Key Education Indicators for NAEP: A Composite Indicator Approach

By Alan Ginsburg and Marshall S. Smith

A NAEP Data Analysis Report
Prepared for the National Assessment Governing Board

June 2014

Alan Ginsburg is an education consultant and analyst. He is former *Director of Policy and Program Evaluation services for the U.S. Department of Education. Marshall S. Smith is former U.S. Under Secretary of Education and former Dean of the Stanford University Graduate School of Education. The data analyses and interpretations in this report are those of the authors and do not necessarily represent the views of the National Assessment Governing Board.*

Key Education Indicators for NAEP:
A Composite Indicator Approach

Table of Contents

EXECUTIVE SUMMARY

KEY EDUCATION INDICATORS FOR NAEP:
A COMPOSITE INDICATORS APPROACH

This report recommends that the National Assessment of Educational Progress (NAEP) develop ten to 15 composite Key Education Indicators (KEIs) that would be regularly reported along with student achievement results. Such indicators would greatly enrich NAEP reporting by adding information on the complex factors that influence student achievement. They also would show how prevalent these conditions are in the various groups and states on which the assessment reports.

Because of their complexity, useful measures of important background conditions frequently require composites that are theoretically and empirically valid, rather than the individual contextual variables on which NAEP now reports. A KEI is best described as a weighted average of several different contextual variables. Preparing such indicators for a range of important topics would extend the idea of a composite for socio-economic status (SES), which has been proposed by an expert panel. The panel said an SES composite would be a much-improved alternative to using data on the percent of students qualifying for free or reduced-priced lunch as NAEP's prime indicator of poverty.

The National Assessment of Educational Progress is the only regularly and predictably administered cross-sectional data set where background information can be directly related to student achievement. It is the only data set where information is regularly gathered from students, teachers and principals in the same schools. These characteristics provide the opportunity for asking questions to help us better understand the reasons for the differences and changes in student achievement. The questions might also provide data to increase our understanding of the status and changes in the quality of school experiences and of the pre-school experiences that prepare young children for kindergarten.

At present NAEP's reporting of contextual variables is limited and appears ad hoc. While there are over 1,400 variables on the NAEP Data Explorer, over 1,000 of them were not administered in the most recent assessments. The only regular reporting is by racial/ethnic categories and eligibility for school-lunch. Almost all of the other background data collected are never formally analyzed nor reported in NAEP publications. Even though the structure of the Data Explorer is sensible, it does not establish priorities. Moreover, unlike the two major international surveys of TIMSS and PISA, each variable is presented only in isolation with no connections made among those addressing similar conditions. The lists in the Data Explorer are confusing and there is no clear rationale for the many changes in the variables collected.

Key education indicators are proposed as theoretically and empirically derived statistics that regularly measure important conditions likely to influence academic achievement. While there are many potential configurations for KEIs, we suggest that a coherent set of

indicators should be clustered in two categories, one focusing on the school, the other on the student.

The school quality component would have five basic school characteristic variables-- location, size, type, socio-economic class composition of the student body and racial composition, and six composite KEIs--teacher quality, teacher professionalism, school climate, quality of implementation of standards and curriculum, effective use of technology, and the use by the school of systematic improvement strategies (Exhibit ES-1).

Exhibit ES-1. Illustrative key education indicators (KEI) for school quality	
Composite Indicators	**Evidence-Based Indicator Components (illustrative)**
1. Teacher quality	• Student view of quality, teacher degree in field, experience, dispositions & mindset
2. Teacher professionalism	• Seek help to improve, support other teachers, seek growth year after year, enjoy work, engaged in professional networks
3. School climate for learning	• Student absenteeism (not excessive), school safety, teacher expectations for students, teacher support for each other, principal trusted, mindset
4. Quality of implementation of standards and curriculum	• Student-centered, aligned rigorous content, teach for understanding, adjust for student learning differences
5. School effectively uses technology to teach	• Access at school and home, use at school and home, effectiveness in technology adding learning value
6. Continuous improvement throughout	• Teachers use formative assessment, professional development focused on improving classroom and administrative processes

The student component represents the individual characteristics of the students. Along with the basic characteristics of sex, race, age, and handicapping conditions, the student KEIs seek to capture the fundamental characteristics of student learning inside and outside the school through six broad indicators--socio-economic status, home/ and neighborhood educational climate, preschool experiences, student engagement with learning, after-school learning opportunities, and non-cognitive contributors to academic achievement (such as self-control and persistence). (Exhibit ES-2).

Exhibit ES-2. Illustrative key education indicators (KEIs) for students	
Composite Indicators	**Evidence-Based Indicator Components (illustrative)**
1. Socio-economic status	• Composite indicator as recommended by NCES expert panel
2. Home and neighborhood educational climate	• Family support, place to study, parents talk with but not at the child, friends respect educational accomplishment
3. Preschool experiences	• Number of years in formal preschool, parent literacy activities with child, parent numeracy activities with child, parent sets boundaries
4. Student engagement with learning	• Student effort, hard work more important than luck, likes and goes to school, believes is learning a lot
5. After-school learning opportunities	• Formal after-school programs; informal after-school programs, parents take child to zoos, museums, etc.
6. Non-cognitive contributors to academic achievement	• Self-control • Persistence (grit or determination)

Illustrative KEI Composite Indicators

The paper illustrates in some detail the development of composite indicators in five of the above areas. Illustrative indicators are presented for three school KEIs—school climate, teacher quality, and education technology; and two student KEIs— socio-economic status (SES) and student engagement. The illustrations were chosen in part based on the capabilities of the NAEP Data Explorer.

Each illustrative indicator is based on theoretical and empirical research that supports its importance for student achievement. The SES KEI reflects the recommendation of the NAEP expert panel for a composite indicator. Development of the other four illustrative KEIs began with identifying an explicit framework of underlying causal variables. From this framework, the NAEP Data Explorer was examined to identify measured proxy variables. For the technology KEI, we concluded that existing NAEP data are insufficient to develop even an illustrative indicator. Instead, we suggest possible variables that could be developed into an indicator. For three of the other KEIs, only the most current data are utilized; for one proposed KEI trends over time are also presented.

As an example of indicator development, this report measures *school climate* as a three-variable KEI consisting of student attendance, school misbehavior, and teacher expectations. However, limitations of the NAEP Data Explorer prevent disaggregating results of the three-variable composite by student and school characteristics. Therefore, a two-variable composite indicator is presented to permit disaggregation. Exhibit ES-3 illustrates the results for grade 8 math of a composite indicator consisting of a two-variable combination of days absent and teacher expectations. The two-variable KEI was constructed because the Data Explorer can display a table of two composite variables along with student or school characteristics. The three-variable composite is at the Data Explorer maximum and the results cannot be disaggregated by school or

Exhibit ES-3. Composite index for average NAEP scores & percentages for math, grade 8, by race/ethnicity showing very positive and very negative teacher expectations for students and 0-2 days absent prior month, 2003

Race/ethnicity	Very positive teacher expectations & 0-2 days absent prior month		Some what positive or negative teacher expectations & 3 or more days absent prior month	
	Average NAEP score	Percentage	Average NAEP score	Percentage
National	**286**	**48**	**259**	**9**
White	293	52	273	8
Black	260	36	240	13
Hispanic	267	39	246	12
Asian/Pacific Islander	300	57	268	4
American Indian/ Alaska Native	270	38	253	15

Source: NAEP Data Explorer
Note: Data for "Two or more races" are not included in the tables.

student characteristics.

