/diploma after high school (such as cosmetology or auto mechanics)	○1	○2

Q24 What is your father currently doing?

(Please darken only one circle.)

Working full-time for pay ○1

Working part-time for pay ○2

Not working, but looking for a job ○3

Other (e.g., home duties, retired) ○4

ST20

Q25 In what country were you and your parents born?

(Please darken only one circle in each column.)

	You	*Mother*	*Father*
United States*	○ 01	○ 01	○ 01
Other country	○ 02	○ 02	○ 02

*NOTE: The "United States" refers to the 50 states, District of Columbia, and U.S. military bases abroad.

ST21

Q26 If <u>you</u> were NOT born in the United States, how old were you when you arrived in the United States?

If you were less than 12 months old, please write zero (0).

If you were born in the United States please skip this question and go to Q27.

_____ *years old*

ST25

Q27 What language do you speak at home most of the time?

(Please darken only one circle.)

English	○ 313
Spanish	○ 156
Other language	○ 859

Q28 Which of the following are in your home?

(Please darken only one circle in each row.)

	Yes	*No*
a) A desk to study at	○1	○2
b) A room of your own	○1	○2
c) A quiet place to study	○1	○2
d) A computer you can use for school work	○1	○2
e) Educational software	○1	○2
f) A link to the Internet	○1	○2
g) Classic literature (e.g., Shakespeare)	○1	○2
h) Books of poetry	○1	○2
i) Works of art (e.g., paintings)	○1	○2
j) Books to help with your school work	○1	○2
k) Technical reference books or manuals	○1	○2
l) A dictionary	○1	○2
m) A dishwasher	○1	○2
n) A DVD player	○1	○2
o) A guest room	○1	○2
p) A high-speed Internet connection	○1	○2
q) A musical instrument	○1	○2

Q29 How many of these are there at your home?

(Please darken only one circle in each row.)

	None	*One*	*Two*	*Three or more*
a) Cell phones	○	○	○	○ 4
b) Televisions	○	○	○	○ 4
c) Computers	○	○	○	○ 4
d) Cars	○	○	○	○ 4
e) Bathrooms with a bathtub or shower	○	○	○	○ 4

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 ○ 1

11-25 books ○ 2

26-100 books ○ 3

101-200 books ○ 4

201-500 books ○ 5

More than 500 books ○ 6

SECTION C: ABOUT LEARNING MATHEMATICS

ST29

Q31 Thinking about your views on mathematics: to what extent do you agree with the following statements?

(Please darken only one circle in each row.)

	Strongly agree	*Agree*	*Disagree*	*Strongly disagree*
a) I enjoy reading about mathematics.	○ 1	○ 2	○ 3	○ 4
b) Making an effort in mathematics is worth it because it will help me in the work that I want to do later on.	○ 1	○ 2	○ 3	○ 4
c) I look forward to my mathematics lessons.	○ 1	○ 2	○ 3	○ 4
d) I do mathematics because I enjoy it.	○ 1	○ 2	○ 3	○ 4
e) Learning mathematics is worthwhile for me because it will improve my career prospects.	○ 1	○ 2	○ 3	○ 4
f) I am interested in the things I learn in mathematics.	○ 1	○ 2	○ 3	○ 4
g) Mathematics is an important subject for me because I need it for what I want to study later on.	○ 1	○ 2	○ 3	○ 4
h) I will learn many things in mathematics that will help me get a job.	○ 1	○ 2	○ 3	○ 4

Q32 Thinking about how people important to you view mathematics: how strongly do you agree with the following statements?

(Please darken only one circle in each row.)

	Strongly agree	*Agree*	*Disagree*	*Strongly disagree*
a) Most of my friends do well in mathematics.	○ 1	○ 2	○ 3	○ 4
b) Most of my friends work hard at mathematics.	○ 1	○ 2	○ 3	○ 4
c) My friends enjoy taking mathematics tests.	○ 1	○ 2	○ 3	○ 4
d) My parents believe it's important for me to study mathematics.	○ 1	○ 2	○ 3	○ 4
e) My parents believe that mathematics is important for my career.	○ 1	○ 2	○ 3	○ 4
f) My parents like mathematics.	○ 1	○ 2	○ 3	○ 4

Q33 How confident do you feel about having to do the following mathematics tasks?

(Please darken only one circle in each row.)

	Very confident	*Confident*	*Not very confident*	*Not at all confident*
a) Using a train schedule to figure out how long it would take to get from one place to another.	○ 1	○ 2	○ 3	○ 4
b) Calculating how much cheaper a TV would be after a 30% discount.	○ 1	○ 2	○ 3	○ 4
c) Calculating how many square feet of tile you need to cover a floor.	○ 1	○ 2	○ 3	○ 4
d) Understanding graphs presented in newspapers.	○ 1	○ 2	○ 3	○ 4
e) Solving an equation like $3x+5= 17$.	○ 1	○ 2	○ 3	○ 4
f) Finding the actual distance between two places on a map with a 1:10,000 scale.	○ 1	○ 2	○ 3	○ 4
g) Solving an equation like $2(x+3) = (x + 3)(x - 3)$.	○ 1	○ 2	○ 3	○ 4
h) Calculating the gas mileage of a car.	○ 1	○ 2	○ 3	○ 4

Q34 Thinking about your mathematics lessons: to what extent do you agree with the following statements?

(Please darken only one circle in each row.)

	Strongly agree	*Agree*	*Disagree*	*Strongly disagree*
a) If I put in enough effort I can succeed in mathematics.	○ 1	○ 2	○ 3	○ 4
b) Whether or not I do well in mathematics is completely up to me.	○ 1	○ 2	○ 3	○ 4
c) Family demands or other problems prevent me from putting a lot of time into my mathematics work.	○ 1	○ 2	○ 3	○ 4
d) If I had different teachers I would try harder in mathematics.	○ 1	○ 2	○ 3	○ 4
e) If I wanted to I could do well in mathematics.	○ 1	○ 2	○ 3	○ 4
f) I do badly in mathematics whether or not I study for my exams.	○ 1	○ 2	○ 3	○ 4

Q35 Suppose that you are a student in the following situation:

Each week, your mathematics teacher gives a short quiz. Recently you have done badly on these quizzes. Today you are trying to figure out why.

How likely are you to have these thoughts or feelings in this situation?

(Please darken only one circle in each row.)

	Very likely	*Likely*	*Slightly likely*	*Not at all likely*
a) I'm not very good at solving mathematics problems.	○ 1	○ 2	○ 3	○ 4
b) My teacher did not explain the concepts well this week.	○ 1	○ 2	○ 3	○ 4
c) This week I guessed badly on the quiz.	○ 1	○ 2	○ 3	○ 4
d) Sometimes the course material is too hard.	○ 1	○ 2	○ 3	○ 4
e) The teacher did not get students interested in the material.	○ 1	○ 2	○ 3	○ 4
f) Sometimes I am just unlucky.	○ 1	○ 2	○ 3	○ 4

Q36 Thinking about the mathematics you do for school: to what extent do you agree with the following statements?

(Please darken only one circle in each row.)

	Strongly agree	*Agree*	*Disagree*	*Strongly disagree*
a) I have my homework finished in time for mathematics class.	○ 1	○ 2	○ 3	○ 4
b) I work hard on my mathematics homework.	○ 1	○ 2	○ 3	○ 4
c) I am prepared for my mathematics exams.	○ 1	○ 2	○ 3	○ 4
d) I study hard for mathematics quizzes.	○ 1	○ 2	○ 3	○ 4
e) I keep studying until I understand mathematics material.	○ 1	○ 2	○ 3	○ 4
f) I pay attention in mathematics class.	○ 1	○ 2	○ 3	○ 4
g) I listen in mathematics class.	○ 1	○ 2	○ 3	○ 4
h) I avoid distractions when I am studying mathematics.	○ 1	○ 2	○ 3	○ 4
i) I keep my mathematics work well organized.	○ 1	○ 2	○ 3	○ 4

Q37 For each pair of statements, please choose the item that best describes you.

a) Please darken only one of the following two circles.

○$_1$ I intend to take additional mathematics courses after I finish high school.

○$_2$ I intend to take additional English courses after I finish high school.

b) Please darken only one of the following two circles.

○$_1$ I plan on majoring in a subject in college that requires mathematics skills.

○$_2$ I plan on majoring in a subject in college that requires science skills.

c) Please darken only one of the following two circles.

○$_1$ I am willing to study harder in my mathematics classes than is required.

○$_2$ I am willing to study harder in my English classes than is required.

d) Please darken only one of the following two circles.

○$_1$ I plan on taking as many mathematics classes as I can during my education.

○$_2$ I plan on taking as many science classes as I can during my education.

e) Please darken only one of the following two circles.

○$_1$ I am planning on pursuing a career that involves a lot of mathematics

○$_2$ I am planning on pursuing a career that involves a lot of science.

Q38 How often do you do the following things at school and outside of school?

(Please darken only one circle in each row.)

	Always or almost always	*Often*	*Sometimes*	*Never or rarely*
a) I talk about mathematics problems with my friends.	○ 1	○ 2	○ 3	○ 4
b) I help my friends with mathematics.	○ 1	○ 2	○ 3	○ 4
c) I do mathematics as an extracurricular activity.	○ 1	○ 2	○ 3	○ 4
d) I take part in mathematics competitions.	○ 1	○ 2	○ 3	○ 4
e) I do mathematics more than 2 hours a day outside of school.	○ 1	○ 2	○ 3	○ 4
f) I play chess.	○ 1	○ 2	○ 3	○ 4
g) I program computers.	○ 1	○ 2	○ 3	○ 4
h) I participate in a mathematics club.	○ 1	○ 2	○ 3	○ 4

SECTION D: ABOUT YOUR PROBLEM SOLVING EXPERIENCES

ST93

Q39 How well does each of the following statements describe you?

(Please darken only one circle in each row.)

	Very much like me	*Mostly like me*	*Somewhat like me*	*Not much like me*	*Not at all like me*
a) When confronted with a problem I give up easily.	○1	○2	○3	○4	○5
b) I put off difficult problems.	○1	○2	○3	○4	○5
c) I remain interested in the tasks that I start.	○1	○2	○3	○4	○5
d) I continue working on tasks until everything is perfect.	○1	○2	○3	○4	○5
e) When confronted with a problem I do more than what is expected of me.	○1	○2	○3	○4	○5

Q40 How well does each of the following statements describe you?

(Please darken only one circle in each row.)

	Very much like me	*Mostly like me*	*Somewhat like me*	*Not much like me*	*Not at all like me*
a) I can handle a lot of information.	○ 1	○ 2	○ 3	○ 4	○ 5
b) I am quick to understand things.	○ 1	○ 2	○ 3	○ 4	○ 5
c) I seek explanations for things.	○ 1	○ 2	○ 3	○ 4	○ 5
d) I can easily link facts together.	○ 1	○ 2	○ 3	○ 4	○ 5
e) I like to solve complex problems.	○ 1	○ 2	○ 3	○ 4	○ 5

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

Q42 *Suppose that you are planning a trip to the zoo with your brother. You don't know which route to take to get there.*

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 read the zoo brochure to see if it says how to get there.	○ 1	○ 2	○ 3	○ 4
b) I study a map and figure out the best route.	○ 1	○ 2	○ 3	○ 4
c) I leave it to my brother to worry about how to get there.	○ 1	○ 2	○ 3	○ 4
d) I know roughly where it is, so I suggest we just start driving.	○ 1	○ 2	○ 3	○ 4

Q43 *Suppose that you arrive at the train station. There is a ticket machine that you have never used before. You want to buy a ticket.*

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 check how similar it is to other ticket machines I have used.	○1	○2	○3	○4
b) I try out all the buttons to see what happens.	○1	○2	○3	○4
c) I ask someone for help.	○1	○2	○3	○4
d) I try to find a ticket office at the station to buy a ticket.	○1	○2	○3	○4

SECTION E: MORE QUESTIONS ABOUT LEARNING MATHEMATICS

ST53

Q44 For each group of three items, please choose the item that best describes your approach to mathematics.

a) Please darken only one of the following three circles.

○1 When I study for a mathematics test, I try to figure out what are the most important parts to learn.

○2 When I study for a mathematics test, I try to understand new concepts by relating them to things I already know.

○3 When I study for a mathematics test, I learn as much as I can by heart.

b) Please darken only one of the following three circles.

○1 When I study mathematics, I try to figure out which concepts I still do not understand completely.

○2 When I study mathematics, I think of new ways to get the answer.

○3 When I study mathematics, I make myself check to see if I remember the work I have already done.

c) Please darken only one of the following three circles.

○1 When I study mathematics, I try to relate the work to things I have learned in other subjects.

○2 When I study mathematics, I start by working out exactly what I need to learn.

○3 When I study mathematics, I go over some problems so often that I feel as if I could solve them in my sleep.

d) Please darken only one of the following three circles.

○1 In order to remember the method for solving a mathematics problem, I go through examples again and again.

○2 I think about how the mathematics I have learned can be used in everyday life.

○3 When I cannot understand something in mathematics, I always search for more information to clarify the problem.

Q45 How many hours do you typically spend per week attending out-of-school-time lessons in the following subjects?

These are only lessons in subjects that you are also learning at school, that you spend extra time learning outside of normal school hours. The lessons may be given at your school, at your home or somewhere else.

(Please darken only one circle in each row.)

	I do not attend out-of-school-time lessons in this subject	*Less than 2 hours a week*	*2 or more but less than 4 hours a week*	*4 or more but less than 6 hours a week*	*6 or more hours a week*
a) English (e.g., language arts, literature, writing)	○ 1	○ 2	○ 3	○ 4	○ 5
b) Mathematics	○ 1	○ 2	○ 3	○ 4	○ 5
c) Science	○ 1	○ 2	○ 3	○ 4	○ 5
d) Other subjects	○ 1	○ 2	○ 3	○ 4	○ 5

Q46 Thinking about all school subjects: on average, how many hours do you spend each week on the following?

When answering, include time spent on the weekend too.

a) Homework or other material assigned by your teachers _____ *hours per week*

b) Out of the time spent in (a), how many hours do you work on your homework with somebody overlooking and providing help if necessary ("guided homework"), either at school or elsewhere? _____ *hours per week*

c) Work with a personal tutor (whether paid or not) _____ *hours per week*

d) Attend out of school classes organized by a commercial company and paid for by your parents _____ *hours per week*

e) Study with a parent or other family member _____ *hours per week*

f) Practice content from school lessons by working on a computer (e.g., learn vocabulary with training software) _____ *hours per week*

Q47 How often have you encountered the following types of mathematics tasks during your time at school?

(Please darken only one circle in each row.)

	Frequently	*Sometimes*	*Rarely*	*Never*
a) Figuring out from a train schedule how long it would take to get from one place to another.	○ 1	○ 2	○ 3	○ 4
b) Calculating how much more expensive a computer would be after adding tax.	○ 1	○ 2	○ 3	○ 4
c) Calculating how many square feet of tile you need to cover a floor.	○ 1	○ 2	○ 3	○ 4
d) Understanding scientific tables presented in an article.	○ 1	○ 2	○ 3	○ 4
e) Solving an equation like $6x^2 + 5 = 29$.	○ 1	○ 2	○ 3	○ 4
f) Finding the actual distance between two places on a map with a 1:10,000 scale.	○ 1	○ 2	○ 3	○ 4
g) Solving an equation like $2(x+3) = (x + 3)(x - 3)$.	○ 1	○ 2	○ 3	○ 4
h) Calculating the power consumption of an electronic appliance per week.	○ 1	○ 2	○ 3	○ 4
i) Solving an equation like $3x + 5 = 17$.	○ 1	○ 2	○ 3	○ 4

Q48 Thinking about mathematical concepts: how familiar are you with the following terms?

(Please darken only one circle in each row.)

	Never heard of it	*Heard of it once or twice*	*Heard of it a few times*	*Heard of it often*	*Know it well, understand the concept*
a) Exponential Function	○ 1	○ 2	○ 3	○ 4	○ 5
b) Divisor	○ 1	○ 2	○ 3	○ 4	○ 5
c) Quadratic Function	○ 1	○ 2	○ 3	○ 4	○ 5
d) Proper Number	○ 1	○ 2	○ 3	○ 4	○ 5
e) Linear Equation	○ 1	○ 2	○ 3	○ 4	○ 5
f) Vectors	○ 1	○ 2	○ 3	○ 4	○ 5
g) Complex Number	○ 1	○ 2	○ 3	○ 4	○ 5
h) Rational Number	○ 1	○ 2	○ 3	○ 4	○ 5
i) Radicals	○ 1	○ 2	○ 3	○ 4	○ 5
j) Subjunctive Scaling	○ 1	○ 2	○ 3	○ 4	○ 5
k) Polygon	○ 1	○ 2	○ 3	○ 4	○ 5
l) Declarative Fraction	○ 1	○ 2	○ 3	○ 4	○ 5
m) Congruent Figure	○ 1	○ 2	○ 3	○ 4	○ 5
n) Cosine	○ 1	○ 2	○ 3	○ 4	○ 5
o) Arithmetic Mean	○ 1	○ 2	○ 3	○ 4	○ 5
p) Probability	○ 1	○ 2	○ 3	○ 4	○ 5

ST69

Q49 How many minutes, on average, are there in a class period for the following subjects?

a) Minutes in a class period in English (e.g., language arts, literature, writing): _____ *minutes*

b) Minutes in a class period in mathematics: _____ *minutes*

c) Minutes in a class period in science: _____ *minutes*

ST70

Q50 How many class periods per week do you typically have for the following subjects?

a) Number of class periods per week in English (e.g., language arts, literature, writing): _____ *class periods*

b) Number of class periods per week in mathematics: _____ *class periods*

c) Number of class periods per week in science: _____ *class periods*

ST71

Q51 In a normal, full week at school, how many class periods do you have in total?

Number of **ALL** class periods _____ *class periods*

ST72

Q52 On average, about how many students attend your English class (e.g., language arts, literature, writing)?

_____ *students*

The next four questions are about your experience with different kinds of mathematics problems at school. You will see descriptions of problems and gray-colored boxes, each containing a mathematics problem.

Please read each problem. You do NOT need to solve it.

ST73

Q53 *In this box are two problems. Each requires you to understand a problem written in text and perform the appropriate calculations. Usually the problem talks about practical situations, but the numbers, people and places mentioned are made up. All the information you need is given. Here are two examples.*

1) Ann is two years older than Betty and Betty is four times as old as Sam. When Betty is 30, how old is Sam?

2) Mr. Smith bought a television and a bed. The television cost $625 but he got a 10% discount. The bed cost $200. He paid $20 for delivery. How much money did Mr. Smith spend?

We want to know your experience with this type of word problem at school. Do not solve them!

(Please darken only one circle in each row.)

