Reporting and Dissemination Committee

February 20, 2024
11:00 am – 12:30 pm
Zoom

AGENDA

11:00 – 11:30 am	**Strategic Communications Update**	
	Laura LoGerfo, Assistant Director for Reporting and Analysis	
11:30 am – 12:00 pm	**Reporting of New Socioeconomic Status (SES) Index**	*Attachment A*
	Ebony Walton, National Center for Education Statistics	
	William Ward, National Center for Education Statistics	
12:00 – 12:30 pm	**Member Discussion**	
	Marty West, Chair	

Reporting of New SES Index

Goal

This session at the Reporting and Dissemination (R&D) Committee meeting on February 20, 2024 will explain the development and testing of a new socioeconomic status index on NAEP.

Overview

Ebony Walton and William Ward of the National Center for Education Statistics (NCES) will describe the new NAEP socioeconomic status (SES) index, which will debut on the 2024 *Nation's Report Card*. They will cover issues that arose with using eligibility for the National School Lunch Program as an indicator for economic disadvantage and efforts to identify better ways to measure SES. Ward will describe the work conducted to build the new SES index, and Walton will describe the current analytic steps to evaluate missingness and trend issues prior to finalizing the reporting plan for the NAEP SES index. Time is allotted for questions.

Background

Three years ago, R&D Committee members learned about three different ways to measure SES, including this new NAEP SES index. To prepare for this session, a summary of that 2021 discussion follows, which is taken directly from the March 1, 2021, committee report:

Vice Chair Marty West convened several panelists to discuss the measurement and interpretation of socioeconomic status in NAEP. West introduced the topic by noting that the committee has long bemoaned NAEP's reliance on the increasingly convoluted indicator of student eligibility for free and reduced-price lunch to capture socioeconomic status (SES). Similarly, the idiosyncratic items about socioeconomic status on the student contextual questionnaire seem weak in comparison to those in other data collections. The Reporting and Dissemination Committee is not alone in their dissatisfaction as members of the Committee on Standards, Design and Methodology as well as the Assessment Development Committee also find these items lacking.

This session intended first to present approaches for improving the measure of SES underway by the NAEP team and second to offer alternative measurement methods. To the first intention, William Ward of NCES and Markus Broer of the American Institutes for Research (a NAEP contractor) shared insights into research and development work within NAEP. To the second intention, Thomas Kane of the Harvard Graduate School of

Education and Rick Hanushek of Stanford University's Hoover Institution (and Governing Board member) lent the committee members insights from their research.

Ward explained the fundamental assumptions which underlie NAEP's construct of SES. A measure of SES for the NAEP program must be useful, relevant to educational outcomes, and work in similar ways across grades 4, 8, and 12. The measure should comprise existing variables, so that any new iteration of SES can be applied to previous data to chart critical trends. To conduct such trend analyses, NAEP needs to measure the same construct over time so that changes in scores reflect changes in what students know and can do and not changes in the variable or construct. This criterion imposes a strict limitation, but panelists did offer some tentative solutions to this constraint.

Ward reminded the committee how NAEP currently captures SES, which is through eligibility for the National School Lunch Program (NSLP). This strategy suffers from variable reliability across grades, large within-category differences, and changes in eligibility across time, among other issues. Ideally, SES comprises three sources of information: (1) parental educational attainment; (2) parental occupational status; and (3) family income. The expert committee which advises NAEP on collecting contextual data recommended measuring school-level SES and neighborhood-level SES, then combining those with student-level SES to build an expanded measure of SES.

However, challenges emerge. Specific items from the contextual questionnaire are vulnerable to change over time, such as outdated references to brand-name technology like Nooks or to items that are so prevalent they no longer distinguish SES categories. Other challenges reside in missing or erroneous data, which derive from a wide range of reasons, from states opting out of the student questionnaire to students' lack of knowledge on particular items, e.g., about a quarter of fourth-graders do not know their parents' education.

Markus Broer then described his investigations into a new and improved measure of SES with extant NAEP data so as to analyze trends from 2003. His measure includes:

- number of books at home;
- students' eligibility for NSLP;
- percent of students eligible for NSLP at school the student attends; and
- parents' highest level of education.

Broer finds that NSLP retains its value, validity, and power to explain variance in assessment performance, despite the aforementioned issues with the variable. At grade 4, the fourth component of the index--parents' highest level of education--is excluded, due to inaccurate reporting or missing data. When analyzed, Broer's index explains

more variance in NAEP scores than NSLP alone and than measures of SES employed by large-scale assessment programs like the [Trends in International Mathematics and Science Study](#) (TIMSS). This index also shows expected correlations with other measures and explains achievement at the national level, at the state level, at the TUDA level.