Exhibit ES-3 displays both the most positive and most negative two-variable combination for a school-climate indicator based on principal reports of teacher expectations for their students and student days absent during the prior month. The table shows NAEP scores and percentages cross-walked with student race/ethnicity.

The *very-positive school climate two-variable combination* consists of students with 0-2 days absent in the past month in schools with principals responding that their teachers mostly hold very positive expectations for student achievement. The year 2003 is used because that is the most recent year in which these background variables were collected.

- Nationally, 48 percent of grade 8 students were in this highly favorable school climate situation.
- By race/ethnicity, Whites and Asians were about 50 percent more likely to be in this highly favorable school climate than Blacks, Hispanics or American Indians.

The highly negative combination consists of students absent three or more days in the prior month and enrolled in schools with principals rating teachers as having only somewhat positive or negative expectations for students.

- Nationally, 9 percent of students were in a very unfavorable school climate situation.
- While only 8 percent of White and 4 percent of Asian-American students had both 3 or more days absent and were in schools with the least favorable teacher expectations, about 50 percent more Black (13%), Hispanic (13%), and American Indian (15%) students were attending schools with the most undesirable school climate.

Over time we hope that having higher percentages of minority students in the more favorable category would help to close achievement gaps.

The three-variable school climate composite indicator measures school climate as the combination of student attendance, school misbehavior, and teacher expectations. It identified 39 percent of all 2003 grade 8 students in a highly favorable school climate. This was a school where a student was absent 0-2 days, with no more than minor discipline problems and a grade-8 math teacher with very positive expectations for student achievement. Unfortunately, these contextual variables where not collected more recently than 2003 so we cannot examine changes in this indicator over time.

The report also illustrates the development of four other KEIs

- A *teacher quality composite* KEI with the NAEP variables of: (1) teachers' knowledge of academic content, (2) teachers' mindset or disposition, and (3) teacher experience

- A *technology composite* KEI as a combination of: (1) student and school access to computers, (2) computer use at school and home for instructional and learning purposes, and (3) effectiveness based on the belief of teachers and students that the technology adds value to learning beyond the impact of teachers and the student's peers. As a different approach to developing KEIs, each sub-indicator will be constructed of three or four questions (variables).

- A *student engagement* composite KEI for reading consisting of three variables: reading is a favorite activity, pages read in school and for homework, and student learns a lot when reading books.

- A *socio-economic status* (SES) KEI would be based on the NCES Expert Panel recommendations to construct an SES composite around three factors: family income and possessions, educational attainment of parents, and parental occupational status.

Recommendations to the National Assessment Governing Board

This report discusses the importance of adopting a consistent set of priority contextual variables for regular NAEP data collection and reporting. Many of these variables should

be components of Key Education Indicators, providing important composite data on factors affecting student achievement. Composite indicators are widely used in other fields, in education by international assessments, and by NAEP to develop achievement scales. They should now be applied to the NAEP contextual variables.

The report makes the following specific recommendations*:*

1. *Convene expert panels to develop frameworks for composite Key Education Indicators in several areas to be selected by the Governing Board. Each framework with accompanying specifications would provide the blueprint for preparing questions and methods of analysis and weighting. The process would be analogous to long-standing arrangements for preparing subject-matter frameworks and test item specifications for NAEP cognitive assessments. However, since each indicator framework would be more limited, the time and expense needed should be much less.*

   a. One of the KEIs should be an SES indicator based on the recommendations of the expert panel that reported to NCES. This indicator should be a composite of at least three factors--family income and possessions, parental educational attainment, and parental occupational status.

   b. Other indicators may be based on the illustrations in this report, as shown in the school and student groups in Exhibits ES-1 and ES-2. Consideration could be given to KEIs for specific assessment subjects and possibly for specific grades. Development should start with a few areas of greatest value and interest.

   c. Each KEI should be validated by research and theory. Before use in reports, each indicator must be tested in field studies along with the individual variables of which it is comprised.

2. *Identify questions previously used that could support developing trends over time for KEIs.*
   a. Consider re-using questions from old assessments, even if dropped more recently, to generate trends for variables likely to have a high priority in developing the KEIs. Examples include the questions on student, teacher, and principal perceptions incorporated in our illustrative KEIs that were last given in 2003. Repeating these questions would provide new information about trends that might help determine how best to create KEIs and effectively measure changes over time.

   b. *Report results for currently administered NAEP contextual variables with trends of ten years or more.* These trend analyses will provide useful information on school, teacher and student changes over at least a decade while offering a better understanding of important trend areas for indicator development.

3. *Consider other actions to support KEI development.*
   a. Conduct psychometric studies on building composite indicators. Conduct exploratory analyses to determine preferred strategies for computing indicator weights.

b. Examine possibilities for coordinating or linking with data from other federal data collections. An example is the SES indicator panel's recommendation to link NAEP measures with U.S. Census collections.

4. *Build a repository of articles and publications that use NAEP variables and indicators, which would be readily available to scholars and the public. A possible model for this repository is the NCES Early Childhood Longitudinal Study Data Products and Publications (2013).*

5. *Improve the NAEP Data Explorer to allow users to focus readily on the most useful and timely variables and dramatically reduce the number of variables routinely shown in searches.*
   a. Recent, useful variables should be placed in a prominent file; old, redundant, or useless variables in a secondary file.
   b. Enable the user to choose to see only those contextual variables available for selected years of interest.

Addendum on Long-Term Trend NAEP

Long-term trend NAEP provides important national mathematics and reading results at ages 9, 13 and 17 dating back to 1970. Although an in-depth examination of contextual variables and possible KEIs for the long-term NAEP assessment was beyond the scope of this review, we believe that the underlying rationale for developing KEIs is equally applicable to the long-term trend NAEP. Unfortunately, about half the contextual variables in long-term trend were eliminated in 2008 and 2012 without a clear rationale. Some of these should be restored to report on trends in important factors affecting academic achievement.

It is recommended that the Governing Board consider the following:

1. *Have the expert panels developing KEI frameworks and specifications for main NAEP also make recommendations for KEIs in the areas under consideration using contextual variables in the long-term trend assessments.*

2. *Restore useful questions that were eliminated in the 2008 and 2012 administrations of long-term NAEP by adding them to the next administration.*
.

