Improving the Measurement of Socioeconomic Status for the National Assessment of Educational Progress:

A THEORETICAL FOUNDATION

Recommendations to the National Center for Education Statistics

NOVEMBER 2012

TABLE OF 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.

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.

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 educational attainment, and parental occupational status. 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

educational attainment, and parental occupation (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 defini-

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.

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

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.

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 income, low educational attainment, or working in low-status jobs 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:

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

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),
- 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), and
- 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) educational attainment of heads of household, and (c) occupational status 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.

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 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.

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.

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 educational attainment, and parental occupational status. 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 education, and parental

occupational status, 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., 4th-grade reading, 8th-grade mathematics, 12th-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 family income, parental educational attainment, parental occupational status), income, educational attainment, and occupational status could be standardized separately within 4th, 8th, and 12th 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 4th-graders versus 8th- or 12th-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.

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 20th, 50th, and 80th 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.

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. *Peabody 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%20 seminar/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.