To present an alternative approach, Professor Tom Kane shared his results from modeling the relationship between income and achievement in the NAEP data. Kane noted disagreements in the field about how achievement gaps by SES have changed over time, with Sean Reardon at Stanford revealing a significant widening in the achievement gap and Rick Hanushek and others seeing a flattening or narrowing of the gap. Kane challenged the foundation of those disagreements by arguing that SES does not serve as a good proxy measure for income and advocated for using income alone.

Kane enumerated weaknesses in different methods to capture income through NAEP, such as imputing income from students' race, maternal education, state, and urbanicity. Other attempts include matching school locations to neighborhood mean income from Census data. But this method is vulnerable to inaccuracy from increasing prevalence of school choice, and the vast majority of variance in parental income lies within schools, not between schools. Kane's third alternative approach -- of adding a parent questionnaire for a subsample of students and schools -- would require changes to several laws.

The strongest, most valid alternative approach to measuring income on NAEP, posited Kane, is by linking NAEP data to Census data. Kane reassured the committee that this approach would safeguard student privacy. Specifically, students would enter their addresses into the secure system provided by NAEP. The device would match the students' address to a neighborhood, for which the mean income would be drawn from Census data. The students' address would be deleted from the device or platform after the match is made and only the neighborhood mean income level would be retained and assigned to students' records.

Following Kane's conclusion, Rick Hanushek presented his thoughts on measuring income and SES. Hanushek and his colleagues have combined outcomes and SES measures from several sources, including Long-Term Trend NAEP, Main NAEP, TIMSS, and the Programme for International Student Assessment (PISA) to find that achievement gaps by SES have declined over the last fifty years.

Hanushek averred that there is no standard or accepted measure of SES. As such, measuring SES gaps over time presents even greater difficulties than maintaining trend in NAEP. Given that, and the challenges explained by Bill Ward, factoring parental education together with items in the home represents the best strategy to capturing

SES. Hanushek admitted the sources of uncertainty in NAEP, such as accuracy in students' replies, the burden on students, and shifted meanings of constructs over time.

To illustrate his points, Hanushek showed how the construct "items in your home" quickly becomes outdated. In 1990, the NAEP questionnaire asked students if they have encyclopedias in their homes and whether their families regularly receive deliveries of newspapers and magazines. Even a NAEP respondent in 2003 would find a question about encyclopedias amusing at best and confusing at worst. He then walked the committee members through graphs showing the relationship between SES and achievement across different datasets.

These illuminating presentations provoked thoughtful questions from the committee members. West invited members to send LoGerfo any additional questions that the panelists could answer via email and/or follow-up conversations. He then summarized the conversation succinctly. Approaches taken by Broer and Hanushek share conceptual similarities--based on information taken from the NAEP student contextual questionnaire--but differ in methodological approach, i.e., constructing an additive index or an index through principal components analysis or using NAEP alone or in combination with other data. Kane's approach differed from those completely and shifted focus away from SES to income and away from continuing historical precedent to innovating a different approach for the future.

West then posed the first query to Ward, questioning why NCES assumes that the SES measure must behave the same across grade levels and rely only on existing questions. Does this lead to an assumption that questions about SES, such as household items, must be understood by both twelfth-graders and fourth-graders? Currently, researchers struggle to use NAEP data and express dissatisfaction with the SES measure. How much should what NAEP does now and in the future be constrained by retaining connections to the past?

Ward suggested that NCES should become more innovative in how questions are phrased so that all students can answer questionnaire items accurately. For example, NAEP is currently exploring how to ask fourth-graders about their family structures, which is often complicated and tricky to capture in a survey. Ward claimed it is incumbent upon NAEP to gather accurate information from students more effectively. However, West pointed out an intrinsic contradiction in Ward's reply; developing new questions nullifies the claim that variables now must be compatible with variables in the past.

West asked Kane if he ever compared the percentage of variance in achievement explained by income only with that explained by SES. Kane has not yet conducted this comparison but predicted that income would capture considerably more variance than

SES. Kane also clarified that he can impute income using school locations through Census data dating back to 1990, which would allow trend calculations. Only the student-level neighborhood income measure would be new.

Kane concluded his response by beseeching the Board to help sort out the muddled picture of SES and achievement. Non-researchers could grasp achievement patterns more easily if NAEP used an index. But, given the variety of SES indices, there is no consensus on the "right" index. Indeed, the "right" SES measure likely differs with the question being asked. Some research shows the gaps widening; others show the gaps narrowing. All of this research shows wide variation in results with SES, so using only income could clarify the issue.

Time expired, and Matthews concluded the session with an enthusiastic thanks to West and to the panelists. She warned that this conversation was intended to foment more robust and longer conversations in the future.