Introduction

The National Assessment of Educational Progress (NAEP) measures and reports on student achievement in U.S. elementary and secondary schools. In mathematics and reading representative samples of students are tested every two years in grades 4 and 8 at the national, state and urban district levels and in grade 12 at the national level only. Every four years science is added. Other content areas--including writing, U.S. history, and civics-- are assessed on a non-regular basis, usually at least twice in each decade. Student performance data are analyzed and reported on and then posted on the NCES website. Full details are made available in a web-based product, the NAEP Data Explorer, which can also support re-analysis.

In each administration of NAEP, contextual information is collected from students, teachers and school principals to enrich the reporting of academic achievement. The contextual information spans a wide variety of student, teacher and school attributes. It is gathered through separate and independent multiple-choice questions. The questions for students are expected to fit into a 10 to 15 minute block of time. The questionnaire for teachers is expected to take no more than 20 minutes to complete, and for principals (or their designee) up to 30 minutes.

The contextual questions cover a wide range of topics, but apart from a core group used to categorize students (by age, ethnicity, gender, etc.), they often have been asked in only one or two collections, which removes the opportunity to track responses over time. On its face there seems to be little logic to the many changes that have been made. Indeed, since main NAEP began in 1990 there have been over 1,400 contextual questions asked in the administrations of mathematics and reading.   The great majority are no longer used. When NAEP presents its results few contextual variables are included in the widely disseminated public release. The only exception is school-lunch eligibility as a measure of poverty status but this has become increasingly flawed.

Moreover, the independence of the questions makes it difficult in the analyses to measure moderately to highly complex concepts that are theoretically and empirically related to the quality of education and that might be used to help explain levels, trends and differences among schools, districts and states in NAEP achievement data. An important example of such a concept is SES (socio-economic status). Last year an

expert panel, convened by the National Center for Education Statistics (NCES), proposed that a composite indicator of SES be prepared for analytic work and reporting. The panel said the SES indicator should combine the results of a set of independent variables gathered from students and principals. It should be based on theory and on empirical information from other sources. The panel said this single, powerful composite indicator of SES should be part of every administration of NAEP.

In this report we propose that NAGB extend the idea of indicators beyond SES to create about 10 to 15 broadly defined composite key education indicators (KEIs). Each KEI would be comprised of a set of independent variables that would combine to form the composite. Selection of individual variables that comprise a KEI would be determined by use of theoretical and empirical knowledge gained from other reliable sources outside the National Assessment. The SES indicator would be one of the KEIs.

Questions for various KEIs would be included in every administration of NAEP but topics should be rotated across different years to allow for many different topics to be covered. Also, within the time allotted, contextual questionnaires should continue to collect other important information, such as student effort on the assessment. The indicators would be used for the analyses carried out when NAEP results are released and should also be included in the NAEP Data Explorer for re-analyses.

The idea of indicators has been around for a long time. The Office of Management and Budget was creating and using them in the 1970s.[1] The National Science Foundation (NSF) is now working on indicators for STEM education and the National Research Council (2012) is creating indicators for a wide variety of sectors, including education. Those who create indicators for NAEP should take advantage of these efforts.

The recommendations in this report are not an effort to increase the data collection burden on students, teachers and principals. They also are not designed to replace or ignore the wide variety of other education data collected by the federal government. NCES, in particular, has a very useful set of publications every year that describe the status and trends of education in the United States.

---

[1] In the 1970's, Marshall Smith, a co-author of this report, commented on the OMB indicators for education while representing the then Office of Education.

The purpose here is to provide a set of theoretically constructed and organized key indicators of educational quality and equality in the United States. Unlike other available data, these indicators would be directly linked with the academic achievement of students at two or three grade levels and two or three content areas. Because they are composites they will likely be more reliable and valid than individual variables. Because they are theoretically and empirically derived they would provide knowledge and insight that might be generalized to other settings. Because they span several grades they promise to show changes in cohorts over time. Because they will include data from students, teachers, and principals in the same schools they would provide a much richer picture of the character of educational experiences in U.S. schools than can other data in which the linkages among actors are not available.

For all of these reasons, we believe that the composite indicators would substantially improve the quality and usefulness of the National Assessment. Over time we would expect the KEIs themselves to become ever more useful as our understanding of their validity improves and changes are made.

The report has six sections.

- Section I discusses the current contextual variables, their organizing structure and the lack of focus on a consistent set of variables within the structure.
- Section II explains how to move from the current contextual variables to composite indicators.
- Section III makes a short argument for indicators and then provides, as an example, a suggested structure that would contain eleven KEIs.
- Section IV presents four examples of school quality KEIs.
- Section V presents two examples of student KEIs.
- Section VI concludes the report with recommendations.

I. Current Contextual Variables

The NAEP Data Explorer provides access to all of the contextual variables that have been administered by NAEP over the past twenty years. The 8th

grade math assessment has a list of 1441 contextual variables. The vast majority of these were not administered in either of the past two assessments. Moreover there is only a small set of variables, most associated with the major student reporting categories, which have been given for each administration since 1990.

The result is a complex pattern with the underlying rationale not always apparent. Among the over 1,400 questions NAEP has asked, over 1,000 were no longer present in the most recent 2013 administration. If there is a systematic strategy for the pattern of questions included, NAGB should make it transparent so users of the Data Explorer may know what to expect and can plan their studies.

While the process for selecting contextual variables lacks clarity, the current structure for organizing them in the NAEP Data Explorer generally makes sense to us. (See Exhibit I-1). We recommend that a set of important contextual variables be carefully selected within each of the categories of the current structure in a systematic, evidence-based, and transparent way to be included in every NAEP administration. Others should be selected for use in every other administration. These

---

Exhibit I-1. Current NAEP Data Explorer structure for contextual variables

- Major Reporting Groups
    - Student factors
    - School Factors
    - Community Factors
- Student Factors
- Instructional Content and Practice
- Teacher Factors
- School Factors
- Community Factors
- Factors Beyond School
- Government Factors

---

predictable variables would be in addition to the KEIs. Although some new variables might be de developed it would be most desirable if some currently asked variables are deemed important enough for continued regular use.

We note that there are some contextual variables given in the past that might be repeated or be part of the standard set of contextual variables in one of the areas of the structure. For example, a contextual variable in 2002 had principals comment on the perceptions of teachers in their school about student ability. This turns out to be highly useful in developing a school climate KEI.

Overall with respect to the contextual variables, we have five suggestions:

1. Develop a transparent and evidence-based approach to using the contextual variables in the National Assessment.
2. Make sure past variables that measure important characteristics of schooling are carefully considered for use in new administrations.
3. Provide users of the Data Explorer with the option of selecting from a list of contextual variables from the current administration only, a list from past administrations, and a combined list. This would reduce the burden of having to search for variables that are currently used.
4. Pay careful attention to variables that may be altered by circumstances. The recent changes in the regulations for the allocation of free and reduced price lunch to all students in school-wide Title I schools reduces the accuracy of this measure as a proxy for school SES.   Attention should be paid to this.
5. Leave room in the contextual questionnaires for the components of between 10 and 15 key education indicators and the individual variables that comprise these composite indicators.