	Frequently	*Sometimes*	*Rarely*	*Never*
a) How often have you encountered this type of problem in your **mathematics lessons**?	○ 1	○ 2	○ 3	○ 4
b) How often have you encountered this type of problem in the **tests you have taken at school?**	○ 1	○ 2	○ 3	○ 4

Q54 *Below are examples of another type of mathematics problem.*

1) Solve 2x + 3 = 7.
2) Find the volume of a box with sides 3m, 4m and 5m.

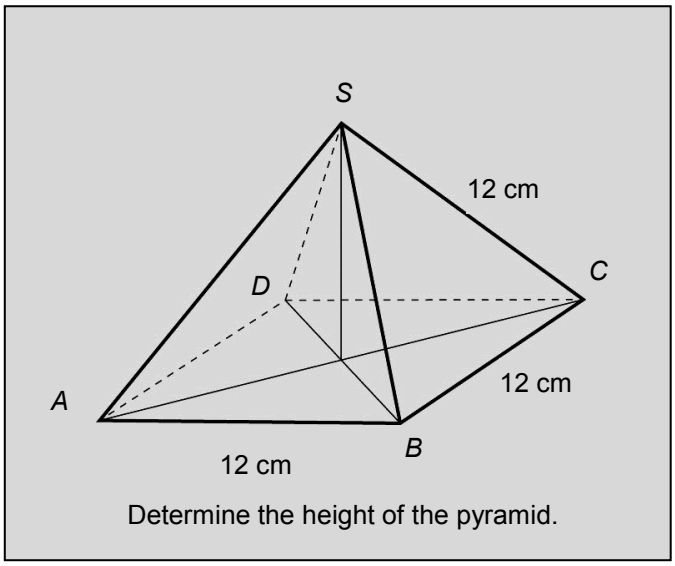
We want to know about your experience with this type of problem at school. Do not solve them!

(Please darken only one circle in each row.)

	Frequently	*Sometimes*	*Rarely*	*Never*
a) How often have you encountered this type of problem in your **mathematics lessons**?	○ 1	○ 2	○ 3	○ 4
b) How often have you encountered this type of problem in the **tests you have taken at school?**	○ 1	○ 2	○ 3	○ 4

Q55 *In the next type of problem, you have to use mathematical knowledge and draw conclusions. There is no practical application provided. Here are two examples.*

1) Here you need to use geometrical theorems:



Determine the height of the pyramid.

2) Here you have to know what a prime number is:

If n is any number: can $(n+1)^2$ be a prime number?

We want to know about your experience with this type of problem at school. Do not solve them!

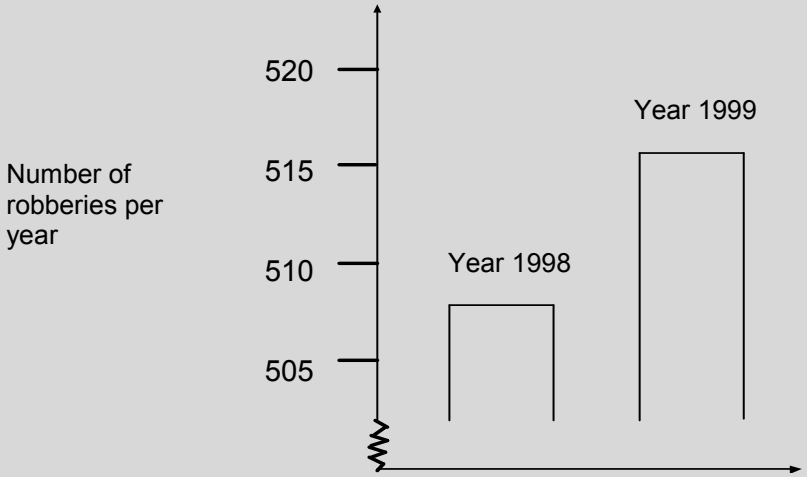
(Please darken only one circle in each row.)

	Frequently	*Sometimes*	*Rarely*	*Never*
a) How often have you encountered this type of problem in your **mathematics lessons**?	○ 1	○ 2	○ 3	○ 4
b) How often have you encountered this type of problem in the **tests you have taken at school?**	○ 1	○ 2	○ 3	○ 4

Q56 *In this type of problem, you have to apply suitable mathematical knowledge to find a useful answer to a problem that arises in everyday life or work. The data and information are about real situations. Here are two examples.*

Example 1:

A TV reporter says "This graph shows that there is a huge increase in the number of robberies from 1998 to 1999."



Do you consider the reporter's statement to be a reasonable interpretation of the graph? Give an explanation to support your answer.

Example 2:

For years the relationship between a person's recommended maximum heart rate and the person's age was described by the following formula:

Recommended maximum heart rate = 220 - *age*

Recent research showed that this formula should be modified slightly. The new formula is as follows:

Recommended maximum heart rate = 208 – (0.7 × *age*)

From which age onwards does the recommended maximum heart rate increase as a result of the introduction of the new formula? Show your work.

We want to know about your experience with this type of problem at school. Do not solve them!

(Please darken only one circle in each row.)

	Frequently	*Sometimes*	*Rarely*	*Never*
a) How often have you encountered this type of problem in your **mathematics lessons**?	○ 1	○ 2	○ 3	○ 4
b) How often have you encountered this type of problem in the **tests you have taken at school?**	○ 1	○ 2	○ 3	○ 4

Thank you very much for your cooperation in completing this questionnaire!

OECD Program for International Student Assessment 2012

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

USA

Date of Test (Main Survey PISA 2012)		
Month	Day	2012

School Questionnaire

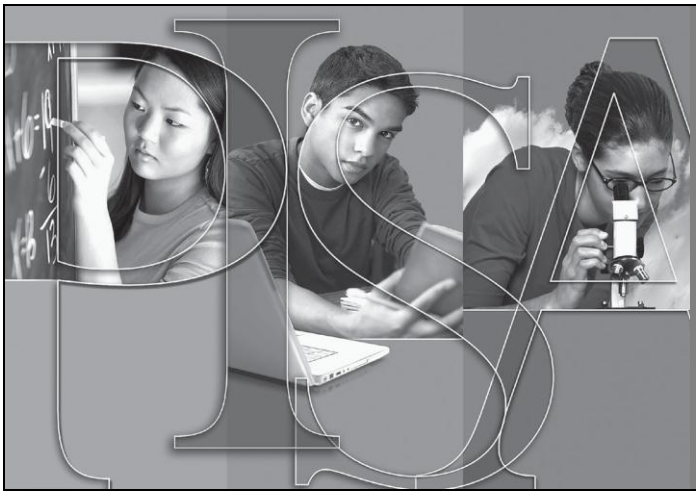
Place Label Here

School ID

English 313

U.S. participation in this study is sponsored by the National Center for Education Statistics (NCES), U.S. Department of Education. Your responses are protected by federal statute (20 U.S.C., § 9573). Your answers may be used only for statistical purposes and may not be disclosed, or used, in identifiable form for any other purpose except as required by law. By law, everyone working on this NCES survey is subject to a jail term of up to 5 years, a fine of up to $250,000, or both if he or she willfully discloses ANY identifiable information about you.

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. The valid OMB control number for this voluntary information collection is 1850-0755. The time required to complete this information collection is estimated to average 30 minutes per response, including the time to review instructions, search existing data resources, gather the data needed, and complete and review the information collection. **If you have any comments concerning the accuracy of the time estimate(s) or suggestions for improving the form, please write to**: U.S. Department of Education, Washington, D.C. 20202-4537. **If you have comments or concerns regarding the status of your individual submission of this form, write directly to**: Program for International Student Assessment (PISA), National Center for Education Statistics, U.S. Department of Education, 1990 K Street, N.W., Washington, D.C. 20006. O.M.B. No. 1850-0755, Approval Expires 11/30/2013



Project Consortium

- Australian Council for Educational Research (ACER)
- cApStAn Linguistic Quality Control (Belgium)
- Deutsches Institut für Internationale Pädagogische Forschung (DIPF, Germany)
- Educational Testing Service (ETS, USA)
- Institutt for Lærerutdanning og Skoleutvikling (ILS, Norway)
- Leibniz - Institute for Science and Mathematics Education (IPN, Germany)
- National Institute for Educational Policy Research (NIER, Japan)
- The Tao Initiative: CRP - Henri Tudor and Université de Luxembourg EMACS (Luxembourg)
- Unité d'analyse des systèmes et des pratiques d'enseignement (aSPe, Belgium)
- Westat (USA)

This questionnaire asks for information including:

- The structure and organization of the school;
- The student body and teachers;
- The school's resources;
- The school's instruction, curriculum and assessment;
- The school climate;
- The school's policies and practices;
- Financial education at school.

This information helps illustrate the similarities and differences between groups of schools in order to better establish the context for students' test results. For example, the information provided may help to establish what effect the availability of resources may have on student achievement – both within and between countries.

The questionnaire should be completed by the principal or their designee. It should take about 30 minutes to complete.

For some questions, specific expertise may be needed. You may consult experts to help you answer these questions.

If you do not know an answer precisely, your best estimate will be adequate for the purposes of the study.

Some questions ask about 10th grade or 10th graders. If you do not have a 10th grade in your school, then answer these questions for the grade in your school that contains the most 15-year-olds.

Your answers will be combined with answers from other principals to calculate totals and averages. All information (or responses) you provide may only be used for statistical purposes and may not be disclosed, or used, in identifiable form for any other purpose except as required by law [Education Science Reform Act of 2002 (ESRA 2002), 20 U.S. Code, Section 9573].

SECTION A: THE STRUCTURE AND ORGANIZATION OF THE SCHOOL

SC01

Q1 Is your school a public or a private school?

(Please check only one box.)

A public school ☐ 1

A private school ☐ 2

SC02

Q2 About what percentage of your total funding for a typical school year comes from the following sources?

(Please write a number in each row. Write 0 (zero) if no funding comes from that source.)

	%
a) Government (includes local, state and federal)	_____
b) Tuition, student fees or school charges paid by parents	_____
c) Benefactors, donations, bequests, sponsorships, parent fundraising	_____
d) Other	_____
Total	100%

SC03

Q3 Which of the following definitions best describes the community in which your school is located?

(Please check only one box.)

- A village, hamlet, or rural area (fewer than 3,000 people) ☐ 1
- A small town (3,000 to about 15,000 people) ☐ 2
- A town (15,000 to about 100,000 people) ☐ 3
- A city (100,000 to about 1,000,000 people) ☐ 4
- A large city (with over 1,000,000 people) ☐ 5

SC04

Q4 *We are interested in the options parents have when choosing a school for their children.*

Which of the following statements best describes the schooling available to students in your location?

(Please check only one box.)

- There are two or more other schools in this area that compete for our students. ☐ 1
- There is one other school in this area that competes for our students. ☐ 2
- There are no other schools in this area that compete for our students. ☐ 3

Q5 What is the average size of English classes in the 10th grade in your school?

(Please check only one box.)

15 students or fewer ☐ 01

16-20 students ☐ 02

21-25 students ☐ 03

26-30 students ☐ 04

31-35 students ☐ 05

36-40 students ☐ 06

41-45 students ☐ 07

46-50 students ☐ 08

More than 50 students ☐ 09

SECTION B: THE STUDENT BODY AND TEACHERS

SC07

Q6 As of September 1, 2012, what was the total school enrollment (number of students)?

(Please write a number on each line. Write 0 (zero) if there are none.)

a) Number of boys: _____

b) Number of girls: _____

USA_SC53A01

Q7 Approximately what percentage of students at this school last year were eligible for free- or reduced-price lunches through the National School Lunch Program?

(*Please write a number on the line. Write 0 (zero) if there are none.)*

Percentage of students eligible _____ %

USA_SC54A01

Q8 About how many students in the 10th grade in your school have a first language that is not English?

(Please check only one box.)

60% or more ☐ 1

40% or more but less than 60% ☐ 2

20% or more but less than 40% ☐ 3

10% or more but less than 20% ☐ 4

More than 0% but less than 10% ☐ 5

None ☐ 6

Q9 How many of the following teachers are on the staff of your school?

Include both full-time and part-time teachers. *A full-time teacher is employed at least 90% of the time as a teacher for the full school year. All other teachers should be considered part time.*

(Please write a number in each space provided. Write 0 (zero) if there are none.)

	Full-time	*Part-time*
a) Teachers in TOTAL	_____	_____
b) Teachers fully certified by the state in their main assignment field	_____	_____
c) Teachers with at least a bachelor's degree	_____	_____

Q10 How many of the following are on the mathematics staff of your school?

Include both full-time and part-time teachers. *A full-time teacher is employed at least 90% of the time as a teacher for the full school year. All other teachers should be considered part time.*

Please count only those teachers who have taught or will teach mathematics during the current school year.

(Please write a number in each space provided. Write 0 (zero) if there are none.)

	Full-time	*Part-time*
a) Teachers of mathematics in TOTAL	_____	_____
b) Teachers of mathematics with a bachelor's or master's degree	_____	_____
c) Teachers of mathematics with a bachelor's or master's degree with a major in mathematics, statistics, physics, or engineering	_____	_____
d) Teachers of mathematics with a bachelor's or master's degree in education	_____	_____
e) Teachers of mathematics with an associate's degree but not a bachelor's or master's degree	_____	_____

SECTION C: THE SCHOOL'S RESOURCES

The goal of the following set of three questions is to gather information about the student-computer ratio for students in the 10th grade at your school.

SC11

Number

Q11a At your school, what is the total number of students in the 10th grade? _____

Q11b Approximately how many computers are available for these students for educational purposes? _____

Q11c Approximately how many of these computers are connected to the Internet? _____

SC13

Q12 In all subjects taken together, for how much of their work does the school expect 10th-grade students to access the Internet?

(Please check only one box in each row.)

	<10%	*10-25%*	*26-50%*	*51-75%*	*>75%*
a) Work during class	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b) Homework	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c) Assignments or projects	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q13 Is your school's capacity to provide instruction hindered by any of the following issues?

(Please check one box in each row.)

	Not at all	*Very little*	*To some extent*	*A lot*
a) A lack of qualified science teachers	☐ 1	☐ 2	☐ 3	☐ 4
b) A lack of qualified mathematics teachers	☐ 1	☐ 2	☐ 3	☐ 4
c) A lack of qualified English teachers	☐ 1	☐ 2	☐ 3	☐ 4
d) A lack of qualified teachers of other subjects	☐ 1	☐ 2	☐ 3	☐ 4
e) Shortage or inadequacy of science laboratory equipment	☐ 1	☐ 2	☐ 3	☐ 4
f) Shortage or inadequacy of instructional materials (e.g., textbooks)	☐ 1	☐ 2	☐ 3	☐ 4
g) Shortage or inadequacy of computers for instruction	☐ 1	☐ 2	☐ 3	☐ 4
h) Lack or inadequacy of Internet connectivity	☐ 1	☐ 2	☐ 3	☐ 4
i) Shortage or inadequacy of computer software for instruction	☐ 1	☐ 2	☐ 3	☐ 4
j) Shortage or inadequacy of library materials	☐ 1	☐ 2	☐ 3	☐ 4
k) Shortage or inadequacy of school buildings and grounds	☐ 1	☐ 2	☐ 3	☐ 4
l) Shortage or inadequacy of heating/cooling and lighting systems	☐ 1	☐ 2	☐ 3	☐ 4
m) Shortage or inadequacy of instructional space (e.g., classrooms)	☐ 1	☐ 2	☐ 3	☐ 4

SECTION D: SCHOOL INSTRUCTION, CURRICULUM, AND ASSESSMENT

SC15

Q14 *Schools sometimes organize instruction differently for students with different abilities and interests in mathematics.* **Which of the following options describe what your school does for 10th-grade students in mathematics classes?**

(Please check one box in each row.)

	For all classes	*For some classes*	*Not for any classes*
a) Mathematics classes study similar content, but at different levels of difficulty.	☐ 1	☐ 2	☐ 3
b) Different classes study different content or sets of mathematics topics that have different levels of difficulty.	☐ 1	☐ 2	☐ 3
c) Students are grouped by ability within their mathematics classes.	☐ 1	☐ 2	☐ 3
d) In mathematics classes, teachers use pedagogy suitable for students with heterogeneous abilities (i.e., students are not grouped by ability).	☐ 1	☐ 2	☐ 3

Q15 In this academic year, which of the following activities does your school offer to students in the 10th grade?

(Please check one box in each row.)

	Yes	*No*
a) Band, orchestra or choir	☐ 1	☐ 2
b) School play or school musical	☐ 1	☐ 2
c) School yearbook, newspaper or magazine	☐ 1	☐ 2
d) Volunteering or community service activities	☐ 1	☐ 2
e) Mathematics club	☐ 1	☐ 2
f) Mathematics competitions	☐ 1	☐ 2
g) Chess club	☐ 1	☐ 2
h) Club with a focus on computers/ Information and Communication Technology	☐ 1	☐ 2
i) Art club or art activities	☐ 1	☐ 2
j) Sports team or sports activities	☐ 1	☐ 2

Q16 In your school, are assessments of students in the 10th grade used for any of the following purposes?

(Please check only one box in each row.)

	Yes	*No*
a) To inform parents about their child's progress	☐ 1	☐ 2
b) To make decisions about students' retention or promotion	☐ 1	☐ 2
c) To group students for instructional purposes	☐ 1	☐ 2
d) To compare the school to district, state, or national performance	☐ 1	☐ 2
e) To monitor the school's progress from year to year	☐ 1	☐ 2
f) To make judgments about teachers' effectiveness	☐ 1	☐ 2
g) To identify aspects of instruction or the curriculum that could be improved	☐ 1	☐ 2
h) To compare the school with other schools	☐ 1	☐ 2

Q17 In your school, are achievement data used in any of the following accountability procedures?

*Achievement data include **aggregated** school or grade-level test scores or grades, or graduation rates.*

(Please check one box in each row.)

	Yes	*No*
a) Achievement data are posted publicly (e.g., in the media)	☐ 1	☐ 2
b) Achievement data are tracked over time by an administrative authority, such as a district, state, or national education agency	☐ 1	☐ 2

SC20

Q18 Does your school offer mathematics lessons or classes in addition to the mathematics classes offered during the usual school hours?

(Please check only one box.)

Yes	☐ 1	*go to the next question*
No	☐ 2	*go to Q20*

SC21

Q19 What is the purpose of these additional mathematics lessons or classes?

(Please check only one box.)

Enrichment mathematics only	☐ 1
Remedial mathematics only	☐ 2
Both enrichment mathematics and remedial mathematics	☐ 3
Without differentiation depending on the prior achievement level of the students	☐ 4

SECTION E: SCHOOL CLIMATE

SC22

Q20 In your school, to what extent is the learning of students hindered by the following phenomena?

(Please check one box in each row.)