II. From Variables to Key Education Indicators (KEI)

What is a Key Education Indicator?

In the context of NAEP, Key Education Indicators (KEIs) are statistics that regularly measure important conditions of the education system and of students that are likely to influence academic achievement over time. A Key Education Indicator (KEI) for this report typically will consist of a composite set of variables that are theoretically and empirically related to each other For example, family income, educational attainment and occupational status are parts of a SES indicator. The contextual variables comprising KEIs should be asked regularly in NAEP assessments and may be viewed as part of a balanced scorecard approach that includes data on

the fairness and quality of education for all students as a complement to NAEP's primary role as an assessment of academic achievement.

Why do we propose KEIs for NAEP?

Although we believe KEIs would add greatly to the usefulness and impact of NAEP, several arguments have been made against this approach.

The first is that no more data is needed to describe the condition of education in the United States. After all, NCES releases an annual report named the "Condition of Education" with hundreds of data elements and a companion data digest with even more data.   Moreover, the National Science Foundation is creating a set of STEM indicators and the NRC is developing a select few indicators as part of a larger project that covers many sectors of American society.

A second argument is that long-term NAEP has successfully existed for 45 years and the main NAEP has been administered for over 20 years, without indicators or other composite variables; there is no need for change.

 It is certainly true that data is collected yearly and in longitudinal surveys on hundreds of aspects of American education. However, NAEP is different from other surveys in three important respects: (1) it links contextual variables to student achievement on a regular basis, giving us important information on how to interpret the levels and gains in achievement results. (2) NAEP gives correlated information about context from students, teachers and principals in the same schools, a characteristic that does not occur elsewhere on a regular basis.   (3) NAEP provides comparable, representative-sample data on a regular basis not only for the nation, but also for states and many large urban school districts.

Although the use of composite indicators by NAEP would be a change from past practice, this change has already begun and may play a crucial role in sustaining NAEP's leading position in educational testing. NCES is working on implementation of the expert study group proposal for an SES indicator. The major international assessment programs, PISA and TIMMS, make use of composite indicators. And indictors would add to the relevance and visibility of NAEP at a time when its role of providing

comparative data on student achievement is being challenged by the development of Common Core state exams.

This report proposes that NAEP develop a set of theoretically and empirically designed 10 to 15 key indicators, which would be regularly constructed and available for all analyses of NAEP achievement data. A coherent set of well constructed indicators collected regularly would give us a better understanding of the condition of education in classrooms and schools at the national, state and many local levels and for a wide range of different school environments. The predictability and reliability of the KEIs over time would provide a far stronger platform than we currently have to understand the levels of quality and inequality in our schools and classrooms.

The KEIs would be designed to measure contextual components that are critical to the success of schools and students. The starting point for developing KEIs is an underlying theoretical picture of the core educational factors that affect student learning at home and school. This school-home focus is consistent with the NAEP survey of students, teachers and principals. Thus, detailed descriptions of important education policy issues around state standards, assessments and governance are outside the scope of the NAEP survey and must come from other sources. However, school-level responses to these policies, such as teacher understandings of t Common Core standards or the inclusion of technology into classrooms are a reasonable part of NAEP data collections.

Exhibit II-1 describes a conceptual way to think about home and school factors in the form of a series of nested factors illustrated by concentric circles. The outer ring is learning that takes place in the home or after school. The remaining rings refer to various school-level factors. These include resources, the climate for learning, and classroom-level factors, such as technology and assessment use and changes in instruction and curriculum. The proposed NAEP composite indicators (KEIs) reflect this framework.

Exhibit II-1. Core Home and School Factors Affecting NAEP Student Outcomes
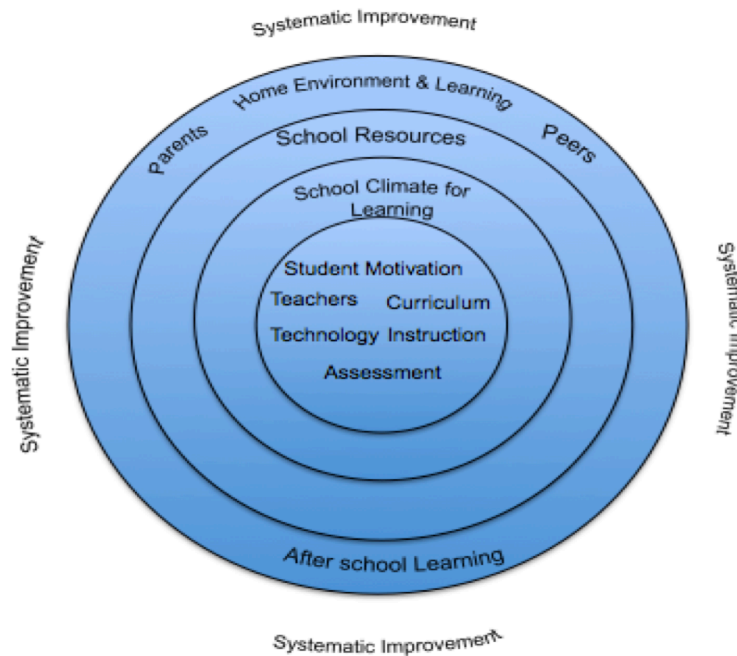
Composite Indicators

A KEI is a composite indicator that combines the results of variables (sub-indicators) that represent different aspects of a complex education phenomenon. Each KEI measures performance on a specific school or student condition critical for NAEP student achievement. Ordinarily these critical conditions are best described in terms of several different underlying variables -- thus a key education indicator will typically be framed as a composite of multiple variables.

A composite-indicators approach has strengths and weakness as outlined in Exhibit II-2. In our opinion, and in the opinion of the expert SES panel, the strengths are compelling. Further, it is proposed that the underlying components of any composite indicator be made public at the same time as the indicator itself to provide an understanding of the elements that comprise it.

Exhibit II-2. Strengths and Weaknesses of Composite Indicators	
Strengths	Weaknesses
• Summarizes complex conditions that are theoretically and empirically related to student achievement. • Makes explicit relationships among different survey questions • Enables a top-level view of a small but very important set of complex constructs. • Provides a composite measure that facilitates understanding & communication about levels and equity of performance on important aspects of education contexts.	• The selection of indicator weights can be arbitrary and lead to varying results. • Some say may lead to overly simplistic policy conclusions. • Focusing on the aggregate may miss serious problems in some dimensions of indicator

Numerous examples of current or proposed composite indices illustrate their use to capture a complex construct. Examples of composite indices outside of education are:

- The Standard and Poor's and Dow Jones stock indexes of large U.S. corporations.
- The United Nations Human Development Index (HDI) combining indicators of life expectancy, educational attainment and income into an index between zero and one.
- The Annie Casey Kids Count state rankings index of child well-being consisting broadly of four domains: (1) economic well-being, (2) education, (3) health and (4) family and community.

Moreover, the NAEP achievement scales for mathematics, reading, and other subjects are themselves an example of a composite index. For example, the NAEP mathematics achievement scale averages the results of 5 sub-scales for numbers, measurement, geometry, data analysis and probability and algebra with weights that vary by grade level. NAEP disaggregated scores for individual mathematics topics are available.

Also, the 2012 TIMSS international assessment has created composite scales from variables describing important educational contexts that affect student achievement. TIMSS indicators for students include early numeracy activities before primary school, home resources, and whether students like learning mathematics. The indicators for schools include

resource availability by subject, safe and orderly environments, and teacher career satisfaction.

The TIMSS approach in creating composite scales is to employ an IRT scaling procedure to develop numeric scales that represent a weighted response to individual items. Exhibit II-3 displays the items in the early numeracy activity scale comprised of parent responses to six questions about activities in the home before primary school. The responses to these

Exhibit II-3. TIMSS creates a composite scale from items about early numeracy activities before beginning primary school

Before your child began primary/elementary school, how often did you or someone else in your home do the following activities with him or her?

		Often	Sometimes	Never or almost never
ASBH02J	1) Say counting rhymes or sing counting songs			
ASBH02K	2) Play with number toys (e.g., blocks with numbers)			
ASBH02L	3) Count different things			
ASBH02M	4) Play games involving shapes (e.g., shape sorting toys, puzzles)			
ASBH02N	5) Play with building blocks or construction toys			
ASBH02O	6) Play board games or card games			

Often 10.3 Sometimes 6.9 Never or Almost Never

Source: Methods and Procedures in TIMSS and PIRLS 2011. Available July 2013 online: http://timssandpirls.bc.edu/methods/t-context-q-scales.html.

six items are pooled through an IRT statistical procedure to yield a scale with a mean across all countries of 10 and a standard deviation of 2. Cut-points were established on the scale to create three categories of doing early numeracy activities often, sometimes, and never or almost never.

This report is limited by the capability of the NAEP Data Explorer, which does not generate IRT scales.

III. A proposal for a Key Education Indicator Framework

This proposal is meant to be indicative rather than final. The underlying framework for developing the KEIs, shown by the nested rings in Exhibit II-1, leads to a suggested set of KEIs that are composites of variables on important education conditions affecting student achievement. The suggested KEI Framework has two groups of theoretically and evidence-based indicators. One group focuses on the school, the other on the student.

The school quality component would have five basic school characteristic variables (place, size, type, social-class composition and racial composition) and six key composite indicators (teacher quality, teacher professionalism, school climate, quality of implementation of standards and curriculum, quality of effective use of technology, and the use by the school of systematic continuous improvement strategies). All of these indicators are firmly based on evidence of their importance for academic achievement (Exhibit III-1).

| **Exhibit III-1. Illustrative key education indicators (KEIs) for school quality** ||
Composite Indicators	Evidence-Based Indicator Components (illustrative)
7. Teacher quality	• Student view of quality, teacher degree in field, experience, dispositions & mindset
8. Teacher professionalism	• Seeks help to improve, supports other teachers, seeks growth year after year, enjoys work, engaged in professional networks
9. School climate for learning	• Excessive student absenteeism, school safety, teacher expectations for students, teachers support each other, principal trusted, mindset
10. Quality of implementation of the standards and the curriculum	• Student centered, aligned rigorous content, teach for understanding, adjust for student learning differences
11. School effectively uses technology to teach	• Access at school and home, use at school and home, effectiveness in technology adding learning value
12. Continuous improvement throughout	• Teachers use formative assessment, professional development focused on improving classroom and admin processes

The student component would endeavor to capture the fundamental characteristics of student learning outside the school and student

perceptions about learning as it affects their experience in school. The component would have four basic student characteristic variables (race/ethnicity, gender, ELL status, and disability status) and five composite key indicators (SES, home and neighborhood educational climate, preschool experiences, student engaged with learning, and after-school educational opportunities). Each of the key indicators is based on extensive evidence and theory about its importance in the learning opportunities for students (Exhibit III-2).

Exhibit III-2. Illustrative key education indicators (KEI's) for students	
Composite Indicators	**Evidence-Based Indicator Components (illustrative)**
7. Socio-economic status	• Composite indicator as recommended by NCES expert panel
8. Home/neighborhood educational climate	• Family support, place to study, parents talk with but not at the child, friends respect educational accomplishment
9. Preschool experiences	• Number of years formal preschool, parent literacy activities with child, parent numeracy activities with child, parent sets boundaries
10. Student engaged with learning	• Student effort, hard work more important than luck, likes and goes to school, believes learning a lot
11. After-school learning opportunities	• Formal after school programs; informal after school programs, parents take child to zoos, museums, etc.

These two components and their indicators are only one way of thinking about how to construct the KEI. There are dozens of other reasonable approaches. We tried to adhere to a number of conditions: evidence based, theory based, parsimony, clarity, interest in indicators that would be valid over a reasonably long time period, and indicators that had variance and that measured constructs that could be improved. We would expect that the variables that were part of the indicators would also be available for analysts to look at separately as well as a variety of other variables selected by NAGB committees.

Our general recommendation here is that NAGB organize a small committee to settle on the structure of the KEIs and then create three or four other committees to construct the indicators that are proposed by the structure committee. This is similar to the approach proposed by the SES expert panel.

IV. School Quality: Examples of Key Education Indicators

Introduction

The following examples were developed for three of the six Key Education Indicators proposed on school quality: school climate for learning, teacher quality, and using technology effectively for teaching. Data for 2013 are available in the NAEP Data Explorer for the teacher quality KEI, but 2003 is the most recent year with data available on school climate. Unfortunately, there is insufficient useful data to develop a KEI for technology but we suggest the variables needed and a methodology to construct it.

Each example KEI consists of three sub-indicators, the maximum allowable in Data Explorer tables. For the teacher quality and school climate KEIs, the sub-indicators are described and data reported by race/ethnicity and the percentage of students qualifying for subsidized school lunch, an indicator of poverty. Then two and three-variable composite indicators are developed. The two-variable composite is also illustrated by student race/ethnicity and the percentage of a school's students on school lunch. The limits of the Data Explorer preclude such breakouts for the three-variable composite.

1. School Climate for Learning KEI

A white paper on *The School Climate Challenge,* jointly prepared by the Center for Social and Emotional Education and the Education Commission of the States, defines a positive school climate as a "safe and supportive school environment in which students have positive social relationships and are respected, engaged and feel competent." Perhaps the largest regular report on school climate is New York City's *School Environment Report (2013).* It assesses school climate based student attendance and on surveys of students, parents, and teachers that evaluate their school's academic expectations, communication, engagement, safety, and respect.