	Not at all	*Very little*	*To some extent*	*A lot*
a) Student truancy	☐ 1	☐ 2	☐ 3	☐ 4
b) Students skipping classes	☐ 1	☐ 2	☐ 3	☐ 4
c) Students arriving late for school	☐ 1	☐ 2	☐ 3	☐ 4
d) Students not attending compulsory school events (e.g., school assemblies) or excursions	☐ 1	☐ 2	☐ 3	☐ 4
e) Students lacking respect for teachers	☐ 1	☐ 2	☐ 3	☐ 4
f) Disruption of classes by students	☐ 1	☐ 2	☐ 3	☐ 4
g) Student use of alcohol or illegal drugs	☐ 1	☐ 2	☐ 3	☐ 4
h) Students intimidating or bullying other students	☐ 1	☐ 2	☐ 3	☐ 4
i) Students not being encouraged to achieve their full potential	☐ 1	☐ 2	☐ 3	☐ 4
j) Poor student-teacher relations	☐ 1	☐ 2	☐ 3	☐ 4
k) Teachers having to teach students of heterogeneous ability levels within the same class	☐ 1	☐ 2	☐ 3	☐ 4
l) Teachers having to teach students of diverse ethnic backgrounds (i.e., language, culture) within the same class	☐ 1	☐ 2	☐ 3	☐ 4
m) Teachers' low expectations of students	☐ 1	☐ 2	☐ 3	☐ 4

	Not at all	*Very little*	*To some extent*	*A lot*
n) Teachers not meeting individual students' needs	☐ 1	☐ 2	☐ 3	☐ 4
o) Teacher absenteeism	☐ 1	☐ 2	☐ 3	☐ 4
p) Staff resisting change	☐ 1	☐ 2	☐ 3	☐ 4
q) Teachers being too strict with students	☐ 1	☐ 2	☐ 3	☐ 4
r) Teachers being late for classes	☐ 1	☐ 2	☐ 3	☐ 4
s) Teachers not being well prepared for classes	☐ 1	☐ 2	☐ 3	☐ 4

SC23

Q21 During the 2011-2012 academic year, what proportion of students left your school without a diploma or an alternative credential (e.g., a GED)?

Only include students who dropped out of school without a diploma or alternative credential (e.g., a GED), not students who moved or transferred to another school.

_____ %

SC24

Q22 Which statement below best characterizes parental expectations towards your school?

(Please check only one box.)

There is *constant pressure* from many parents who expect our school to set very high academic standards and to have our students achieve them. ☐ 1

Pressure on the school to achieve higher academic standards among students comes from a *minority of parents.* ☐ 2

Pressure from parents on the school to achieve higher academic standards among students is *largely absent.* ☐ 3

Q23 During the 2011-2012 academic year, what proportion of students' parents participated in the following school-related activities?

(Please write a number in each row. Write 0 (zero) if no parents participated in the activity. Write 100 (one hundred) if all parents participated in the activity.)

	%
a) Discussed their child's behavior with a teacher on their own initiative.	_____
b) Discussed their child's behavior on the initiative of one of their child's teachers.	_____
c) Discussed their child's academic progress with a teacher on their own initiative.	_____
d) Discussed their child's academic progress on the initiative of one of their child's teachers.	_____
e) Volunteered for physical activities (e.g., building maintenance, carpentry, gardening or yard work).	_____
f) Volunteered for extra-curricular activities (e.g., book club, school play, sports, field trip).	_____
g) Volunteered in the school library or media center.	_____
h) Assisted a teacher in the school.	_____
i) Appeared as a guest speaker.	_____
j) Participated in local school government (e.g., parent teacher association or parent advisory council).	_____
k) Assisted in fundraising for the school.	_____
l) Volunteered in the school cafeteria.	_____

Q24 Think about the teachers in your school. How much do you agree with the following statements?

(Please check one box in each row.)

	Strongly agree	*Agree*	*Disagree*	*Strongly disagree*
a) The morale of teachers in this school is high.	☐ 1	☐ 2	☐ 3	☐ 4
b) Teachers work with enthusiasm.	☐ 1	☐ 2	☐ 3	☐ 4
c) Teachers take pride in this school.	☐ 1	☐ 2	☐ 3	☐ 4
d) Teachers value academic achievement.	☐ 1	☐ 2	☐ 3	☐ 4

SC27

Q25 How much do you agree with these statements about teachers in your school?

(Please check one box in each row.)

	Strongly agree	*Agree*	*Disagree*	*Strongly disagree*
a) Mathematics teachers are interested in trying new methods and teaching practices.	☐ 1	☐ 2	☐ 3	☐ 4
b) There is a preference among mathematics teachers to stay with well-known methods and practices.	☐ 1	☐ 2	☐ 3	☐ 4

SC28

c) There is consensus among mathematics teachers that academic achievement must be kept as high as possible.	☐ 1	☐ 2	☐ 3	☐ 4
d) There is consensus among mathematics teachers that it is best to adapt academic standards to the students' levels and needs.	☐ 1	☐ 2	☐ 3	☐ 4

SC29

e) There is consensus among mathematics teachers that the social and emotional development of the students is as important as their acquisition of mathematical skills and knowledge in mathematics classes.	☐ 1	☐ 2	☐ 3	☐ 4
f) There is consensus among mathematics teachers that the development of mathematical skills and knowledge in students is the most important objective in mathematics classes.	☐ 1	☐ 2	☐ 3	☐ 4

Q26 During the last year, have any of the following methods been used to monitor the practice of <u>mathematics</u> teachers at your school?

(Please check one box in each row.)

	Yes	*No*
a) Tests or assessments of student achievement	☐ 1	☐ 2
b) Teacher peer review (of lesson plans, assessment instruments, lessons)	☐ 1	☐ 2
c) Principal or senior staff observations of lessons	☐ 1	☐ 2
d) Observation of classes by inspectors or other persons external to the school	☐ 1	☐ 2

Q27 To what extent have appraisals of and/or feedback to teachers directly led to the following?

(Please check one box in each row.)

	No change	*A small change*	*A moderate change*	*A large change*
a) A change in salary	☐ 1	☐ 2	☐ 3	☐ 4
b) A financial bonus or another kind of monetary reward	☐ 1	☐ 2	☐ 3	☐ 4
c) Opportunities for professional development activities	☐ 1	☐ 2	☐ 3	☐ 4
d) A change in the likelihood of career advancement	☐ 1	☐ 2	☐ 3	☐ 4
e) Public recognition from you	☐ 1	☐ 2	☐ 3	☐ 4
f) Changes in work responsibilities that make the job more attractive	☐ 1	☐ 2	☐ 3	☐ 4
g) A role in school development initiatives (e.g., curriculum development group, development of school objectives)	☐ 1	☐ 2	☐ 3	☐ 4

SECTION F: SCHOOL POLICIES AND PRACTICES

SC32

Q28 How often are the following factors considered when students are admitted to your school?

(Please check one box in each row.)

	Never	*Sometimes*	*Always*
a) Student's record of academic performance (including placement tests)	☐ 1	☐ 2	☐ 3
b) Recommendation of feeder schools	☐ 1	☐ 2	☐ 3
c) Parents' endorsement of the instructional or religious philosophy of the school	☐ 1	☐ 2	☐ 3
d) Whether the student requires or is interested in a special program	☐ 1	☐ 2	☐ 3
e) Preference given to family members of current or former students	☐ 1	☐ 2	☐ 3
f) Residence in a particular area	☐ 1	☐ 2	☐ 3
g) Other	☐ 1	☐ 2	☐ 3

Q29 Regarding your school, who has a considerable responsibility for the following tasks?

(Please check as many boxes as appropriate in each row.)

	Principal	*Teachers*	*School-level governing board*	*Local education agency*	*State education agency*	*U.S. Dept. of Education*
a) Selecting teachers for hire	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1
b) Firing teachers	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1
c) Establishing teachers' starting salaries	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1
d) Determining teachers' salary increases	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1
e) Formulating the school budget	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1
f) Deciding on budget allocations within the school	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1
g) Establishing student disciplinary policies	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1
h) Establishing student assessment policies	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1
i) Approving students for admission to the school	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1
j) Choosing which textbooks are used	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1
k) Determining course content	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1
l) Deciding which courses are offered	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1	☐ 1

Q30 Below are statements about your management of this school. Please indicate the frequency of the following activities and behaviors in your school during the 2011-2012 academic year.

(Please check only one box in each row.)

	Did not occur	*1-2 times during the year*	*3-4 times during the year*	*Once a month*	*Once a week*	*More than once a week*
a) I work to enhance the school's reputation in the community.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
b) I use student performance results to develop the school's educational goals.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
c) I make sure that the professional development activities of teachers are in accordance with the teaching goals of the school.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
d) I ensure that teachers work according to the school's educational goals.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
e) I promote teaching practices based on recent educational research.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
f) I praise teachers whose students are actively participating in learning.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
g) When a teacher has problems in his/her classroom, I take the initiative to discuss matters.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
h) I draw teachers' attention to the importance of students' development of critical and social capacities.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

	Did not occur	*1-2 times during the year*	*3-4 times during the year*	*Once a month*	*Once a week*	*More than once a week*
i) I pay attention to disruptive behavior in classrooms.	☐1	☐2	☐3	☐4	☐5	☐6
j) I provide staff with opportunities to participate in school decision-making.	☐1	☐2	☐3	☐4	☐5	☐6
k) I engage teachers to help build a school culture of continuous improvement.	☐1	☐2	☐3	☐4	☐5	☐6
l) I ask teachers to participate in reviewing management practices.	☐1	☐2	☐3	☐4	☐5	☐6
m) When a teacher brings up a classroom problem, we solve the problem together.	☐1	☐2	☐3	☐4	☐5	☐6
n) I discuss the school's academic goals with teachers at faculty meetings.	☐1	☐2	☐3	☐4	☐5	☐6
o) I refer to the school's academic goals when making curricular decisions with teachers.	☐1	☐2	☐3	☐4	☐5	☐6
p) I discuss academic performance results with the faculty to identify curricular strengths and weaknesses.	☐1	☐2	☐3	☐4	☐5	☐6
q) I lead or attend in-service activities concerned with instruction.	☐1	☐2	☐3	☐4	☐5	☐6
r) I set aside time at faculty meetings for teachers to share ideas or information from in-service activities.	☐1	☐2	☐3	☐4	☐5	☐6

***Q30 Continued*...Below are statements about your management of this school. Please indicate the frequency of the following activities and behaviors in your school during <u>the 2011-2012 academic year</u>.**

(Please check only one box in each row.)

	Did not occur	*1-2 times during the year*	*3-4 times during the year*	*Once a month*	*Once a week*	*More than once a week*
s) I conduct informal observations in classrooms on a regular basis (informal observations are unscheduled, last at least 5 minutes, and may or may not involve written feedback or a formal conference).	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
t) I review work produced by students when evaluating classroom instruction.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
u) I evaluate the performance of staff.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

SC35

Q31 During the last three months, what percentage of teaching staff in your school has attended a program of professional development with a focus on mathematics?

A program of professional development here is a formal program designed to enhance teaching skills or pedagogical practices. It may or may not lead to a recognized qualification. The program must last for at least one day in total and have a focus on mathematics teaching and education.

a) All staff at your school _____ %

b) Staff who teach mathematics at your school _____ %

Q32 Which of the following measures aimed at quality assurance and improvement do you have in your school?

(Please check one box in each row.)

	Yes	*No*
a) Written specification of the school's curricular profile and educational goals	☐ 1	☐ 2
b) Written specification of student performance standards	☐ 1	☐ 2
c) Systematic recording of data including teacher and student attendance and graduation rates, test results and professional development of teachers	☐ 1	☐ 2
d) Internal evaluation/self-evaluation	☐ 1	☐ 2
e) External evaluation	☐ 1	☐ 2
f) Seeking written feedback from students (e.g., regarding classes, teachers or resources)	☐ 1	☐ 2
g) Teacher mentoring	☐ 1	☐ 2
h) Regular consultation aimed at school improvement with one or more experts over a period of at least six months	☐ 1	☐ 2
i) Implementation of a standardized policy for mathematics (i.e., school curriculum with shared instructional materials accompanied by staff development and training)	☐ 1	☐ 2

Q33 Which of the following statements apply in your school?

A policy refers to written rules known to those concerned with the policy.

(Please check one box in each row.)

	Yes	*No*
a) The school has a policy on how to use computers in mathematics instruction (e.g., amount of computer use in mathematics classes, use of specific mathematics computer programs).	☐ 1	☐ 2
b) All 10th grade mathematics classes in the school use the same textbook.	☐ 1	☐ 2
c) Mathematics teachers in the school follow a standardized curriculum that specifies content at least on a monthly basis.	☐ 1	☐ 2

Q34 In your school, how likely is it that a student in 10th grade would be transferred to another school for the following reasons?

(Please check one box in each row.)

	Not likely	*Likely*	*Very likely*
a) Low academic achievement	☐ 1	☐ 2	☐ 3
b) High academic achievement	☐ 1	☐ 2	☐ 3
c) Behavioral problems	☐ 1	☐ 2	☐ 3
d) Special learning needs	☐ 1	☐ 2	☐ 3
e) Parents' or guardians' request	☐ 1	☐ 2	☐ 3
f) Other	☐ 1	☐ 2	☐ 3

SECTION G: FINANCIAL EDUCATION AT SCHOOL

The following five questions are about financial education/personal finance in your school. Financial education/personal finance involves the development of students' knowledge, confidence and skills relating to topics such as money and income; budgeting and long-term planning; saving and spending; credit and debt; investment and insurance; the potential risks and benefits of financial products; and the financial landscape (including consumer rights and responsibilities and understanding of the wider financial, economic and social system).

SC47

Q35 Which of the statements below best describes the situation for students in 10th grade regarding the availability of financial education in your school?

(Please check only one box.)

Financial education is not available. ☐ 1

Financial education has been available for less than two years. ☐ 2

Financial education has been available for two years or more. ☐ 3

SC45

Q36 Is financial education compulsory in your school?

(Please check only one box.)

Yes ☐ 1

No ☐ 2

Q37 Which of the statements below describe the teaching of financial education in your school?

For each statement, please indicate the number of hours of financial education of this type for students in 10th grade during the 2011-2012 academic year.

(Please check one box in each row.)

	Not at all	*1-4 hours a year*	*5-19 hours year*	*20-49 hours a year*	*50 or more hours a year*
a) It is taught as a separate subject.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b) It is taught as a cross-curricular subject.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c) It is taught as part of business or economics courses.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
d) It is taught as part of mathematics.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
e) It is taught as part of other social sciences and humanities subjects and/or literature/language (e.g., history, geography, home economics, civics).	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
f) It is available as an extracurricular activity.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
g) It is taught as part of homeroom.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

SC52

Q38 Who provides financial education in your school?

(Please check one box in each row.)

	Yes	*No*
a) Teachers	☐ 1	☐ 2
b) People from private sector, for-profit institutions (e.g., commercial bank, insurance company)	☐ 1	☐ 2
c) People from non-profit or not-for-profit organizations	☐ 1	☐ 2
d) People from public sector institutions	☐ 1	☐ 2

SC51

Q39 During the last twelve months, what percentage of teaching staff in your school has attended a program of professional development with a focus on financial education?

A program of professional development here is a formal program designed to enhance teaching skills or pedagogical practices. It may or may not lead to a recognized qualification. The program must last for at least one day in total and have a focus on the teaching of financial education.

a) Staff who teach financial education in your school _____ %

b) All other teaching staff in your school _____ %

Thank you very much for your cooperation in completing this questionnaire!

THIS PAGE INTENTIONALLY LEFT BLANK

About TIMSS
Countries Participating
TIMSS & PIRLS in 2011
Schedule
Frameworks
Encyclopedia
Mathematics Results
Science Results
Methods and Procedures
Contextual Questionnaires
International Database
Released Items

There are numerous contextual factors that affect students' learning, for example, type of school, school resources, instructional approaches, teacher characteristics, student attitudes, and home support for learning contribute heavily to student learning and achievement. For a fuller appreciation of what the TIMSS achievement results mean and how they may be used to improve students learning in mathematics and science, it is important to understand the contexts in which students learn. In every cycle, TIMSS collects a range of information about these contexts for learning by administering background questionnaires to students, teachers, school principals, and curriculum experts, which, together with assessing students' performance in mathematics and science, provide a rich source of data on student achievement. Countries that administered the TIMSS and PIRLS 2011 fourth grade assessments to the same sample of students also administered a Home Questionnaire *(Learning to Read Survey)*, which was completed by the students' parents or guardians. Each of the questionnaires is described below.

Student Questionnaires

Each student in the selected class completed a Student Questionnaire. The questionnaire asked about aspects of students' home and school lives, including basic demographic information, their home environment, school climate for learning, and self-perception and attitudes toward learning mathematics and science. In countries teaching science as separate subjects at eighth grade, students completed a version of the Student Questionnaire with questions specific to each subject (e.g., biology, chemistry, physics, and earth science).

Home Questionnaire (*Learning to Read Survey*)

For countries that administered the TIMSS and PIRLS 2011 fourth grade assessments to the same sample of students, the parents or guardians of each student completed the *Learning to Read Survey*, often referred to as the Home Questionnaire. The questionnaire asked about preparations for primary schooling, including attendance in preschool and literacy- and numeracy-centered activities in the home before the child began school, such as reading books, singing songs, or playing with number toys. Parents answered questions about home resources in addition to information about their highest level of education and employment situations.

Teacher Questionnaires

Teachers of the assessed classes responded to the Teacher Questionnaire. The questionnaire asked students' teachers about their education, professional development, and experience in teaching. It also asked about coverage of the mathematics and science curriculum and about the instructional activities

and materials used in the class of students selected for the TIMSS assessment.

School Questionnaires

The principal of each school sampled for TIMSS 2011 completed a School Questionnaire. Principals answered questions about student demographic characteristics, the availability of resources, types of programs, and environments for learning in their schools.

Curriculum Questionnaires

The TIMSS 2011 National Research Coordinator within each country was responsible for completing the Curriculum Questionnaire. Questions primarily centered on the organization and content of the curriculum in mathematics and science.

Download Questionnaires

Each of the TIMSS 2011 contextual questionnaires can be downloaded free of charge below.

The TIMSS 2011 Contextual Questionnaires are available in Portable Document Format (PDF). If needed, click here to download Adobe Reader® to read and print these materials.

Download all questionnaires for:

- Fourth Grade
- Eighth Grade

Download the questionnaires separately:

Fourth Grade

- Student Questionnaire
- Home Questionnaire (Learning to Read Survey)
- Teacher Questionnaire
- School Questionnaire
- Curriculum Questionnaire

Eighth Grade

- Student Questionnaire—General/Integrated Science Version
- Student Questionnaire—Separate Sciences Version
- Teacher Questionnaire—Mathematics
- Teacher Questionnaire—Science
- School Questionnaire
- Curriculum Questionnaire

Do Not Turn Page Until Instructed To Do So.

TIMSS 2011

Student Questionnaire

Grade 8

TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College

Directions

In this booklet, you will find questions about yourself. Some questions ask for facts while other questions ask for your opinion.

Each question is followed by a number of answers. Fill in the oval next to or under the answer of your choice as shown in Examples 1, 2, and 3.

Example 1

Do you go to school?

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

Yes -- ●

No -- ②

Example 2

How often do you do these things?