Exhibit IV-1. School climate for learning composite indicator
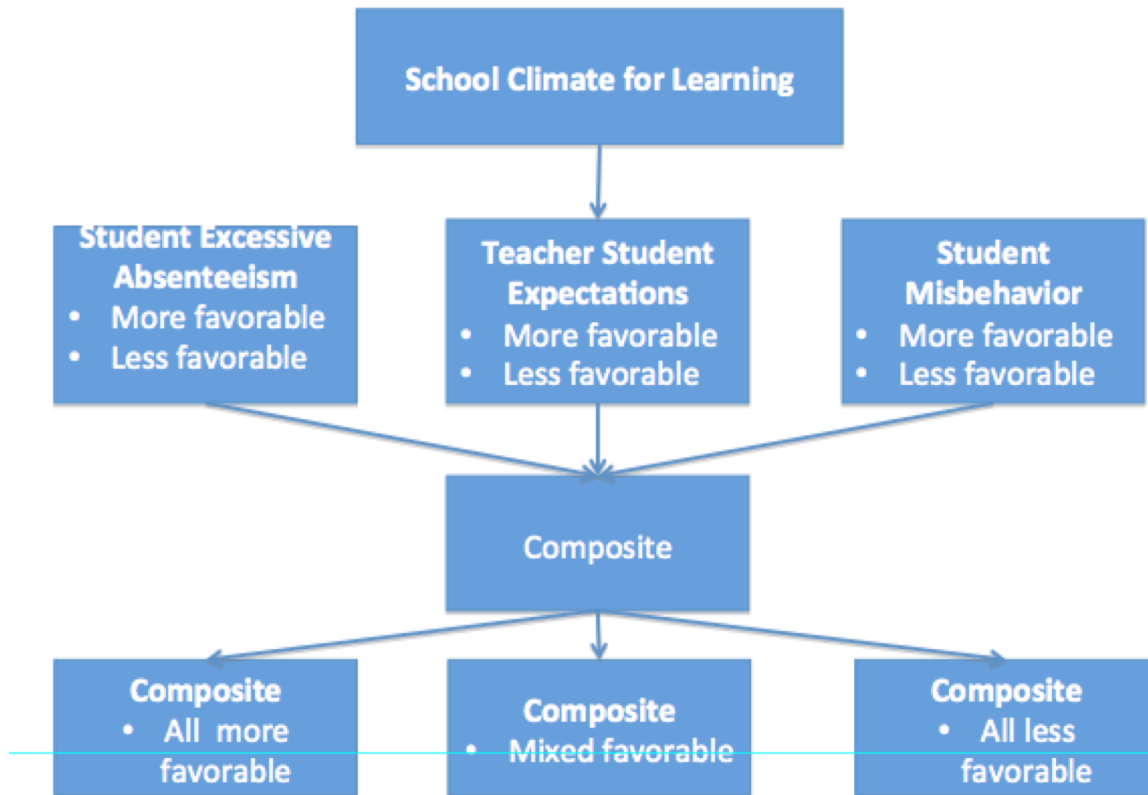
With the NAEP Data Explorer limited to a three-variable display, this report approximates the measurement of school climate as the three-variable composite of student attendance, school misbehavior and teacher expectations (Exhibit IV-1). Because 2003 is the latest year in which NAEP asked about teacher expectations for students, that year is chosen for the data for all three sub-indicators of the school climate for learning KEI.

Sub-indicator 1. Student Attendance

Schools that offer a student-friendly environment and monitor and respond to excessive student absenteeism encourage students to have good attendance. A solid body of research has identified harmful consequences associated with decreased school attendance (Gottfried, 2011). Students who are excessively absent receive less classroom instruction and their performance declines on exams in the same year (Chen & Stevenson, 1995; Nichols, 2003). Consistently low attendance over several years in the early grades is associated with later problems of non-promotion and dropping out (Neild & Balfanz, 2006).

NAEP reports average student attendance both by school (percent absent on an average day) and for individual students (by number of days absent during the prior month). We believe excessive absenteeism is more accurately reflected in individual data on student days absent the prior month than by the school-wide averages. A prior report to the Governing Board (Ginsburg & Chudowski, 2012) showed a sharp fall-off in achievement occurring between students reporting two or less days absent the prior month compared with three or more days absent. This break will be used for indicator construction to demarcate the category of excessive absenteeism (Exhibit IV-2).

Exhibit IV-2. Average NAEP scores and percentages for mathematics, grade 8 by days absent from school in the last month and percent in school eligible for national school lunch program: 2003

Percent eligible for National School Lunch Program	0-2 days		3 or more days	
	Average NAEP score	Percentage	Average NAEP score	Percentage
National	**281**	**78**	**267**	**22**
0-25%	293	80	282	20
26-50%	281	77	267	23
51-75%	269	77	256	23
76-100%	255	74	241	26

Source: NAEP Data Explorer

Exhibit IV-2 shows a consistent decline on NAEP grade 8 math scores for students between 0-2 days absent the prior month compared with those absent 3 or more days at both the national level and within each of the four school poverty categories. The declines range from 11 NAEP score points (roughly one grade level) for the lowest poverty schools (0-25%) to 14 points for schools with the highest proportion of low-income students.

Exhibit IV-3. Average NAEP scores and percentages for mathematics, grade 8, by students' days absent from school in the last month and race/ethnicity: 2003

Student race/ ethnicity	0-2 days		3 or more days	
	Average NAEP score	Percentage	Average NAEP score	Percentage
National	**281**	**78**	**267**	**22**
White	291	79	278	21
Black	256	76	242	24
Hispanic	263	75	249	25
Asian/Pacific Islander	294	89	271	11
American Indian/Alaska Native	267	67	255	33

Source: NAEP Data Explorer
Note: Data for "Two or more races" are not included in the tables.

Exhibit IV-3 displays the same information about excessive absenteeism for different racial/ethnic groups, showing a consistent fall-off in NAEP grade 8 math scores as days absent during the prior month rise from 2 or less to 3 or more. The score declines are similar across all racial/ethnic groups.