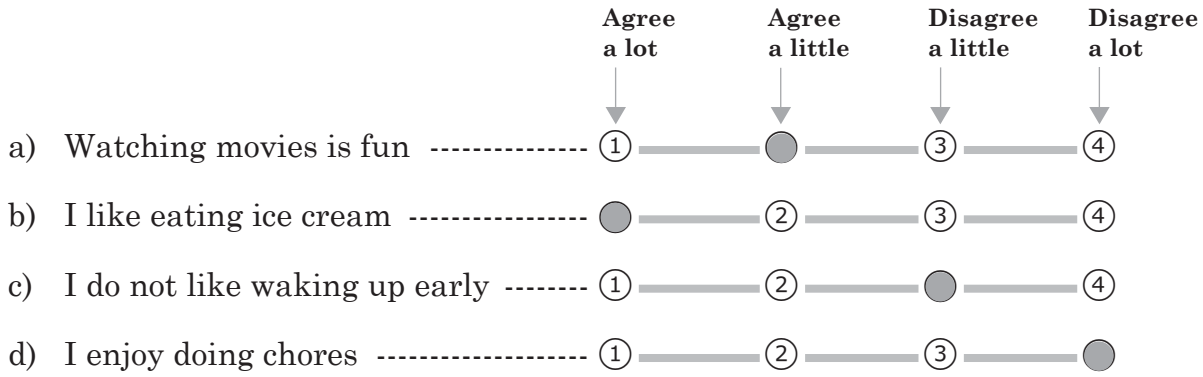
*Fill in **one** oval for each line.*

	Every day or almost every day	**Once or twice a week**	**Once or twice a month**	**Never or almost never**
a) I talk with my friends	●	②	③	④
b) I play sports	①	●	③	④
c) I ride a skateboard	①	②	●	④

Example 3

What do you think? Tell how much you agree with these statements.

*Fill in **one** oval for each line.*



- Read each question carefully, and pick the answer you think is best.
- Fill in the oval next to or under your answer.
- If you decide to change your answer, completely erase your first choice. Then, fill in the oval next to or under your new answer.
- Ask for help if you do not understand something or are not sure how to answer.

About you

1

A. Are you a girl or a boy?

Fill in one oval only.

Girl -- ①

Boy -- ②

B. Are you Hispanic or Latino?

Fill in one oval only.

Yes, I am Hispanic or Latino -- ①

No, I am not Hispanic or Latino -- ②

C. Which of the following best describes you?

Fill in one or more ovals.

White -- ①

Black or African American -- ①

Asian -- ①

American Indian or Alaska Native -- ①

Native Hawaiian or other Pacific Islander -- ①

2

When were you born?

Fill in the ovals next to the month and year you were born.

a) Month	b) Year
January -- ⓪	1993 -- ①
February -- ⓪	1994 -- ②
March -- ⓪	1995 -- ③
April -- ⓪	1996 -- ④
May -- ⓪	1997 -- ⑤
June -- ⓪	1998 -- ⑥
July -- ⓪	1999 -- ⑦
August -- ⓪	2000 -- ⑧
September -- ⓪	2001 -- ⑨
October -- ⓪	Other -- ⓪
November -- ⓪	
December -- ⓪	

3

A. How often do you speak English at home?

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

Always ---------- ① *If **Always**, please go to question 4* →

Almost always ---------- ②

Sometimes ---------- ③

Never ---------- ④

*If **Almost always, Sometimes, Never,** please go to question 3B* ↓

B. What language do you speak at home (other than English)?

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

Spanish ---------- ①

Other ---------- ② Please specify _____

4

A. Was your mother (or stepmother or female legal guardian) born in the United States? ("United States" includes the 50 states, its territories, the District of Columbia, and U.S. military bases abroad.)

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

Yes -- ①

No -- ②

B. Was your father (or stepfather or male legal guardian) born in the United States?

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

Yes -- ①

No -- ②

5

A. Were you born in the United States?

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

Yes -- ①

No -- ②

B. If you were not born in the United States, how old were you when you came to the United States?

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

Older than 10 years old -- ①

5 to 10 years old -- ②

Younger than 5 years old -- ③

6

How many days were you absent from school in the last month?

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

None -- ①

1 or 2 days -- ②

3 or 4 days -- ③

5 to 10 days -- ④

More than 10 days -- ⑤

7

Have you ever repeated a grade?

*Fill in **one** oval for each line.*

	Yes	No
a) In elementary school	①	②
b) In middle or junior high school	①	②

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

9

Do you have any of these things at your home?

*Fill in **one** oval for each line.*

	Yes	No
a) Computer	①	②
b) Study desk/table for your use	①	②
c) Books of your very own (do not count your school books)	①	②
d) Your own room	①	②
e) Internet connection	①	②
f) Encyclopedia (a book or CD)	①	②
g) PlayStation®, Game Cube®, XBox®, or other TV/Video game system	①	②
h) VCR, DVD or Blu-ray player	①	②

10

A. What is the highest level of education completed by your mother (or stepmother or female legal guardian)?

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

Some elementary or junior/high middle school ---------- ①

Completed junior high/middle school ---------- ②

Some high school ---------- ③

Completed high school ---------- ④

Completed a vocational/technical certificate after high school ---------- ⑤

Completed an Associate's degree (AA) in vocational/technical program ---------- ⑥

Completed a 2-year or 4-year college or university degree (i.e., Associate's or Bachelor's degree) ---------- ⑦

Completed an academic Master's degree, teaching certificate program, or first professional degree (e.g., law, medicine, dentistry) or higher ---------- ⑧

I don't know ---------- ⑨

10 (continued)

B. What is the highest level of education completed by your father (or stepfather or male legal guardian)?

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

Some elementary or junior/high middle school ---------- ①

Completed junior high/middle school ---------- ②

Some high school ---------- ③

Completed high school ---------- ④

Completed a vocational/technical certificate after high school ---------- ⑤

Completed an Associate's degree (AA) in vocational/technical program ---------- ⑥

Completed a 2-year or 4-year college or university degree (i.e., Associate's or Bachelor's degree) ---------- ⑦

Completed an academic Master's degree, teaching certificate program, or first professional degree (e.g., law, medicine, dentistry) or higher ---------- ⑧

I don't know ---------- ⑨

11

How far in your education do you expect to go?

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

- Finish junior high/middle school ①
- Finish high school ②
- Finish a vocational/technical certificate after high school ③
- Finish an Associate's degree (AA) in vocational/technical program ④
- Complete a 2-year or 4-year college or university degree (i.e., Associate's or Bachelor's degree) ⑤
- Beyond a bachelor's degree ⑥
- I don't know ⑦

12

How often do you use a computer in each of these places?

*Fill in **one** oval for each line.*

	Every day or almost every day	Once or twice a week	Once or twice a month	Never or almost never
a) At home	①	②	③	④
b) At school	①	②	③	④
c) Some other place	①	②	③	④

13

How often do the following things happen at home?

*Fill in **one** oval for each line.*

	Every day or almost every day	Once or twice a week	Once or twice a month	Never or almost never
a) My parents ask me what I am learning in school	①	②	③	④
b) I talk about my schoolwork with my parents	①	②	③	④
c) My parents make sure that I set aside time for my homework	①	②	③	④
d) My parents check if I do my homework	①	②	③	④

Your School

14

What do you think about your school? Tell how much you agree with these statements.

*Fill in **one** oval for each line.*

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I like being in school	①	②	③	④
b) I feel safe when I am at school	①	②	③	④
c) I feel like I belong at this school	①	②	③	④

15

During this year, how often have any of the following things happened to you <u>at school?</u>

*Fill in **one** oval for each line.*

	At least once a week	Once or twice a month	A few times a year	Never
a) I was made fun of or called names	①	②	③	④
b) I was left out of games or activities by other students	①	②	③	④
c) Someone spread lies about me	①	②	③	④
d) Something was stolen from me	①	②	③	④
e) I was hit or hurt by other student(s) (*e.g., shoving, hitting, kicking*)	①	②	③	④
f) I was made to do things I didn't want to do by other students	①	②	③	④

16

How much do you agree with these statements about learning mathematics?

*Fill in **one** oval for each line.*

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I enjoy learning mathematics	①	②	③	④
b) I wish I did not have to study mathematics	①	②	③	④
c) Mathematics is boring	①	②	③	④
d) I learn many interesting things in mathematics	①	②	③	④
e) I like mathematics	①	②	③	④
f) It is important to do well in mathematics	①	②	③	④

17

How much do you agree with these statements about your <u>mathematics lessons</u>?

*Fill in **one** oval for each line.*

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I know what my teacher expects me to do	①	②	③	④
b) I think of things not related to the lesson	①	②	③	④
c) My teacher is easy to understand	①	②	③	④
d) I am interested in what my teacher says	①	②	③	④
e) My teacher gives me interesting things to do	①	②	③	④

18

How much do you agree with these statements about mathematics?

*Fill in **one** oval for each line.*

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I usually do well in mathematics	①	②	③	④
b) Mathematics is more difficult for me than for many of my classmates	①	②	③	④
c) Mathematics is not one of my strengths	①	②	③	④
d) I learn things quickly in mathematics	①	②	③	④
e) Mathematics makes me confused and nervous	①	②	③	④
f) I am good at working out difficult mathematics problems	①	②	③	④
g) My teacher thinks I can do well in mathematics classes with difficult materials	①	②	③	④
h) My teacher tells me I am good at mathematics	①	②	③	④
i) Mathematics is harder for me than any other subject	①	②	③	④

18 (continued)

How much do you agree with these statements about mathematics?

*Fill in **one** oval for each line.*

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
j) I think learning mathematics will help me in my daily life	①	②	③	④
k) I need mathematics to learn other school subjects	①	②	③	④
l) I need to do well in mathematics to get into the college or university of my choice	①	②	③	④
m) I need to do well in mathematics to get the job I want	①	②	③	④
n) I would like a job that involves using mathematics	①	②	③	④

19

How much do you agree with these statements about learning science?

*Fill in **one** oval for each line.*

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I enjoy learning science	①	②	③	④
b) I wish I did not have to study science	①	②	③	④
c) I read about science in my spare time	①	②	③	④
d) Science is boring	①	②	③	④
e) I learn many interesting things in science	①	②	③	④
f) I like science	①	②	③	④
g) It is important to do well in science	①	②	③	④

20

How much do you agree with these statements about your <u>science lessons</u>?

*Fill in **one** oval for each line.*

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I know what my teacher expects me to do	①	②	③	④
b) I think of things not related to the lesson	①	②	③	④
c) My teacher is easy to understand	①	②	③	④
d) I am interested in what my teacher says	①	②	③	④
e) My teacher gives me interesting things to do	①	②	③	④

21

How much do you agree with these statements about science?

*Fill in **one** oval for each line.*

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I usually do well in science	①	②	③	④
b) Science is more difficult for me than for many of my classmates	①	②	③	④
c) Science is not one of my strengths	①	②	③	④
d) I learn things quickly in science	①	②	③	④
e) Science makes me confused and nervous	①	②	③	④
f) I am good at working out difficult science problems	①	②	③	④
g) My teacher thinks I can do well in science classes with difficult materials	①	②	③	④
h) My teacher tells me I am good at science	①	②	③	④
i) Science is harder for me than any other subject	①	②	③	④

21 (continued)

How much do you agree with these statements about science?

*Fill in **one** oval for each line.*

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
j) I think learning science will help me in my daily life	①	②	③	④
k) I need science to learn other school subjects	①	②	③	④
l) I need to do well in science to get into the college or university of my choice	①	②	③	④
m) I need to do well in science to get the job I want	①	②	③	④
n) I would like a job that involves using science	①	②	③	④

Homework

22

A. How often does your teacher give you homework in mathematics?

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

Every day -- ①

3 or 4 times a week -- ②

1 or 2 times a week -- ③

Less than once a week -- ④

Never -- ⑤

B. When your teacher gives you mathematics homework, about how many minutes do you usually spend on your homework?

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

My teacher never gives me homework in mathematics -- ①

1–15 minutes - ②

16–30 minutes -- ③

31–60 minutes -- ④

61–90 minutes -- ⑤

More than 90 minutes -- ⑥

23

A. How often does your teacher give you homework in science?

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

Every day -- ①

3 or 4 times a week -- ②

1 or 2 times a week -- ③

Less than once a week -- ④

Never -- ⑤

B. When your teacher gives you science homework, about how many minutes do you usually spend on your homework?

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

My teacher never gives me homework in science -- ①

1–15 minutes - ②

16–30 minutes -- ③

31–60 minutes -- ④

61–90 minutes -- ⑤

More than 90 minutes -- ⑥

THIS PAGE LEFT BLANK ON PURPOSE


Thank You!
Thank you for filling out the questionnaire!

NAEP
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

Mathematics
Student Questionnaire

2011
Grade 8

SECTION 3

In this section, please tell us about yourself and your family. The section has 12 questions. Mark your answers in your booklet. Fill in only **one** oval for each question except where instructed otherwise.

VB331330

1. Are you Hispanic or Latino? Fill in **one or more ovals.**

   Ⓐ No, I am not Hispanic or Latino.

   Ⓑ Yes, I am Mexican, Mexican American, or Chicano.

   Ⓒ Yes, I am Puerto Rican or Puerto Rican American.

   Ⓓ Yes, I am Cuban or Cuban American.

   Ⓔ Yes, I am from some other Hispanic or Latino background.

VB331331

2. Which of the following best describes you? Fill in **one or more ovals.**

   Ⓐ White

   Ⓑ Black or African American

   Ⓒ Asian

   Ⓓ American Indian or Alaska Native

   Ⓔ Native Hawaiian or other Pacific Islander

Section 3

For the rest of the questions in this section, fill in only **one** oval for each question.

VB331334

3. Does your family get any magazines regularly?

Ⓐ Yes

Ⓑ No

Ⓒ I don't know.

VB331335

4. 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)

VB331336

5. Is there a computer at home that you use?

Ⓐ Yes

Ⓑ No

VB331337

6. Is there an encyclopedia in your home? It could be a set of books, or it could be on the computer.

Ⓐ Yes

Ⓑ No

Ⓒ I don't know.

TB001101

7. About how many pages a day do you have to read in school and for homework?

Ⓐ 5 or fewer

Ⓑ 6–10

Ⓒ 11–15

Ⓓ 16–20

Ⓔ More than 20

VB331339

8. How often do you talk about things you have studied in school with someone in your family?

Ⓐ Never or hardly ever

Ⓑ Once every few weeks

Ⓒ About once a week

Ⓓ Two or three times a week

Ⓔ Every day

VB331447

9. How many days were you absent from school in the last month?

Ⓐ None

Ⓑ 1 or 2 days

Ⓒ 3 or 4 days

Ⓓ 5 to 10 days

Ⓔ More than 10 days

VB330870

10. How far in school did your mother go?

Ⓐ She did not finish high school.

Ⓑ She graduated from high school.

Ⓒ She had some education after high school.

Ⓓ She graduated from college.

Ⓔ I don't know.

VB330871

11. How far in school did your father go?

Ⓐ He did not finish high school.

Ⓑ He graduated from high school.

Ⓒ He had some education after high school.

Ⓓ He graduated from college.

Ⓔ I don't know.

VB331451

12. How often do people in your home talk to each other in a language other than English?

Ⓐ Never

Ⓑ Once in a while

Ⓒ About half of the time

Ⓓ All or most of the time

STOP

SECTION 4

Section 4

This section has 18 questions. Mark your answers in your booklet. Fill in only **one** oval for each question except where instructed otherwise.

VB543277

1. What math class are you taking this year?

   Ⓐ Geometry

   Ⓑ Algebra II

   Ⓒ Algebra I (one-year course)

   Ⓓ First year of a two-year Algebra I course

   Ⓔ Second year of a two-year Algebra I course

   Ⓕ Introduction to algebra or pre-algebra

   Ⓖ Basic or general eighth-grade math

   Ⓗ Integrated or sequential math

   Ⓘ Other math class

VB543278

2. What math class do you expect to take next year?

   Ⓐ Geometry

   Ⓑ Algebra II

   Ⓒ Algebra I (one-year course)

   Ⓓ First year of a two-year Algebra I course

   Ⓔ Second year of a two-year Algebra I course

   Ⓕ Introduction to algebra or pre-algebra

   Ⓖ Basic or general math

   Ⓗ Integrated or sequential math

   Ⓘ Business or consumer math

   Ⓙ Other math class

   Ⓚ I don't know.

VC290281

3. Do you study or do work for math at an after-school or tutoring program?

   Ⓐ Yes

   Ⓑ No

Section 4

VC497572

4. How often do you feel the following way in your math class? Fill in **one** oval on each line.

	Never or hardly ever	Sometimes	Often	Always or almost always	
a. I have a clear understanding of what my math teacher is asking me to do.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VC497573
b. The math work is too easy.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VC497574
c. The math work is challenging.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VC497575
d. The math work is engaging and interesting.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VC497576
e. I am learning.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VC497577

VC189706

5. Please indicate how much you DISAGREE or AGREE with the following statements. Fill in **one** oval on each line.

	Strongly disagree	Disagree	Agree	Strongly agree	
a. Because math is fun, I wouldn't want to give it up.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VC189707
b. I like math.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VC189710
c. Math is one of my favorite subjects.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VC189711

VB517159

6. How often do you use these different types of calculators in your math class?

	Never use	Sometimes, but not often	Usually use	
a. Basic four-function (addition, subtraction, multiplication, division)	Ⓐ	Ⓑ	Ⓒ	VB517160
b. Scientific (not graphing)	Ⓐ	Ⓑ	Ⓒ	VB517161
c. Graphing	Ⓐ	Ⓑ	Ⓒ	VB517282

VB517158

7. When you take a math test or quiz, how often do you use a calculator?

   Ⓐ Never

   Ⓑ Sometimes

   Ⓒ Always

VB543269

8. For each of the following activities, how often do you use a **calculator**? Fill in **one** oval on each line.

	Never or hardly ever	Once every few weeks	About once a week	Two or three times a week	Every day or almost every day	
a. To check your work on math homework assignments	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VB543270
b. To calculate the answers to math homework problems	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VB543271
c. To work in class on math lessons led by your teacher	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VB543272

VB543267

9. What kind of **calculator** do you usually use when you are **not in math class**?

   Ⓐ None

   Ⓑ Basic four-function (addition, subtraction, multiplication, division)

   Ⓒ Scientific (not graphing)

   Ⓓ Graphing

VB525162

10. How often do you use a computer for math at school?

    Ⓐ Never or hardly ever

    Ⓑ Once every few weeks

    Ⓒ About once a week

    Ⓓ Two or three times a week

    Ⓔ Every day or almost every day

Section 4

VC497143

11. Do you use a computer for math homework at home?

Ⓐ Yes

Ⓑ No

VB543148

12. On a typical day, how much time do you spend doing work for math class on a computer? Include work you do in class and for homework.