Sub-indicator 2. Teacher Expectations

Teacher expectations are described by how teachers gauge students in terms of their belief as to who will be successful in the classroom. While teachers need to adjust their teaching to challenge students at their individual levels, low-expectations for some students can become self-fulfilling prophecies. In a classic 1968 study, Pygmalion in the Classroom (Rosenthal and Jacobson), teachers were given incorrect information about students' IQ. The result was that students whose teachers expected them to perform better did in fact perform better, regardless of their actual IQ, and those expected to perform poorly achieved less well, regardless of actual IQ. The Education Commission of the States (2012) has cited

Exhibit IV-4. Average NAEP scores and percentages for mathematics, grade 8, by teachers' expectations for achievement and school percent of students eligible for school lunch: 2003 (school reported)

Percent eligible for National School Lunch Program	Very positive		Somewhat positive or negative	
	Average score	Percentage	Average score	Percentage
National	283	60	269	40
0-25%	293	75	284	25
26-50%	280	55	275	45
51-75%	269	47	263	53
76-100%	255	40	249	60

Source: NAEP data Explorer

similar associations between teacher expectations and the rate of improvement in student test scores in four studies published in academic journals since 2006 (Rubie-Davies, et.al, 2006; Tenenbaum & Ruck, 2007; McKown & Weinstein, 2008; van den Bergh, et.al., 2010).

In 2003, NAEP asked principals to respond to the following question about the expectations of teachers in their school:

> Question: How would you characterize each of the following within your school? Teachers' expectations for student achievement (school-reported)
> Responses: Very positive, Somewhat positive, Somewhat negative, Very negative

The advantage of asking school principals about teacher expectations instead of the teachers themselves is that the principals are more likely to give a valid response because the teachers may be reluctant to admit to low expectations for their students.

Exhibit IV-4 shows that nationally 40 percent of the students attended a school in which principals would characterize teachers as having less than very positive expectations about their students. The distribution varies considerably by the percentage of low-income students in a school. Among lowest-poverty schools, only 25 percent of students are in schools with teachers holding less than very positive expectations. By contrast, among the highest-poverty school group, 60 percent of students are in schools with teachers having only somewhat positive or negative expectations for their students.

Exhibit IV-5. Average NAEP scores and percentages for mathematics, grade 8 by teachers' expectations for achievement and student race / ethnicity: 2003 (school reported)

Student race/ ethnicity	Very positive		Somewhat positive or negative	
	Average score	Percentage	Average score	Percentage
National	**283**	**60**	**269**	**40**
White	291	66	282	34
Black	257	47	249	53
Hispanic	263	51	255	49
Asian/Pacific Islander	297	64	279	36
American Indian/Alaska Native	265	56	260	44

Source: NAEP data Explorer
Note: Data for "Two or more races" are not included in the tables.

Exhibit IV-5 shows a pattern of large differences in teacher expectations across different racial/ethnic groups. Only 34 percent of White and 36 percent of Asian /Pacific Islander students attend schools with teachers characterized as having less than very positive expectations. However, among Black students 53 percent are in schools with less positive expectations, among Hispanic students, 49 percent.

Sub-indicator 3. Student Misbehavior

A consistent body of research identifies a strong negative relationship between student misbehavior and student performance at both the individual student and school-wide level. An IES practice guide presented a research synthesis on *Reducing Behavior Problems in the Elementary School Classroom* (2008) estimated that "one-third of students fail to learn because of psycho-social problems" which lead to behaviors that interfere with learning.

Bryk (2010) reports on a 15-year longitudinal study of Chicago public schools that distinguished schools that improve from schools that fail to improve. This report concludes: "At a minimum, improving schools establish a safe and orderly environment — the most basic prerequisite for learning."

NAEP has at various times asked a range of questions about student behavior, including tardiness, cutting classes, drug and alcohol use, physical conflicts, and gang activity. For purposes of developing a

Exhibit IV-6. Average NAEP scores and percentages for mathematics, grade 8 by degree to which student misbehavior is a problem and percent in school eligible for national school lunch program: 2003

Percent in school eligible for National School Lunch Program	Not or minor problem		Moderate or serious	
	Average NAEP score	Percentage	Average NAEP score	Percentage
National	**282**	**72**	**266**	**28**
0-25%	292	87	282	13
26-50%	279	68	276	32
51-75%	268	60	263	40
76-100%	256	49	246	51

NAEP Data Explorer

composite index, a summary NAEP question is selected that captures a wide-range of misbehavior:

Question: To what degree is each of the following a problem in your

Responses: Not a problem, Minor, Moderate, Serious

Exhibit IV-6 shows nationally that 28 percent of the students attend schools where misbehavior in the classroom is considered a moderate or serious problem. The percentage directly varies with school poverty. Among

Exhibit IV-7. Average NAEP scores and percentages for mathematics, grade 8, by degree to which student misbehavior is a problem and race/ethnicity: 2003

Student race / ethnicity	Not or minor problem		Moderate or serious	
	Average NAEP score	Percentage	Average NAEP score	Percentage
National	**282**	**72**	**266**	**28**
White	289	79	281	21
Black	257	50	248	50
Hispanic	261	68	254	32
Asian/Pacific Islander	295	74	277	26
American Indian/Alaska Native	266	64	257	36

"Source: NAEP data Explorer
Note: Data for ""Two or more races"" are not included in the tables."

students in low-poverty schools, only 13 percent attend a school in which student misbehavior is considered a moderate or serious problem compared with 51 percent of students in high-poverty schools that have such problems. Among high-poverty schools, there is a 10-point differential, about one full-grade, in NAEP test scores between schools for

which student misbehavior is not or a minor problem and for those with moderate or serious levels of student misbehavior.

Exhibit IV-7 shows a similar wide disparity in the incidence of student misbehavior by race/ethnicity. Half of all grade 8 Black students are in schools with moderate or serious misbehavior problems--far higher than the proportion of Hispanic or American Indian students, and about double the rates for Whites and Asians.

Two and Three-Variable Composite indicators

Two and three-variable composite indicators illustrate combinations of the separate variables. Each composite is formed as a three-category combination of indicators with highly-favorable responses, highly-

Exhibit IV-8. Composite index for average NAEP scores & percentages for math, grade 8, by race/ethnicity showing very positive and very negative teacher expectations for students and 0-2 days absent prior month, 2003

Race/ethnicity	Very positive teacher expectations & 0-2 days absent prior month		Some what positive or negative teacher expectations & 3 or more days absent prior month	
	Average NAEP score	Percentage	Average NAEP score	Percentage
National	**286**	**48**	**259**	**9**
White	293	52	273	8
Black	260	36	240	13
Hispanic	267	39	246	12
Asian/Pacific Islander	300	57	268	4
American Indian/ Alaska Native	270	38	253	15

Source: NAEP Data Explorer
Note: Data for "Two or more races" are not included in the tables.

unfavorable responses, and all other.

Exhibit IV-8 illustrates a two-variable combination for grade-8 math of days absent and teacher expectations. The exhibit shows NAEP scores and percentages cross-walked with students' race/ethnicity for highly favorable and highly-unfavorable composite response categories.

The *highly-favorable combination* consists of students with 0-2 days absent in schools with principals responding that their teachers mostly hold very positive student expectations. In 2003,

- Nationally, 48 percent of grade 8 students were in this highly favorable school climate situation.
- By race/ethnicity, whites and Asians were about 50 percent more likely to be in this highly favorable school climate than Blacks, Hispanics or American Indians.