Ⓐ None

Ⓑ Half an hour or less

Ⓒ About 1 hour

Ⓓ About 2 hours

Ⓔ More than 2 hours

VB543155

13. When you are doing math for school or homework, how often do you use these **different types of computer programs**? Fill in **one** oval on each line.

	Never or hardly ever	Once every few weeks	About once a week	Two or three times a week	Every day or almost every day	
a. A spreadsheet program for math class assignments	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VB543157
b. A program to practice or drill on math facts (addition, subtraction, multiplication, division)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VB543158
c. A program that presents new math lessons with problems to solve	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VB543160
d. The Internet to learn things for math class	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VB543159
e. A calculator program on the computer to solve or check problems for math class	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VB543161
f. A graphing program on the computer to make charts or graphs for math class	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VB543262
g. A statistical program to calculate patterns such as correlations or cross tabulations	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VB517157
h. A word processing program to write papers for math class	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VB543156
i. A program to work with geometric shapes for math class	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VC466133

Section 4

VC189591

14. Do you use the Internet at home?

Ⓐ Yes

Ⓑ No

VC189613

15. How often do you use e-mail, instant messages, blogs, or text messages to do any of the following? Fill in **one** oval on each line.

	Never or hardly ever	Once every few weeks	About once a week	Two or three times a week	Every day or almost every day
a. Talk online with friends about math work	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
b. Get help with math from someone other than your teacher, family, classmates, or friends	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ

VC299268

VC189627

VB595182

16. How hard was this test compared to most other tests you have taken this year in school?

Ⓐ Easier than other tests

Ⓑ About as hard as other tests

Ⓒ Harder than other tests

Ⓓ Much harder than other tests

VC034559

17. How hard did you try on this test compared to how hard you tried on most other tests you have taken this year in school?

Ⓐ Not as hard as on other tests

Ⓑ About as hard as on other tests

Ⓒ Harder than on other tests

Ⓓ Much harder than on other tests

VB595184

18. How important was it to you to do well on this test?

Ⓐ Not very important

Ⓑ Somewhat important

Ⓒ Important

Ⓓ Very important

Technology and Engineering Literacy Background Questionnaire

Because of the growing importance of technology and engineering in the educational landscape, and to support America's ability to contribute to and compete in a global economy, the National Assessment Governing Board initiated development of the first national assessment in Technology and Engineering Literacy (TEL). In 2014, the NAEP TEL assessment will be administered to a national sample of eighth-grade students in public and private schools. The TEL background questionnaire is currently being piloted as part of the large-scale psychometric pilot in 2013, leading to the 2014 assessment. TEL has both a student questionnaire and a school administrator questionnaire. The TEL assessment will not include a teacher questionnaire since technology and engineering literacy is not taught as a course in most schools. NCES would not be able to link student responses with teacher responses as the majority of students taking the assessment would not have a "TEL teacher."

TEL background questions are designed to cover the assessment's three target areas: 1) Technology and Society, 2) Design and Systems, and 3) Information and Communication Technology. There are questions that cut across these three target areas to assess students' experience and interest in technology and engineering. Within the target areas, questions are designed to measure what students have done or are currently doing as it applies to technology and engineering, where they have acquired their knowledge, and students' self-efficacy. In contrast to other NAEP subject-area questionnaires, a higher percentage of TEL questions ask students about their experiences outside of school. Additionally, the school questionnaire provides a platform for assessing the school's curricular focus and available resources relating to TEL.

The TEL questionnaires underwent several stages of iterative development, including Governing Board reviews[1]; cognitive interviews; exploratory research; a large-scale tryout of the student questionnaires; and a pilot of both the student and school questionnaires (which is currently being administered). While Governing Board reviews and cognitive interviews are standard practice for NAEP background questions, exploratory research and a large-scale tryout represented additional work to help further refine background questions. The overall purpose of

[1] The Governing Board's Assessment Development Committee (ADC) first reviewed the TEL questions in December 2011 and provided comments and recommendations. ADC reviewed the TEL questions a second time in April 2012, prior to these questions being prepared for pilot testing. ADC will review these questions again in spring 2013, prior to the 2014 assessment.

exploratory research was to ascertain what could be asked, how these questions should be constructed, and how findings might be reported. The item tryout served as a pre-pilot activity to determine what additional changes were needed before the pilot. Below is a more in-depth description of the purpose of the exploratory research and the item tryout:

- **Exploratory research**, via focus groups conducted with student and parents in fall 2011 and with teachers in winter 2012, was used to:
    - Investigate what kinds of technology and engineering activities students are doing, particularly outside of school;
    - Identify (and if necessary reduce) any potential gender imbalance in the current sub-item lists of technology and engineering activities;
    - Understand issues around access to (and learning of) technology and engineering; and
    - Determine how to best report TEL background questionnaire findings to varied audiences (e.g., parents, educators, policymakers)
- **Item Tryout:** conducted in March 2012 with nearly 3,000 respondents. The goals of the tryout for the TEL student questionnaire were:
        - Pretest new and revised questions that resulted from cognitive interviews and previous reviews;
        - Validate the appropriate range of response choices for certain questions; and
        - Confirm the time taken to complete questions.

The information obtained from the tryout was used to make decisions about which questions to retain, drop, or modify for the pilot. The student questionnaire items in the pilot are being administered via a spiral design, meaning students will not be given every question. This design allows NCES to assess which questions perform best in the pilot. Following the pilot, we will propose to the ADC Committee a set of questions to be included in the 2014 TEL assessment.

At the February/March meeting, the TEL background presentation will highlight questionnaire development activities that have led up to the pilot. This presentation will focus specifically on the exploratory research, the item tryout, and pilot (including the 2013 spiral design). Additionally, there will be a summary of lessons learned from these various activities.

VE682225

1. **In school**, how often have you learned about or discussed the following? Select **one** circle in each row.

	Never	Rarely	Sometimes	Often	
a. The use and purpose of tools, machines, or devices	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682226
b. The care or maintenance of tools, machines, or devices	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682227
c. Designing or creating something to solve a problem	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682228
d. Designing something when there is limited time, money, or materials	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682229
e. Figuring out how to fix something	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682230
f. Finding the right people to work with or get help from to fix something	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682231

VE682232

2. **Outside of school**, how often have you learned about or discussed the following? Select **one** circle in each row.

	Never	Rarely	Sometimes	Often	
a. The use and purpose of tools, machines, or devices	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682233
b. The care or maintenance of tools, machines, or devices	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682234
c. Designing or creating something to solve a problem	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682238
d. Designing something when there is limited time, money, or materials	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682236
e. Figuring out how to fix something	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682237
f. Finding the right people to work with or get help from to fix something	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682235

VE639166

3. How interested are you in learning about the following areas of technologies? Select **one** circle in each row.

	Not at all interested	Not too interested	Somewhat interested	Very interested	
a. Information and communication (for example, computers, Internet, social networking sites)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639168
b. Transportation (for example, cars, planes, trains, traffic analysis)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639169
c. Construction (for example, architecture, building a bridge)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639171
d. Power and energy (for example, dams, power plants, batteries)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639173
e. Environmental and green technologies (for example, recycling, renewable energy sources such as sunlight and wind)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639174
f. Agriculture (for example, farming, food chemistry)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639175
g. Medical technologies (for example, vaccines, drugs, surgical tools, heart monitors, x-ray machines)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639176
h. Home and domestic (for example, air conditioning, cleaning, cooking, heating, plumbing, sewing)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VF009755
i. Manufacturing (for example, what goes on in factories, developing or improving products)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639170

VE639123

4. **For school work**, how often do you use a computer or other digital technology for the following activities? Select **one** circle in each row.

	Never or almost never	A few times a year	Once or twice a month	Once or twice a week	Every day or almost every day	
a. Send or receive messages (for example, chat, e-mail, instant messages, text messages)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639125
b. View or download digital media (for example, art, books, games, mobile apps, music, pictures, software, videos)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639127
c. Create, edit, or organize digital media	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639130
d. Send, share, present, or upload digital media	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639131
e. Create a presentation	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639137
f. Create a spreadsheet (a table or grid that displays data into columns and rows and may be used to create charts and graphs)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639136

VF025108

5. In this question, please think about activities you do that are not related to your school work. How often do you use a computer or other digital technology for the following activities **not for school work**? Select **one** circle in each row.

	Never or almost never	A few times a year	Once or twice a month	Once or twice a week	Every day or almost every day	
a. Send or receive messages (for example, chat, e-mail, instant messages, text messages)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF025109
b. View or download digital media (for example, art, books, games, mobile apps, music, pictures, software, videos)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF025110
c. Create, edit, or organize digital media	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF025112
d. Send, share, present, or upload digital media	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF025113
e. Create a presentation	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF025117
f. Create a spreadsheet (a table or grid that displays data into columns and rows and may be used to create charts and graphs)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF025116

VE682276

6. Do you think that you would be able to do each of the following? Select **one** circle in each row.

	I definitely can't	I probably can't	Maybe	I probably can	I definitely can	
a. Build a model using a kit	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE682278
b. Build a model without using a kit	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE682280
c. Use tools or materials to fix something	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE682281
d. Take something apart in order to fix it or see how it works	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE682284
e. Design a computer program	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE682286

VE639025

7. Technology refers to all the things people make and do to their natural environment in order to get the things they want and need. How much do you disagree or agree with the following statements about technology? Select **one** circle in each row.

	Strongly disagree	Disagree	Neither disagree nor agree	Agree	Strongly agree	
a. Technology is important to society.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639028
b. Technology is important to my daily life.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639043
c. Learning about technology will help me in the future.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639048
d. Learning about technology will help me do (or get) the job I want.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639046
e. I enjoy learning about technology.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639053
f. I enjoy using technology.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF009048

VF009050

8. Engineering refers to using skills or knowledge to solve problems that meet people's wants and needs. How much do you disagree or agree with the following statements about engineering? Select **one** circle in each row.

	Strongly disagree	Disagree	Neither disagree nor agree	Agree	Strongly agree	
a. Engineering is important to society.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF009051
b. Engineering is important to my daily life.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF009052
c. Learning about engineering will help me in the future.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF009053
d. Learning about engineering will help me do (or get) the job I want.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF009054
e. I enjoy learning about engineering.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF009055
f. I enjoy solving problems.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF009056
g. I enjoy fixing things.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF009061
h. I enjoy creating, building, or designing things.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF009064
i. I enjoy figuring out how things work.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF009065
j. I do things that I would describe as engineering.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF009066

VE682274

9. Who taught you most of what you know about building things, fixing things, or how things work?

Ⓐ I taught myself.

Ⓑ Family members

Ⓒ Friends

Ⓓ Teachers

Ⓔ Someone else

VE682217

10. Do you think that you would be able to do each of the following? Select **one** circle in each row.

	I definitely can't	I probably can't	Maybe	I probably can	I definitely can	
a. Publish or maintain a personal website or blog	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE682218
b. Create presentations with sound, pictures, or video	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE682219
c. Organize information into a chart, graph, or spreadsheet	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE682221
d. Compare products using the Internet	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE682222

VE682215

11. Who taught you most of what you know about using computers or other digital technology for collecting or sharing information?

Ⓐ I taught myself.

Ⓑ Family members

Ⓒ Friends

Ⓓ Teachers

Ⓔ Someone else

VE639847

12. Have you ever studied technology or engineering topics in any of the following classes or subjects in school? Select **one or more** squares.

Ⓐ Mathematics

Ⓑ Science

Ⓒ Social studies or history

Ⓓ I have not studied technology or engineering in any of the classes or subjects listed above.

VE638956

13. In school, how often have you ever done the following activities? Select **one** circle in each row.

	Never	Once or twice	Three to five times	More than five times	
a. Used tools or materials to fix or build something	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE638957
b. Used different tools, materials, or machines to see which are best for a given purpose	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE638959
c. Built or tested a model to see if it solves a problem	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE638963
d. Figured out why something is not working in order to fix it	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682247
e. Taken something apart in order to fix it or see how it works	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE638965
f. Designed a computer program	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682248

VE638983

14. **Outside of school**, how often have you ever done the following activities? Select **one** circle in each row.

	Never	Once or twice	Three to five times	More than five times	
a. Used tools or materials to fix or build something	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE638986
b. Used tools or materials to plan or design something (for example, cake recipe, party)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VF009777
c. Used different tools, materials, or machines to see which are best for a given purpose	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE638998
d. Built or tested a model to see if it solves a problem	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639038
e. Figured out why something is not working in order to fix it	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682267
f. Taken something apart in order to fix it or see how it works	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639042
g. Designed a computer program	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682268

VF238958

15. **For school work**, how often do you use a computer or other digital technology for the following activities? Select **one** circle in each row.

	Never or almost never	A few times a year	Once or twice a month	Once or twice a week	Every day or almost every day	
a. Participate in online discussion forums, social networking sites, or virtual communities	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF238965
b. Work with others to solve a problem	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF238968
c. Get information from experts (people with strong skills or knowledge in a subject)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF238969
d. Maintain a website or blog	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF238973
e. Search for information (for example, browse the Internet or check out websites)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF238974
f. Play games or run simulations	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VF238975

VE639842

16. Have you ever taken or are you currently taking any of the following classes or subjects in school? Select **one or more** squares.

Ⓐ Industrial technology (for example, auto mechanics, carpentry)

Ⓑ Engineering (for example, robotics, bridge building, rocketry)

Ⓒ Any class that involves learning to use, program, or build computers

Ⓓ Any other technology-related class (for example, electronics, sewing, farming)

Ⓔ I have not taken any of the classes listed above.

VE639172

17. In school or outside of school, how often do you do the following? Select **one** circle in each row.

	Never or almost never	A few times a year	Once or twice a month	Several times a month	At least once a week	
a. Participate in clubs, camps, or competitions about technology or engineering (for example, digital art and editing, design, programming, robotics, science)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639177
b. Go to museums or events to learn about technology or engineering	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639178
c. Edit digital photographs or other graphic images	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639179
d. Create, build, or design things (for example, robots, clothes, science projects, recipes)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639180
e. Work in a shop or garage with industrial technologies (for example, auto mechanics, machining, metalworking, construction, carpentry)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639181
f. Work with drafting or design tools (for example, computer aided design [CAD], systems analysis)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639182
g. Take online classes to learn more about technology or engineering	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE639183
h. Watch video or listen to audio to learn more about technology or engineering (video or audio includes online videos, movies, television shows, podcasts, radio programs)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE677642

VE681624

18. In school, how often do you learn about or discuss the following? Select **one** circle in each row.

	Never	Rarely	Sometimes	Often	
a. How to judge reliability of sources (for example, how a website might be biased or inaccurate)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE681629
b. How to credit others for their ideas (for example, citing sources, using endnotes and footnotes in reports)	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE681632

VE638999

19. In school, how often have you learned about or discussed the following? Select **one** circle in each row.

	Never	Rarely	Sometimes	Often	
a. Inventions that change the way people live	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639002
b. Choices people make that affect the environment	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639004
c. Conditions that influence the use or availability of machines or devices	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639005
d. The ways people work together to solve problems in their community or the world	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682300

VE639008

20. **Outside of school**, how often have you learned about or discussed the following? Select **one** circle in each row.

	Never	Rarely	Sometimes	Often	
a. Inventions that change the way people live	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639012
b. Choices people make that affect the environment	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639013
c. Conditions that influence the use or availability of machines or devices	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE639014
d. The ways people work together to solve problems in their community or the world	Ⓐ	Ⓑ	Ⓒ	Ⓓ	VE682314

VE682315

21. Who taught you most of what you know about how technology, people, and the environment are related to each other?

Ⓐ I taught myself.

Ⓑ Family members

Ⓒ Friends

Ⓓ Teachers

Ⓔ Someone else

VE682317

22. Do you think that you would be able to do each of the following? Select **one** circle in each row.

	I definitely can't	I probably can't	Maybe	I probably can	I definitely can	
a. Describe how inventions change society	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE682321
b. Compare how different activities affect the environment	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE682323
c. Explain why people have different tools, machines, or devices in different parts of the world	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ	VE682324

VF009358

23. Before today, had you ever taken an interactive computer test similar to the one you just took? Select **one** circle in each row.

	Yes	No	
a. I had taken an interactive computer test in school.	Ⓐ	Ⓑ	VF009360
b. I had taken an interactive computer test outside of school.	Ⓐ	Ⓑ	VF009361

VE401773

24. How hard was this test compared to most other tests you have taken this year in school?

Ⓐ Easier than other tests

Ⓑ About as hard as other tests

Ⓒ Harder than other tests

Ⓓ Much harder than other tests

VE401776

25. How hard did you try on this test compared to how hard you tried on most other tests you have taken this year in school?

Ⓐ Not as hard as on other tests

Ⓑ About as hard as on other tests

Ⓒ Harder than on other tests

Ⓓ Much harder than on other tests

VE401779

26. How important was it to you to do well on this test?

Ⓐ Not very important

Ⓑ Somewhat important

Ⓒ Important

Ⓓ Very important

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, is working on two follow-up studies to his exploratory analysis of NAEP data on time and learning that was presented at the first Ad Hoc Committee meeting in November.

One study mines the NAEP Data Explorer to provide state and district-level data on student absenteeism, homework assignments, and instructional time. The analysis shows how this detail, available through NAEP, may be used to report on how well various jurisdictions are trying to meet the needs of target groups of at-risk students.

The second study examines background information obtained in the special NAEP 2011 assessment of 8th grade science. It not only examines instructional time, but other factors, such as student interest, teacher preparation, and instructional methods, that are related to science achievement.

At this meeting a preview of the studies will be presented along with the consultant's recommendations for additional topics and questions on which data may be gathered and special reports prepared in future NAEP assessments.

Draft reports and data anlyses will be provided before the meeting.

EDUCATION WEEK
Published Online: December 11, 2012
Published in Print: December 12, 2012, as **Absenteeism Linked to Low Achievement in NAEP Time Study**

FOCUS ON: TIME

School Absences Translate to Lower Test Scores, Study Says

Minutes spent on homework on the rise

By Sarah D. Sparks

Washington

Back to Story

Missing even a few days of school seems to make a difference in whether 8th graders perform at the top of their game, according to a new analysis of results from the National Assessment of Educational Progress.

The report, the first of a planned **series of analyses** of NAEP's background-survey data, looks at how 4th and 8th graders use existing school time, including their attendance, instructional time, and homework. It was previewed here at a Nov. 29 meeting of the National Assessment Governing Board, which sets policy for NAEP. The study found that instructional time in reading, math, music, and the visual arts is on the rise nationwide, and that teachers are expecting more homework from their middle school students. As schools ramp up their academic focus, however, the analysis shows the cost of missing school may be greater.

Fifty-six percent of 8th graders who performed at the advanced level in NAEP reading in 2011 had perfect attendance in the month before the test, compared with only 39 percent of students who performed below the basic level.

In comparison, nearly one in five 8th graders at the basic level and more than one in four below basic in reading had missed three or more days in the past month, according to Alan L. Ginsburg, a research consultant for the governing board and a co-author of the report with Naomi Chudowsky of Caldera Research in Bend, Ore.

"Three days, if you multiply that out by nine months, is five weeks a year," Mr. Ginsburg said. "You've got more than a quarter of the below-basic kids who are going to miss five weeks of school a year or more," he said, noting that only 8 percent of students at the advanced level had missed that much school. "That, to me, would be something that if you are a chief state school officer or a superintendent, you might worry about."

The analysis contributes to mounting evidence that absenteeism puts students at greater risk of poor academic achievement and eventually dropping out of high school.

"For those of us in schools, this reflects what we've been saying all along: In order to advance, in order to learn, you have to be there," said Doris Hicks, a governing board member and the principal and chief executive officer of the Dr. Martin Luther King Jr. Charter School for Science and Technology, in New Orleans.

Low Performers

Academic expectations seem to be increasing for middle school students both in school and at home, the researchers found. But the bulk of the additional instructional time happened before the 2001 passage of the federal No Child Left Behind Act, with its new demands for academic progress, and the students who most needed extra time weren't always the ones to get it.

Teachers reported that from 1996 to 2000, 18 percent of 8th graders moved from having less than four hours of mathematics instruction each week to four or more hours a week, and from 2005 to 2011, another 6 percent of students started receiving five or more hours of math each week.

While 8th graders performing at or below basic in math on the 2011 NAEP were more likely than advanced students to receive seven hours of math instruction a week or more, the researchers found that more than half of 8th graders performing below basic in math received less than an hour of math each day on average.

"To me, this is [about] opportunity to learn," Mr. Ginsburg said. "Are the kids getting the amount of instruction they need to succeed?

"At grade 8, prealgebra, where we have most kids getting less than an hour a day on average," he said, "does that make sense? ... You have a group of kids who are below basic, who are in need of help, and they are getting less than an hour a day of instruction."

Some educators have voiced concern that extending math and reading instructional time could crowd out other subjects, but the researchers actually found a slight increase in arts instruction in middle school. Fifty-seven percent of 8th graders had music instruction three or more times a week in 2008, up from 49 percent in 1994. During the same time, 47 percent of 8th graders had visual arts at least three times a week, 5 percentage points higher than in 1994.

Moreover, the analysis found that teachers are expecting students to do more work outside of class to bolster their class time. From 1996 to 2011, the percentage of 8th graders assigned an hour or more of math homework each night rose more than fourfold, from 4 percent to 17 percent.

U.S. Secretary of Education Arne Duncan last week called for more expanded school days and years. But the NAEP background questionnaire does not include questions on school length, so researchers were not able to include such data in the report.

The assessment governing board has also released an analysis of charter school attendance and achievement, and it is planning as many as a dozen reports intended to "develop a portrait of American education."

"You're raising questions with this data for the field that I think will be very useful," Mr. Ginsburg said.

MEMORANDUM

UNITED STATES DEPARTMENT OF EDUCATION
INSTITUTE OF EDUCATION SCIENCES
National Center for Education Statistics

TO: Larry Feinberg
National Assessment Governing Board

FROM: Jamie Deaton

DATE: January 31, 2013

RE: NCES Response to the Background Information Framework for the National Assessment of Educational Progress

The purpose of this memorandum is to present NCES' response to the Background Information Framework. NCES appreciates the opportunity to review and provide comments to this framework which was adopted August 1, 2003. Additionally, NCES acknowledges the hard work of the Ad Hoc Committee on Background Questionnaires in looking at procedures and methodologies to strengthen our survey questions.

We understand that the Ad Hoc Committee will be charged with reviewing this framework and possibly updating it too. To aid in this endeavor, we have focused our comments on areas of the framework that we have concerns about.

Reporting Trend Data

The framework acknowledges NAEP's primary mission as "to provide a fair and accurate measure of student achievement and on achievement trends over time" (p. 1). However, the framework does not substantially discuss the importance of preserving trend.

NCES also notes the current framework language advises, "Those questions showing little change over time and/or a stable relationship to achievement should be deleted or asked less frequently and to limited samples, unless required to ensure precision of NAEP results" (p. 55). We would advise caution regarding this statement. There may be questions that show little change over time and a stable relationship with achievement, but provide valuable information to policymakers. These questions can serve value in providing long trend lines, revealing stability in one area even amidst significant changes to other surrounding variables. NCES acknowledges learning that something has not changed over time is sometimes just as important as learning something has changed. We would recommend that the language found in the Board's implementation guidelines is generally more appropriate: "A thorough review will be conducted to eliminate duplicative or low-priority questions. Unproductive topics and questions will be dropped[1]." The term "unproductive topics" is somewhat vague. Therefore, we would recommend revising the framework to provide examples for clarification. Suggestions to clarify the term "unproductive topics" may include those that are: 1) of interest to an extremely narrow

[1] Guideline #3.

population, 2) produce little response variance, 3) not usefully reported on, 4) outdated, or 5) duplicative.

NCES recognizes the tension between having the "best" questions possible and having questions which offer researchers and policy makers the ability to track changes over time. We fully support the Board's recent implementation guideline[2] that states, "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."

Comprehensive reviews over the last several years by NCES, contractor staff, subject-area standing committees, and the Governing Board combined with cognitive interview testing and analysis from pilot and operational assessments have been used to inform changes and improve background questions. These reviews and activities resulted in more reliable questions that are understood as intended by our respondents. However, this work has also created an overall pool of questions in which a significant amount do not have trend data or offer trend data going back to only the previous assessment administered. NCES would advise that if the Background Information Framework were to be revised, a section be included that elaborates on the need to maintain consistent wording.

Sampling Design of Questionnaires

Comparable to the Policy Statement on NAEP Background Questions and the Use of Contextual Data in NAEP Reporting, the Background Information Framework expresses a desire to expand the pool of questions via subsamples. The framework states, "For each subject 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" (p. 5). We understand the desire to collect more information, but also recognize that administering questions to smaller subsamples will decrease the individual sample size for each question. The standard errors associated with the results will increase. The exact impact of this increase in standard errors would need to be investigated to evaluate if it would have significant reporting implications for subgroups within jurisdictions. NCES wants to make sure that using smaller subsamples does not compromise our ability to report background data, especially for our minority populations.

Administering some questions periodically, rather than for every administration, is certainly feasible. NCES would recommend avoiding this approach for assessments that are administered every four years or more. Using periodic administrations of questions is most easily done for mathematics and reading at grades 4 and 8, since these assessments are conducted every two years. Over an 8-year span, these assessments would be administered four times. The benefit of administering some questions periodically would be the collection of more information. Clusters of questions could be developed on topics of continuing interest or questions from international assessments, might be included.[3] One caveat is using this approach heightens the need to maintain trend (when possible) as over four administrations of an assessment, some questions have less than four data points.

[2] Guideline #5.
[3] This approach would address guidelines #1, 4, and 8 from the Policy Statement on NAEP Background Questions and the Use of Contextual Data in NAEP Reporting, adopted by the Board on August 4, 2012.

A revised framework could include a section that discusses the costs and benefits of various student questionnaire rotation designs and the value of further exploring these innovative designs for NAEP. This revised version could also include a discussion about how the comparability of background questions used in NAEP and questions used in international assessments could be increased.

Increasing Topic Focus

NCES would also suggest that 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. A more topic-focused questionnaire could also align future NAEP assessments closer with the third key principle of the current framework which states, "the collection on background data shall be designed to obtain information that is objective, valid, reliable, and of consistently high quality" (p. vi).

Related to the idea of more topic-focused questionnaires are special studies. There is no direct reference to the implementation or integration of special studies into the NAEP assessment schedule. It seems, however, that the framework recognizes the need for special assessment opportunities. The Policy Statement on NAEP Background Questions and the Use of Contextual Data in NAEP Reporting stated "Modules will be prepared for special one-item studies to provide descriptive information on issues of current policy interest[4]." NCES would recommend the framework be updated with a section specifically concerning special studies.

Timing of Questionnaires

NCES recommends updating the language in the Background Information Framework concerning the length of the questionnaires. The framework states, "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: ten minutes for each student . . ." (p. 6). Time changes are being implemented as new computer-based assessments are being introduced. Our current computer-based assessments, which include writing and the Technology and Engineering Literacy (TEL) pilot, have a maximum background questionnaire time of 15 minutes[5] – five minutes for core questions and 10 minutes for subject-specific questions. NCES recommends that the Board should continue to recognize that time burden is a concern, given that NAEP questionnaires are a voluntary supplement.

[4] Guideline #2.

[5] This aligns with guideline #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.

Updating Outdated Language

In addition to revising language concerning the timing of the questionnaires, NCES recognizes that a framework revision should look comprehensively at the entire document, updating text as appropriate. For example, there are references to a NAEP question on television watching (p. 39 and 44). This question is longer asked.

Using Outside Data Sources

The framework advocates that data collected outside NAEP should be considered to improve reporting (e.g., U.S. Census, Quality Education, Data, Inc., Common Core of Data, and School of Staffing Survey). NCES agrees with using outside sources when possible. We would like to add that using state data may be another possibility in the future. The use of external data may increase the validity of our data and reduce respondent burden. However, a key consideration when using other sources is timeliness. We must confirm that all of the data is coming from the same school year.

Emerging Contexts for NAEP

The assessment field is rapidly evolving. Much has changed since the Background Information Framework was published nearly a decade ago. NCES recommends that any revisions to the framework include a discussion of emerging contexts and the ramifications for NAEP. International assessments are playing an increasingly prominent role in the educational landscape. As interest in international comparisons on educational factors has increased, educators, policymakers, and economists have increasingly acknowledged the importance of non-cognitive student factors for success in education and the workplace. In addition to these international comparisons, the Common Core is another emerging area. The framework may also want to address that NAEP, as well as many other assessments, are shifting from a paper and pencil mode to a computer-based platform. These are a few evolving issues that a revised framework may want to examine.

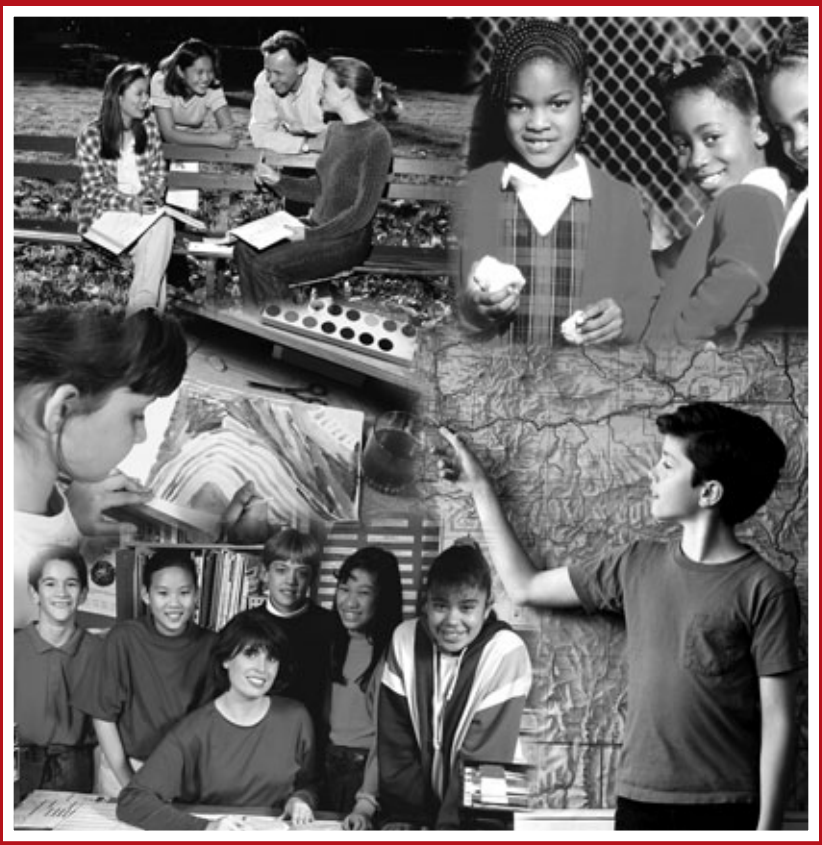
Changing the name "Background Questions"

The Ad Hoc Committee may wish to discuss revising the term "background questions." The word "background" may cause some to infer that NAEP survey questions ask respondents very personal or invasive questions, ones that might be found on a background check. Recent NAEP field publications have avoided using this term. For example, we advise using the term "student questionnaire" in place of "background questionnaire." Comparably, background questions for teachers and school administrators could be described as "teacher questionnaire" and "school administrator questionnaire."

Reviewing the Background Information Framework

We believe the Background Information Framework is a valuable resource for the Governing Board, NCES, and the NAEP contractors as we continue questionnaire development activities

with each assessment cycle. This framework's use is unique and frequent since questionnaires are administered annually, albeit for different subject areas. Consequently, NCES would recommend viewing this guide as a living document, one that might be updated "as needed" on a smaller level. We fully support a thorough review from the Ad Hoc Committee in the near future and then subsequent reviews, approximately every ten years as was originally planned.



Background Information Framework for the National Assessment of Educational Progress

National Assessment Governing Board
U.S. Department of Education

What Is NAEP?

The National Assessment of Educational Progress (NAEP) is the only nationally representative and continuing assessment of what American students know and can do in various academic subjects. It is a congressionally mandated project of the U.S. Department of Education's National Center for Education Statistics. NAEP surveys have been conducted on a national sample basis since 1969 in reading, mathematics, writing, science, and other elementary and secondary school subjects. State-level assessments have been conducted since 1990.

The National Assessment Governing Board

The National Assessment Governing Board (NAGB) was created by Congress in 1988 to formulate policy for NAEP. Among the Board's responsibilities are determining the content of NAEP and designing the assessment methodology. The Board has final authority on the appropriateness of all test items.

Darvin M. Winick, Chair
President
Winick & Associates
Dickinson, Texas

Amanda P. Avallone
Assistant Principal and
8th Grade Teacher
Summit Middle School
Boulder, Colorado

Daniel A. Domenech
Superintendent of Schools
Fairfax County Public Schools
Fairfax, Virginia

Edward Donley
Former Chairman
Air Products & Chemicals, Inc.
Allentown, Pennsylvania

Honorable Dwight Evans
State Legislator
Philadelphia, Pennsylvania

Thomas H. Fisher
Former Director
Student Assessment Services
Florida Department of Education
Tallahassee, Florida

Sheila M. Ford
Principal
Horace Mann Elementary School
Washington, D.C.

Edward H. Haertel
Professor
School of Education
Stanford University
Stanford, California

Catherine (Katy) Harvey
Principal
Bethesda-Chevy Chase
High School
Bethesda, Maryland

Juanita Haugen
Local School Board Member
Pleasanton, California

Honorable Dirk Kempthorne
Governor of Idaho
Boise, Idaho

Kim Kozbial-Hess
4th Grade Teacher
Fall-Meyer Elementary School
Toledo, Ohio

Honorable Ronnie Musgrove
Governor of Mississippi
Jackson, Mississippi

Mark D. Musick
President
Southern Regional Education
Board
Atlanta, Georgia

Honorable Jo Ann Pottorff
State Legislator
Wichita, Kansas

Diane Ravitch
Research Professor
New York University
New York, New York

Sister Lourdes Sheehan, R.S.M.
Associate General Secretary
United States Catholic Conference
Washington, D.C.

Honorable Raymond J. Simon
Director
Arkansas Department of Education
Little Rock, Arkansas

John H. Stevens
Executive Director
Texas Business and Education
Coalition
Austin, Texas

Deborah Voltz
Associate Professor
Department of Special Education
University of Louisville
Louisville, Kentucky

Honorable Michael E. Ward
State Superintendent of Public
Instruction
Public Schools of North Carolina
Raleigh, North Carolina

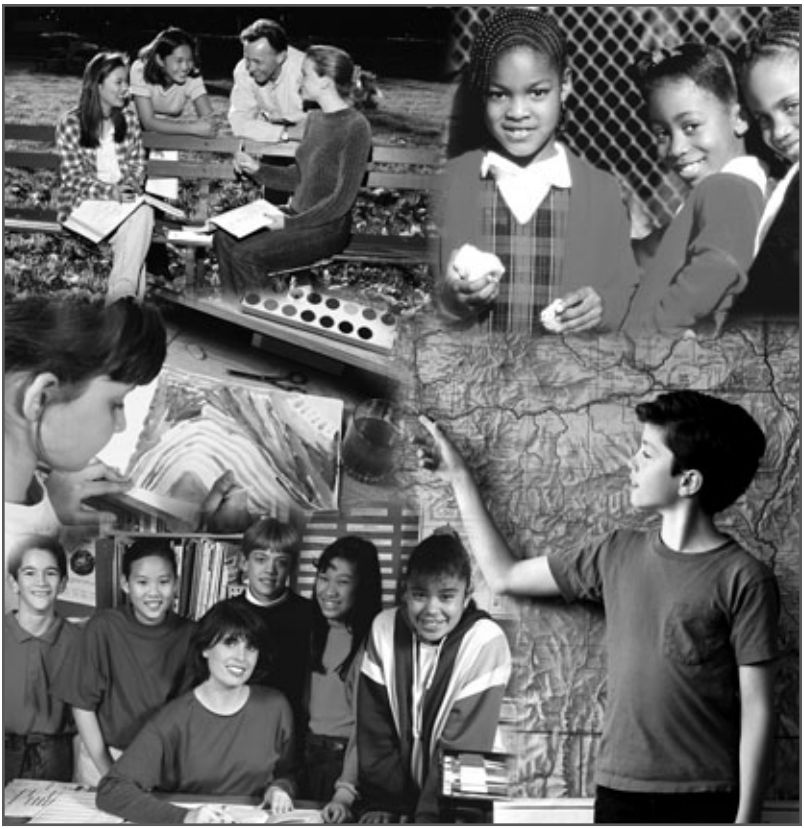
Marilyn A. Whirry
12th Grade English Teacher
Mira Costa High School
Manhattan Beach, California

Dennie Palmer Wolf
Director
Annenberg Institute for
School Reform
Brown University
Providence, Rhode Island

Grover (Russ) Whitehurst
(Ex-Officio)
Director
Institute of Education Sciences
U.S. Department of Education
Washington, D.C.

Charles E. Smith
Executive Director, NAGB
Washington, D.C.

Background Information Framework for the National Assessment of Educational Progress



National Assessment Governing Board
Adopted August 1, 2003

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

For further information, contact the
National Assessment Governing Board:

800 North Capitol Street, N.W.
Suite 825
Washington, DC 20002
www.nagb.org

Table of Contents

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.

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 the 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 in the 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 information framework will refocus the collection of non-cognitive variables on NAEP's primary mission: to provide 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, but its role in this respect

is limited and the framework is not a comprehensive list of possible factors to explore.

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 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.

(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. Although many non-cognitive variables may be of interest, they must be limited to meet the needs of NAEP reporting. In all cases, they must be clearly related to academic achievement or to the fair presentation of achievement results.

(3) ***Subject-specific background information*** should be gathered when 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.

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 noncognitive 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 for 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.

Data Collection

Whenever possible, NAEP should use information from school records and other reliable data sources 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 about 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 number of reading materials in the home. The framework provides that NAEP explore development of a composite index for SES derived from the proxy variables currently collected. To the extent that the index can be sharpened by additional data from readily available sources, such as zip codes and 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 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 each subject 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.

The National Center for Education Statistics (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.

In constructing questionnaires it is important to place strict limits on the burden they impose on respondents. 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: ten minutes for each student, 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 in 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) are of great public interest and can be related to academic results.

In addition, supplemental reports may be prepared that focus on particular aspects of the background data collected. 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.

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.

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. However, it cannot properly be used to establish direct cause-and-effect relationships. Still, 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 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 educational research. The NCES program of secondary analysis grants for researchers to analyze NAEP data should continue. Education 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.

Chapter One

Introduction

The National Assessment of Educational Progress is the only continuous long-term measure of student achievement in U.S. elementary and secondary schools. The primary purpose of NAEP 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: (1) academic assessments, which measure the achievement of students on a broad range of content, and (2) non-cognitive survey questions, which 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 the United States. The assessment does not report results for individual students, but only for groups with large, representative samples, for example, 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 among the states, districts, 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 "background questions," the phrase usually used by NAEP in the past for items designed to collect non-academic information. However, non-cognitive is also less readily understandable than background information, and so the two terms are 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 (SES)—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 opportunities 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.

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, and 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 standard and comprehensive 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.

Although 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, socioeconomic 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.

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.

(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. Although these types of non-cognitive variables are of interest, they must be limited so that they meet the needs of NAEP reporting. In all cases, they must be clearly related to academic achievement or to the fair presentation of achievement results.

(3) ***Subject-specific background 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.

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, as required by the No Child Left Behind Act, may become very 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.***

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 socio-economic status (SES), contextual factors of interest to public policy, and subject-specific variables.

Demographic Reporting Categories

The demographic variables currently collected by NAEP are 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. In addition, information is obtained on disability status and on students who are classified as limited English proficient. All of this information is collected on an administration roster, completed from school records in advance of testing. In addition, data on race/ethnicity is also collected on the NAEP student questionnaire, and students are asked to report on two other SES variables: the number of reading materials at home and the highest level of each parent's education.

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 or 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 by type of location, based on Census Bureau definitions, such as central city, suburban/large town, and rural/small town.

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 are 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, student responses are 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, and Zelenak, 1999).

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, which was 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. 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 ought to be made to find an approximation that is more reliable than the current set of individual proxy variables.***

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, 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 recently deleted from the NAEP student questionnaire 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 Board will have to weigh the considerations involved, and may decide there is value in using them periodically or in limited samples.

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. Research should be conducted into creating a composite index of SES.

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 education—though steps must be taken to ensure that such data are reliable. An expanded set of SES variables may be included periodically and administered to limited samples, including such factors as number of siblings and parents at home, possession of computers, and parent occupation.

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 well.

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, television 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. 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. Information 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 Indicators Report* (Mayer, Mullens, and 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 achievement 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, and 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 of questions for school administrators.

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

A design for doing this was presented to the Board in the 1996 report of the Design/Feasibility Team of prominent researchers (Forsyth 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. 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 2009, the years in which revised frameworks are first used, and then should be repeated every eight years. In the intervening years, only the more limited core modules will be administered. Similar patterns should be established for the school and teacher questionnaires.

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 subsamples. With states now 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.

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.

Table 1 presents a model schedule for comprehensive and core sets of subject-related variables through 2013. It is based on the schedule of assessments approved by the Board in May 2003.

Table 1. Model Data Collection Schedule for Comprehensive and Core Sets of Non-Cognitive Variables by Subject Area

Subject Area	Data Collection Year for Comprehensive Set of Variables	Data Collection Year for Core Variables Only
Reading	2003, 2009,	2005, 2007, 2011, 2013
Mathematics	2003, 2005, 2013	2007, 2009, 2011
Foreign Language (12)	2004, 2012	
World History (12)	2010	TBD
Economics (12)	2006	TBD
Civics	1998, 2012	2006
Writing	2002, 2011	2007
Arts (8)	1997, 2008	
Science	2000, 2009	2005
US History	2001, 2006	
Geography	2001, 2010	

NOTE: Based on schedule approved by NAGB on May 17, 2003.

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 4, therefore, the total NAEP sample approximates 436,000 students. The grade 8 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 12th 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. With implementation of the school report card requirements of the No Child Left Behind Act, much more information should be available soon 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.

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, Jones, and Mitchell, 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

Figure 1. Overview of Current NCES Data Collections

Data and Design Elements	NAEP	NELS	ELS	ECLS	TIMSS	CCD	PSUS	SASS	NHES
Data Elements									
Student achievement	x	x	x	x	x				
Student background characteristics	x	x	x	x	x	x	x	x	x
Home and community support for learning	x	x		x	x				x
Standards and curricula					x				
Instructional practices and learning resources	x	x		x	x			x	
School organization/governance					x			x	
Teacher education and professional development	x				x	x	x	x	
Financial resources					x	x	x	x	
School climate	x	x		x	x			x	x
Design Elements									
Type of design (CS=cross-sectional; L=longitudinal)	CS, L	L	L	L	CS	L	L	CS,L	CS
Periodicity (TBD=to be determined)	2, 4, or 6 yrs	2–6 yrs	TBD	TBD	TBD	Annual	Biennial	2–5 yrs	2–3 yrs
Unit of observation (S=student; T=teacher; A=administrator; P=parent; SC=schools; D=district; ST=states; H=households)	S,T,A	S,T,A	S,A,P	S,T,A,P	S,T,A,P	SC,D,ST	SC	T,A,SC	H
Data collection method (S=survey; R=record analysis; I=interview; V=video; C=case study; O=other)	S	S,R	S,O	S,O	S,R,V,C	S,R	S	S	I
Population of inference (N=national; S=state; G=demographic group)	N,S,G	N,G	N,G	N,G	N				

NELS: National Education Longitudinal Study of 1988
ELS: Educational Longitudinal Study of 2002
ECLS: Early Childhood Longitudinal Study
TIMSS: Third International Mathematics and Science Study
CCD: Common Core of Data
PSUS: Private School Universe Survey
SASS: Schools and Staffing Survey
NHES: National Household Education Survey

NOTE: From *Grading the Nation's Report Card: Evaluating NAEP and Transforming the Assessment of Educational Progress* (pp. 36–37), by J.A. Pellegrino, L.R. Jones, and K.J. Mitchell, 1999, Washington, DC: National Academy Press. Copyright 1999 by the National Academy of Sciences. Reprinted with permission.

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 as up-to-date as possible.

NAEP Background Information Matrix

The types of descriptive information NAEP collects, reporting units, and data sources are summarized in the NAEP Background Information Matrix, which is displayed as Figure 2. The matrix is intended to assist in conceptualizing NAEP background information collections. It is organized by priorities—both for types of information and for how data should be obtained. Note that in each case information is to be obtained from reliable official records before it is sought through questionnaires.

The entries in the cells are illustrative, showing the kinds of information that are currently collected by NAEP and the various data sources (records and questionnaires) that are used. As the principles of this framework are implemented, more information will come from records, less from questionnaires. The sources with higher reliability and less respondent burden should be utilized in priority order.

The Ad Hoc Committee on NAEP Background Questions considered a proposal by Paul Barton (2002) to permit states or groups of states to add customized sets of questions to the background questionnaires. Although these might track progress on topics of particular interest and increase support for NAEP, the Committee felt strongly that the proposal should not be pursued because any customization of NAEP questionnaires would create serious logistical and quality control problems.

In constructing questionnaires it is important to place strict limits on the respondent burden they impose. 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: ten minutes for each student, 20 minutes for each teacher, and 30 minutes for each school.

Figure 2. NAEP Background Information Framework

Reporting Unit and Data Sources	Type of Information			
	Student Reporting Categories	**Socio-Economic Status—Core or Expanded**	**Other Contextual Information**	**Subject-Specific Information**
STUDENT				
School Records	Gender Race/ethnicity SD/LEP	Free/RP lunch participation Title I	New enrollee Type/degree of disability	Course taking in mathematics
Questionnaire	Race/ethnicity	Parent education Reading materials in home	Daily reading Discuss school work TV-watching Absenteeism Language in home	Time spent on math homework Good in math?
SCHOOL				
Dist/State Recds **School Records**	School type (public, private, charter, etc.) School ach. data	% Free/RP lunch participation Title I funding	Grade structure Days of instruction Enrollment % LEP	Graduation requirements in math/science Higher level math courses Graduation testing
CCD/Census **Questionnaire**	Community type		% Students absent % Teachers absent Enrollment mobility Grade retention Teacher retention Graduation rates Post-secondary ed rates	Extracurricular options in math and English Availability of computers for writing
TEACHER				
School Records **Dist/State Recds**			Race and gender Experience Credentials	
Questionnaire			Undergrad/grad content training Professional devel	Correct for spelling and grammar? Frequency of lab work
STATE				
CCD/Census **State Records** **Questionnaire**	Region		Non-NAEP contextual variables	
DISTRICT				
CCD/Census **State Records** **District Records** **Questionnaire**		Community type (urban, rural, etc.)		

NOTE: Information type and data sources are arranged in priority order.

Chapter Five

Using Background Data to Report NAEP Results

This chapter discusses the descriptive information that NAEP should provide, the levels of disaggregation now possible with merged national and state samples, and the importance of minimizing causal interpretations.

Use of Descriptive Information in NAEP

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 the patterns and trends of non-cognitive variables known to have a relationship to academic achievement.

In addition, supplemental reports may be prepared that focus on particular aspects of the background data collected. 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.

All background questions and data collected by NAEP should be made available on the Internet at the time of the initial release of the principal academic results or soon afterwards so the public may be able to consider them in discussing results. Complete data files should be available to researchers for further analysis.

Implementing No Child Left Behind

The intent of the No Child Left Behind Act of 2001 (P.L.107-110) is to hold public schools accountable for closing the achievement gaps between different groups of students. NAEP is asked to contribute to this end by providing an accurate measure of the current levels of student achievement and to monitor change over time.

Descriptive information about all students, but particularly on low-performing groups, would contribute powerfully to the dialogue on the challenges before American education. For example, the NAEP achievement levels focus on the segments of the performance distribution that are at or above *Basic, Proficient,* and *Advanced.* Information should also be provided about those *Below Basic,* who clearly have been "left behind": e.g., the proportion having qualified teachers, receiving free or reduced-price lunch, or moving to different schools frequently, as measured by attending the same school for less than two years.

Such profiles of low-performing or high-performing students would not attempt to ascribe causation, but they would provide information on the distribution of practices and resources that are of concern to the public and policy-makers. Periodic collections of such background data could be used to track change in the distribution of these factors over time. Do the trends seem favorable or adverse to educational progress?

Disaggregation of NAEP Data

For more than three decades NAEP has provided data disaggregated by race/ethnicity, gender, school type (e.g., public/private), and community type (e.g., urban/rural). The No Child Left Behind law calls for disaggregation by major subgroups (when feasible) of race, ethnicity, and gender, and also by socio-economic status, disability, and limited English proficiency.

Because of the large size of the recently combined national and state NAEP samples, NAEP reports should be able to provide information disaggregated at a much greater level of detail than was previously possible. Pooling the data from all states, which now are required to provide NAEP samples in 4th and 8th grade reading and mathematics, will produce a much-enlarged national sample that will sharply reduce the number of empty cells in any cross-tabulations. Such disaggregation might add to the richness of NAEP reporting even with only a limited set of non-cognitive questions. Disaggregation is also very important for reporting on the distribution of student characteristics within the different achievement levels, as described above.

Minimizing Causal Interpretations

NAEP has often reported on the average performance of students by particular non-cognitive variables. One example, presented in many NAEP reports, is the average scale score of students who watch different amounts of television each day, cf. *The Nation's Report Card: Reading, 2000* (Donahue et al., 2001). Another example is the average scale scores for 12th graders who report different amounts of time working at a part-time job, cf. *The Nation's Report Card: Mathematics, 2000* (Braswell et al., 2001).

While there may be a correlation between television watching and reading performance, or between hours working outside school and math results, NAEP is not designed to prove cause-and-effect relationships. As a cross-sectional survey, nearly all of its data are on current activities and practices—not on the complex chain of experience in school and outside, of prior learning and achievement that all contribute heavily to current academic performance. Yet, NAEP has encouraged simple causal inferences by reporting average scores for varying amounts of time spent on current activities.

There is one important exception to the absence of data on learning-related activity over time. This is the information NAEP collects on the transcripts of high school seniors and its questionnaires on courses that students have taken and schools provide. These do show prior instruction before current exams. The trends in course taking have been of great public interest and it is reasonable to relate them to student achievement.

NAEP reports should present information on the patterns and trends of non-cognitive variables known from other sound research to have a relationship to academic achievement. These presentations should be straightforward and impartial, and care must be taken to avoid stating conclusions as to cause and effect relationships. Further analysis of any relationships should be left to researchers.

Chapter Six

Using NAEP in Educational Research

As a cross-sectional survey without longitudinal data, the National Assessment of Educational Progress 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, 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 doing its main task of monitoring achievement well, NAEP has provided sound data for researchers to use. NAEP results have been critical in identifying hypotheses for other research to pursue. 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 the National Assessment can be of value to educational research and that the interest of researchers in the assessment should be encouraged. The NCES program of secondary analysis grants for researchers to use NAEP data should continue. Education researchers should be involved, under the auspices of NCES and its contractors, in developing NAEP background questionnaires and other data collection efforts to carry out the provisions of this framework.

This chapter considers the limitations and strengths of NAEP for educational research and discusses research that has made use of NAEP data. The chapter draws on papers by David Grissmer, senior research scientist at RAND, who has used NAEP extensively in analyzing educational factors and trends.

NAEP's Limitations and Strengths for Research

The primary purpose of NAEP is to ***accurately and fairly monitor achievement*** over time ***and accurately and fairly compare achievement across states and important sub-groups of students.***

Beyond providing such data, any research based on NAEP, particularly into the causes of academic achievement, is severely limited by its design.

As a representative sample survey, in which no individual student takes more than a small part of the full exam, NAEP has shortcomings in most of the elements commonly used to evaluate academic achievement (Podgursky, 2002):

- It provides no prior data on student achievement, and cannot be changed to gather longitudinal data.
- It can only collect contemporaneous information on school practices and resources, and has no way of ascertaining how students were taught or what school experiences they may have had in previous years.
- There is considerable measurement error in survey responses obtained from teachers and schools because they may well give the expected "right" answers rather than report accurately what they do.
- The current classroom practices that teachers report may be a response to student achievement levels, not the cause of such achievement, and it is difficult to disentangle causation.
- It is difficult for NAEP to get good information on socioeconomic status or family background factors, but these are powerfully correlated with academic achievement, and must be controlled for in any analysis of school effects.

On the other hand, NAEP does have unique strengths and comparative advantages (Grissmer, 2003), and thus has the potential to address some important research and public policy questions with its cognitive data and background information:

- NAEP is the only data set on student achievement that has collected data from nationally representative samples of students continuously from 1969–70 to the present.
- It is the only data set that has collected academic achievement data simultaneously, repeatedly, and consistently from three separate age groups.
- It is the only data set that collects from statistically reliable samples at the state level, and within states for different types

of communities (central city, suburban and rural) and for racial/ethnic groups within most states.

- NAEP has far larger sample sizes than any other nationally representative survey of student achievement, such as the National Education Longitudinal Study (NELS) and the Early Childhood Longitudinal Study (ECLS). These surveys are only approximately 10 to 20 percent as large as NAEP in any single application, and 1 to 5 percent as large as NAEP in any repeated data collection.
- NAEP is the only survey that tests a wide range of academic subjects.
- NAEP achievement measures at 4th and 8th grade fill an important void in measuring the well-being of children during this developmental period.
- NAEP generally incorporates a higher quality and unique design of test instruments, administrative procedures, and scoring methodology, compared to other data sets.

Previous Use of NAEP in Research

As a result of its strengths, NAEP has been used in important educational research by authors such as David Grissmer, Alan Krueger, David Armor, and Christopher Jencks. These studies point to an important comparative advantage of NAEP, namely, that it is the only representative sample data in existence on student achievement in the United States from 1969 to 2002. Thus, research into important historical questions about the effects of changing families, communities, and schools on achievement almost require NAEP data. Without NAEP, it is unlikely that the significant narrowing of the black-white score gap would be known and its possible causes the subject of research.

Similarly, NAEP data have been used to help analyze the effects of differences in resources, systemic reform initiatives, differential opportunity for learning, and other educational policies on state-level academic achievement. Such research has concluded that the rates of improvement in achievement varied markedly across states in the 1990s, and that changing resources or demographics cannot account for the gains in the states with most rapid improvement. This

research points to another strong comparative advantage of NAEP. State NAEP is the only survey that includes representative samples of students in many different states, and thus plays a central role in monitoring and explaining the differences in academic achievement and achievement trends across the states. NAEP can identify where positive trends are occurring so researchers can puzzle out causation.

A review of research studies using NAEP (Grissmer, 2003) suggests that only a small proportion of the non-cognitive items collected by the assessment have been utilized in productive research. Also, such research has often supplemented NAEP with data from other sources, such as the U.S. Census and the Common Core of Data (CCD) and Schools and Staffing Survey (SASS), both conducted by the National Center for Education Statistics. However, the National Assessment played such a crucial role in these studies that they could not have been conducted without NAEP data, including some of its non-cognitive variables, principally those on socio-economic status, family structure, and school resources.

On the other hand, NAEP data have also been misused for simplistic and weak research. Many background items on school practice and student behavior have been used in a simplistic way to imply a direct, causal relationship to achievement while ignoring the complex mix of other, more fundamental factors that may well have a stronger impact. NAEP has encouraged such associations by presenting one-way tabulations in its reports, such as average scale score by hours of television watched, type of reading instruction, or books read per week, and these have been disseminated widely to support particular beliefs or public policy positions. Simple, single-variable linkages can often be misleading because of the strong correlations between many background variables, particularly with socio-economic status, prior academic achievement, or family background. They should only be included in NAEP reports when there is strong justification based on previous research.

Also, most of the hundreds of background questions in NAEP have never been used for either public reporting or research. Many come from the early 1980s, and would be difficult to justify in a sound research design today.

Secondary Analysis Grants and District Samples

For many years NCES has been making awards to education researchers for secondary analyses of NAEP data. These have explored a range of topics, often in combination with other data sets. Many of the studies have focused on state-to-state differences in student achievement and the impact of state-level policies, relying on NAEP academic data, a few background questions for SES controls, and much additional information from other sources. The program has been valuable as a means of encouraging the use of NAEP for research, and, in a few cases, notably the Grissmer studies, has had considerable impact. As in any grant program, all findings are the responsibility of the individual researchers, not of the agency making the grant.

The program should continue, and now that NCES has become part of the Institute for Education Sciences, the leadership of the new agency should ensure that the analysis grants are aligned with the research priorities of the Institute.

In addition, data from the school district NAEP samples in the Trial Urban District Assessment, started in 2002, will provide important new opportunities for research. NAEP results for school districts can readily be combined with Census data, which include pertinent information on family background and socio-economic status. The school district samples can also be tied to important education policy variables, such as per-pupil spending, for which information is available at this level but not for schools.

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.

Chapter Seven

Review and Improvement of Non-Cognitive Questions

This chapter discusses several mechanisms for the review and improvement of NAEP's non-cognitive questions and for implementation of the NAEP Background Information Framework.

Independent Validity Studies

Since the early 1990s NAEP has had the benefit of independent outside advice on topics of urgency or interest. These studies have been very helpful to the Governing Board and NCES as they made decisions about the future of the NAEP program. For example, several years ago research was conducted on the possibility of combining the national and state samples in NAEP to achieve greater efficiency and cost-savings. Starting in 2003 NAEP has moved in that direction. The decisions surrounding such change, however, can only be as good as the research that informs them. The work of the current NAEP Validity Panel, in conjunction with the current NAEP operations contractors, contributed significantly to making the change possible.

The value of this kind of applied research cannot be overestimated. Neither can the value of the independent nature of such work. The NAEP program is very large and complex and demands a commitment of many resources from the NAEP contractors. NAEP contractors should not be burdened with conducting simultaneous research studies while carrying out the requirements of the operations contracts. There is a precedent for this approach in the current separation of responsibilities for operations and research in separate NAEP contracts. There are two reasons why independent validity studies on topics associated with the non-cognitive framework are recommended. First, there are some non-cognitive variables that will need validation, particularly if those variables are new or are new composite indices of existing variables. Second, following the approach already

established for the cognitive components of NAEP, recommendations from research studies should be truly independent and free from any conflict of interest.

Review of Background Information Framework

This background information framework should be reviewed on a periodic basis. The NAEP cognitive frameworks are reviewed every ten years. This policy was adopted at the time of the NAEP redesign in 1996. Reviewing a NAEP framework can result in major revision, minor revision, or even no revision and re-adoption. Since the background framework is a new undertaking, a required review after five years is appropriate with additional reviews every ten years thereafter.

References

Allen, N.L., Carlson, J.E., and Zelenak, C.A. (1999). *The NAEP 1996 technical report.* Washington, DC: U.S. Government Printing Office.

Armor, D.J. (December 18, 2002). *Comments on NAEP non-cognitive questions.* Available from the National Assessment Governing Board, Washington, DC.

Barton, P.E. (2002, September). *Perspectives on non-cognitive questions in the National Assessment of Educational Progress.* Paper presented at workshop on NAEP background questions. National Assessment Governing Board, Washington, DC. Available at http://nagb.org.

Braswell, J.S., Lutkus, A.D., Grigg, W.S., Santapau, S.L., Tay-Lim, B., and Johnson, M. (2001). *The Nation's report card: Mathematics 2000.* Washington, DC: National Center for Education Statistics.

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D., and York, R.L. (1966). *Equality of Educational Opportunity.* Report of the Equality of Educational Opportunity Commission (EEOC). Washington, DC: National Center for Education Statistics.

Donahue, P.L., Finnegan, R.J., Lutkus, A.D., and Allen, N.L. (2001). *The Nation's report card: Reading 2000.* Washington, DC: National Center for Education Statistics.

Donahue, P.L., Voelkl, K.E., Campbell, J.R., and Mazzeo, J. (1999). *NAEP 1998 reading report card for the nation and the states.* Washington, DC: National Center for Education Statistics.

Forsyth, R., Hambleton, R.K., Linn, R., Mislevy, R., and Yen, W. (1996). *Design/feasibility team report to National Assessment Governing Board.* Washington, DC: National Assessment Governing Board. Available at http://nagb.org.

Greenberg, E., Stancavage, F., Farr, B., and Bohrnstedt, G. (January 2001 draft). *Making comparisons: Using background information from the Nation's report card to understand student achievement.* Washington, DC: American Institutes for Research.

Grissmer, D. (2002, September). *Improving NAEP for research and policymaking.* Paper presented at workshop on NAEP background

questions. National Assessment Governing Board, Washington, DC. Available at http://nagb.org.

Grissmer, D. (2003, January). *Utilizing NAEP for educational research.* Paper prepared for National Assessment Governing Board, Washington, DC. Available at http://nagb.org.

Mayer, D.P., Mullens, J.E., and Moore, M.T. (2001). *Monitoring school quality: An indicators report.* Washington, DC: National Center for Education Statistics.

Mullis, I.V.S. (2002, September). *Non-cognitive questions in TIMSS and PIRLS: An overview.* Paper presented at workshop on NAEP background questions. National Assessment Governing Board, Washington, DC. Available at http://nagb.org.

National Assessment Governing Board. (1996, August). *Policy statement on redesigning the National Assessment of Educational Progress.* Washington, DC: Author.

National Assessment Governing Board. (2002, May). *Policy statement on collection and reporting of non-cognitive data by the National Assessment of Educational Progress.* Washington, DC: Author.

Pellegrino, J.W., Jones, L.R., and Mitchell, K.J. (eds.) (1999). *Grading the Nation's report card: Evaluating NAEP and transforming the assessment of educational progress.* Washington, DC: National Academy Press.

Podgursky, M. (2002, September). *NAEP non-cognitive questions: What can we learn from NAEP about the effect of schools and teachers on student achievement?* Paper presented at workshop on NAEP background questions. National Assessment Governing Board, Washington, DC. Available at http://nagb.org.

Public Law 107-110. *No Child Left Behind Act of 2001.* Section 602. Available at http://nagb.org.

Reckase, M.D. (2002, September). *Contributions of non-cognitive questions to improving the precision of NAEP results.* Paper presented at workshop on NAEP background questions. National Assessment Governing Board, Washington, DC. Available at http://nagb.org.

Reese, C.M., Miller, K.E., Mazzeo, J., and Dossey, J.A. (1997). *NAEP 1996 mathematics report card for the nation and the states.* Washington, DC: National Center for Education Statistics.

Rust, K. (September 26, 2002). *Technical memo: Sample design for 2003 NAEP.* Available from NAGB Board Meeting briefing book, November 14–16, 2002. National Assessment Governing Board, Washington, DC.

Walberg, H.J. (2002, September). *The National Assessment's most useful non-cognitive items.* Paper presented at workshop on NAEP background questions. National Assessment Governing Board, Washington, DC. Available at http://nagb.org.

Appendix

Adopted May 18, 2002

National Assessment Governing Board

Policy Statement on Collection and Reporting of Background Data by the National Assessment of Educational Progress

Introduction

As the Nation's Report Card, the National Assessment of Educational Progress (NAEP) is an on-going, Congressionally-authorized program to collect data through surveys on the academic knowledge and skills of American students. Its primary goal is to report fair and accurate information on student achievement in reading, mathematics, and other subjects taught in elementary and secondary schools. This information is to be made available in a clear and timely manner to members of the public, policymakers, and educators throughout the country.

Since it began in 1969–70, NAEP has administered, in addition to cognitive questions, background questionnaires that provide information for reporting categories and collect non-cognitive data on students, their family background, teachers, and schools. These have

enriched reporting of the National Assessment and increased the precision of NAEP results. The background data have also been used in secondary analyses. However, because NAEP tests a cross-section of students at a particular time with no follow-up of the students tested, the assessment can only show correlations or associations rather than causal relationships between background factors and achievement.

By statute (P.L. 107-110), the National Assessment shall include, "whenever feasible, information collected, cross-tabulated, compared, and reported by race, ethnicity, socioeconomic status, gender, disability, and limited English proficiency." The statute provides that NAEP may "not evaluate or assess personal or family beliefs and attitudes" and may "only collect information that is directly related to the appraisal of academic achievement and to the fair and accurate presentation of such information." These provisions are intended to prevent intrusive, inappropriate, or unnecessary questions being asked about students and their families.

The law requires that the Governing Board take steps to ensure that all NAEP questions are "free from racial, cultural, gender, or regional bias, and are secular, neutral, and non-ideological." However, a House-Senate Conference report, accompanying the legislation, says the law does not preclude the use of "non-intrusive, non-cognitive questions," with a direct relationship to academic achievement.

The National Assessment is conducted by the Commissioner of Education Statistics under the policy guidance of the National Assessment Governing Board. The Board's specific areas of responsibility include: (1) assessment objectives and test specifications; (2) the methodology of the assessment; (3) guidelines for reporting and disseminating results; and (4) "appropriate actions needed to improve the form, content, use, and reporting" of the National Assessment. Under the statute, the Board has "final authority" on the appropriateness of all NAEP items—both cognitive and non-cognitive.

To carry out these responsibilities, the National Assessment Governing Board hereby adopts guiding principles, policies, and procedures for the collection and reporting of background data by the National Assessment of Educational Progress.

Guiding Principles

1. Background data on students, teachers, and schools is needed to fulfill the statutory requirement that NAEP include information, whenever feasible, disaggregated by race or ethnicity, socioeconomic status, gender, disability, and limited English proficiency. In addition, background data is collected to enrich the reporting of NAEP results by examining factors related to academic achievement. However, the collection of such data should be limited, and the burden on respondents kept to a minimum. It must always be considered in light of NAEP's primary purpose: providing sound, timely information on the academic achievement of American students.
2. All background questions must be directly related to academic achievement or to the fair and accurate presentation of achievement results.
3. Issues of cost, benefit, appropriateness, and burden shall be carefully considered in determining the background questions to be asked and the samples to which they shall be administered.
4. In accordance with law, questions shall be non-intrusive and free from bias, and must be secular, neutral, and non-ideological.
5. No personally identifiable information shall be included in NAEP reports or data releases.
6. Decisions on the retention or addition of background items shall draw on technical studies on the reliability and validity of current and proposed questions and on the contribution such items make to the precision of NAEP results.
7. Consideration should be given to obtaining background information from non-NAEP sources and to avoiding duplication with other federal surveys.
8. Questionnaires should be revised to keep background questions timely and related to academic achievement. Those questions showing little change over time and/or a stable relationship to achievement should be deleted or asked less frequently and to limited samples, unless required to ensure the precision of NAEP results.

9. Questions should not address personal feelings and attitudes.
10. Since security considerations do not apply, background questionnaires shall be readily available to the public.
11. Interpretation of results shall be limited in official NAEP reports and must be strongly supported by NAEP data. Because of the survey nature of the assessment, reports may show correlations and generate hypotheses, but may not state conclusions as to cause and effect relationships.
12. Background questions for NAEP assessments shall be prepared in accordance with frameworks and specifications adopted by the Governing Board.
13. The Governing Board shall review and approve all background items before they are administered in NAEP surveys or pilot and field tests.

Policies and Procedures

1. Framework and Specifications

The Governing Board will adopt a general framework for background questionnaires and specifications for the questions on selected topics and in specific subject areas.

Since this is a new area of responsibility for the Board, the process of developing a framework for background questions and specifications will begin with commissioned papers on relevant issues, such as the reliability and validity of current background questions, their contribution to improving the precision of NAEP results, their value and limitations for educational research, and changes that may be needed in response to the No Child Left Behind legislation. Following consideration of these issues, the Board will define the scope of background questionnaires and adopt a process for preparing a framework and specifications. This work will include the active participation of teachers, education researchers, state and local school administrators, assessment specialists, parents of children in elementary and secondary schools, and interested members of the public.

2. Background Question Development

In preparing background questions, the National Center for Education Statistics shall follow adopted frameworks and specifications, and consider the review criteria adopted by the Governing Board. NCES may use cognitive laboratories of students, teachers, and school officials to help determine the clarity and burden of proposed questions. Ad hoc advisory committees may also be established, comprised of teachers, parents, technical experts, and others interested in NAEP. Steps shall be taken to determine the reliability of questions used.

3. Governing Board Review and Approval of Background Questions

Background questions for all NAEP pilot tests, field tests, and operational use shall be reviewed and approved by the Governing Board. The category of respondents, e.g. students, schools, and grade level, shall clearly be designated, as will the NAEP samples, e.g. national, state, or district, in which the questions will be asked.

For each questionnaire there shall be an explanation of its intended use in NAEP reporting and analysis and of the hypothesized relationships between the background items and student achievement that demonstrates the need to know such information. Technical data shall be presented on the reliability and validity of questions and, if applicable, on their contribution to improving the precision of NAEP results. The Board will use the explanations and data presented along with the review criteria in this policy statement in determining the appropriateness of background questions.

The Reporting and Dissemination Committee will have primary responsibility for the review and approval of background questions. The Assessment Development Committee will participate in the approval of questions relating to specific subject-matter assessments. Ad hoc committees of Board members may be established by the Board Chairman for background question review. Questions may also be reviewed by external advisors, including teachers, parents, and technical experts. Recommendations on background questionnaires shall be subject to final approval by the full Governing Board.

4. Criteria for Governing Board Review

The following criteria for review and approval of background questions are based on the most recent revision of the authorizing statute of the National Assessment of Educational Progress (P.L. 107-110) and the Guiding Principles of this policy statement:

A. Background information is needed to fulfill the statutory requirement that NAEP report and analyze achievement data, whenever feasible, disaggregated by race or ethnicity, gender, socio-economic status, disability, and limited English proficiency. Non-cognitive data may enrich the reporting and analysis of academic results, but the collection of such data should be limited and the burden on respondents kept to a minimum.

B. All background questions must be related to the primary purpose of NAEP: the fair and accurate presentation of academic achievement results.

C. Any questions on conditions beyond the school must be non-intrusive and focused on academic achievement and related factors.

D. Questions shall be free from racial, cultural, gender, or regional bias.

E. All questions must be secular, neutral, and non-ideological. Definitions of these terms, accompanied by clarifying examples, are presented in Attachment A, as adopted in the Governing Board Policy on NAEP Item Development and Review.

F. NAEP must not evaluate or assess personal feelings or family beliefs and attitudes unless such questions are non-intrusive and have a demonstrated relationship to academic achievement.

G. Issues of cost, benefit, appropriateness, and burden shall be carefully considered in determining which questions to include in background questionnaires. These factors must also be considered in determining the frequency with which various questions shall be administered and whether they shall be included in both national and state samples.

H. Background questions that do not differentiate between students or have shown little change over time should be deleted or asked less frequently and to limited samples.

5. Public Access to Background Questions

Since security considerations do not apply, all background questionnaires shall be readily available to parents, teachers, state and local officials, and interested members of the public. Such questionnaires shall be available before field tests and operational assessments or at any other time members of the public wish to obtain them. Background questions in operational use shall be posted on the Internet prior to each assessment, accompanied by explanations and rationales.

6. Reporting of Background Information

The presentation of background data in official NAEP reports shall be straightforward and impartial. Because of the survey nature of the assessment, reports may show correlations and generate hypotheses, but may not state conclusions as to cause and effect relationships. Any composite indices including demographic and socioeconomic factors shall be presented to the Board for approval before use as reporting categories in NAEP data releases and reports.

Background data should be available for extensive secondary analyses by scholars and researchers, who are responsible for conclusions reached. Responses to background questions shall be presented and tabulated on the Internet, although, if necessary, posting may be delayed for a brief period after release of the principal NAEP results.

Attachment A

Definitions of Secular, Neutral, and Non-Ideological Item Review Criteria

From Governing Board Policy on NAEP Item Development and Review—5/18/02

Items shall be secular, neutral, and non-ideological. Neither NAEP nor its questions shall advocate a particular religious belief or political stance. Where appropriate, NAEP questions may deal with

religious and political issues in a fair and objective way. The following definitions shall apply to the review of all NAEP test questions, reading passages, and supplementary materials used in the assessment:

Secular—NAEP questions will not contain language that advocates or opposes any particular religious views or beliefs, nor will items compare one religion unfavorably to another. However, items may contain references to religions, religious symbolism, or members of religious groups where appropriate.

Examples: The following phrases would be acceptable: "shaped like a Christmas tree," "religious tolerance is one of the key aspects of a free society," "Dr. Martin Luther King, Jr., was a Baptist minister," or "Hinduism is the predominant religion in India."

Neutral and Non-Ideological—Items will not advocate for a particular political party or partisan issue, for any specific legislative or electoral result, or for a single perspective on a controversial issue. An item may ask students to explain both sides of a debate, or it may ask them to analyze an issue, or to explain the arguments of proponents or opponents, without requiring students to endorse personally the position they are describing. Item writers should have the flexibility to develop questions that measure important knowledge and skills without requiring both pro and con responses to every item.

Examples: Students may be asked to compare and contrast positions on states rights, based on excerpts from speeches by X and Y; to analyze the themes of Franklin D. Roosevelt's first and second inaugural addresses; to identify the purpose of the Monroe Doctrine; or to select a position on the issue of suburban growth and cite evidence to support this position. Or, students may be asked to provide arguments either for or against Woodrow Wilson's decision to enter World War I. A NAEP question could ask students to summarize the dissenting opinion in a landmark Supreme Court case.

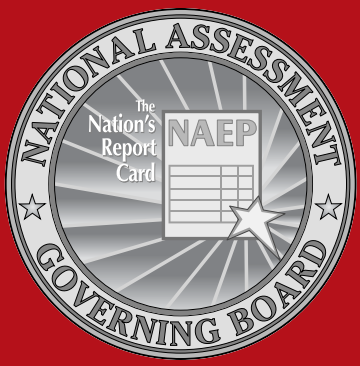
The criteria of neutral and non-ideological also pertain to decisions about the pool of test questions in a subject area, taken as a whole. The Board shall review the entire item pool for a subject area to ensure that it is balanced in terms of the perspectives and issues presented.

Acknowledgments

The Ad Hoc Committee on NAEP Background Questions of the National Assessment Governing Board was chaired by Board member John H. Stevens. He also serves as chairman of the Board's standing Committee on Reporting and Dissemination, which has responsibility for reviewing all core NAEP background questionnaires and making recommendations on them to the full Board.

The Ad Hoc Committee also included members of the Board's two other standing committees—the Assessment Development Committee and the Committee on Standards, Design, and Methodology—with a wide range of backgrounds and perspectives: Amanda Avallone, Dwight Evans, Thomas Fisher, Sheila Ford, Jo Ann Pottorff, and Sister Lourdes Sheehan. The Board chairman, Darvin Winick, participated in many of the lively discussions that marked the Committee's deliberations.

Among the many discussants and presenters at the workshop and public forum, we wish to recognize the care and thoughtfulness of Robert Mislevy, of the University of Maryland, and Harold Wenglinsky, of Baruch College of the City University of New York. The comments submitted by Sandra Feldman, president of the American Federation of Teachers, were particularly perceptive and helpful. This project also benefited greatly from the continuing advice and insight of Paul Barton and David Grissmer, both of whom have used NAEP data for many years to understand and explain American education to its public.


NATIONAL ASSESSMENT
GOVERNING BOARD
The
Nation's
Report
Card
NAEP