The highly-unfavorable response category consists of students with 3 or more days absent in schools with principals rating teachers as having only somewhat positive or negative expectations about students. While only 8 percent of White and 4 percent of Asian-American students had 3 or more days absent and were in schools with the least favorable teacher expectations, about 50 percent more Black (13%), Hispanic (12%) and American Indian (15%) were in these most undesirable school climate situations.

The three variable composite measures school climate as the combination of student attendance, school misbehavior and teacher expectations. It identified 39 percent of all 2003 grade 8 students in a highly favorable school climate. These students were in a school where an average student was absent 0-2 days, had no more than minor discipline problems and had a grade 8 math teacher with very positive expectations for students. The NAEP Data Explorer does not permit further disaggregation of the three-composite index by student race or school poverty classification, as contrasted with the two-variable composite (Exhibit IV-8).

2. Teacher Quality KEI

A considerable body of research suggests that the quality of teachers is a highly important factor influencing student achievement (Hanushek, 2005). Teacher quality can be assessed several ways. Improvements in student test-score gains directly measure one important aspect of teacher quality, but they fail to indicate or monitor the factors that make for effective teachers. NAEP survey questions offer a way to monitor teacher quality

attributes which research suggests are related to quality.

Three research-supported quality-related attributes of teachers may be approximated through the 2013 NAEP survey: (1) teachers' knowledge of academic and pedagogical content, (2) years of experience, and (3) mindset or disposition (Exhibit IV-9).

Exhibit IV-9. Building a Teacher Quality Composite NAEP Indicator

Teacher Quality

Teacher Education
• More favorable
• Less favorable

Teacher Experience
• More favorable
• Less favorable

Student Perception of Teacher Quality
More favorable
• Less favorable

Composite

Composite
• All more favorable

Composite
• Mixed favorable

Composite
• All less favorable
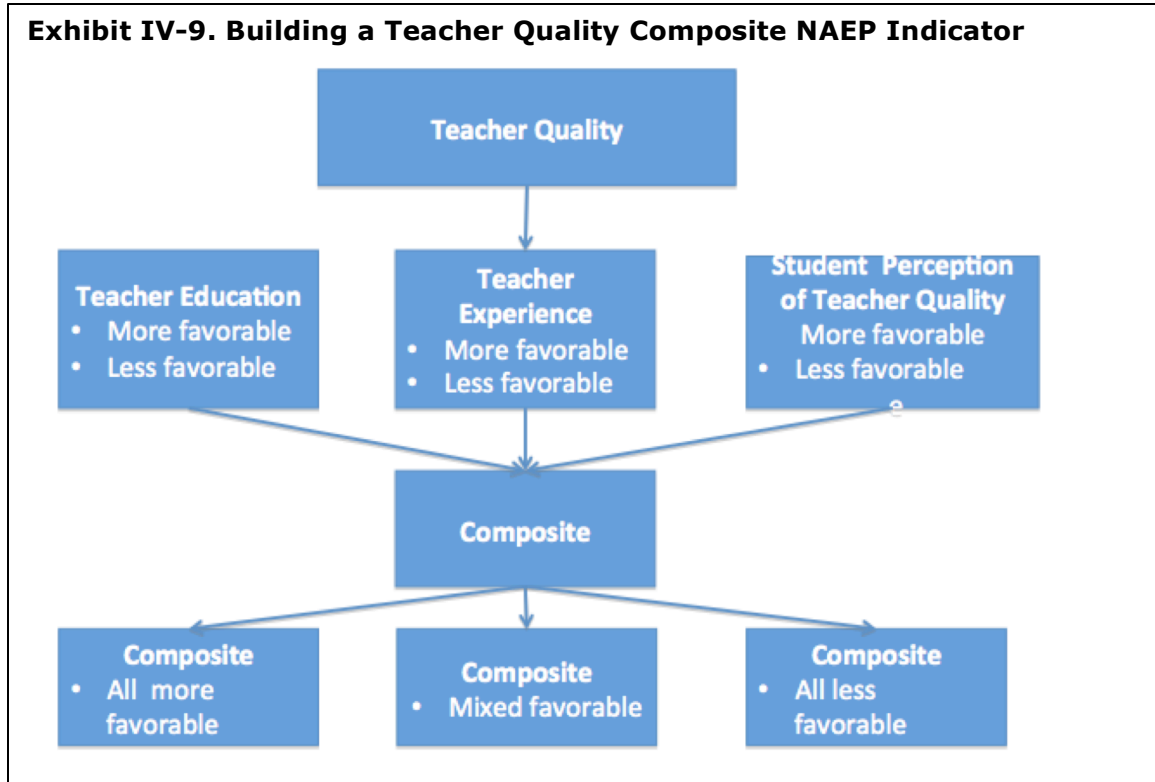
Other teacher-quality composites are possible. Some potential teacher quality factors collected by NAEP surveys, such as teacher attendance being a problem, were omitted, as the Data Explorer is limited to analysis of three factors. If the Data Explorer were enhanced, this KEI could be a composite of four or more variables. Also, NAEP surveys do not provide information on other teacher quality factors, such as ability to differentiate instruction, identify student mistakes, control the class, or make learning exciting. If an expert group were to design a teacher quality KEI their proposal might include these variables.

In the following discussion each variable selected as part of the indicator is explored in three ways: brief highlights are given of the supporting

2013 data are reported by school poverty and student race/ethnicity. Two and three-variable composite indicators are presented.

Sub-indicator 1. Teacher Knowledge

Teacher subject-matter knowledge and pedagogical content knowledge are prominently featured in both the Council of Chief State School Officers (CCSSO) model teacher standards (2013) and the Standards for Accreditation of Educator Preparation written by the Council for the Accreditation of Educator Preparation (2013).

The Council for Accreditation of Educator Preparation (CAEP), the body accrediting teacher preparation institutions, built its standards around research-based "areas of teacher preparation identified by the National Research Council 2010 report, *Preparing Teachers: Building Evidence for Sound Policy.* The NRC report concludes that research has identified two key elements in the capacity to teach (p.73):

- "Subject-matter expertise that encompasses a deep foundation of factual knowledge, understanding of how that knowledge fits in the conceptual framework of the field of study, and an internal organization of that knowledge that facilitates retrieval and application of his or her knowledge;" and

- "Pedagogical content knowledge in a given subject-matter field, that is, an understanding of how students' learning develops in that field, the kinds of misconceptions students may develop, and strategies for addressing students' evolving needs."

NAEP at times has asked different questions pertinent to teacher knowledge. In 2000, 68 percent of grade 8 students had a math teacher with a college course in calculus. Calculus may be sufficient preparation to teach grade 8 math, which suggests they may have mastered mathematics sufficiently through high school to effectively teach through tenth grade. NAEP also asked about teacher perceptions of how well prepared they are to teach different math content areas. Responses indicated that 61 percent of grade 8 students had teachers who felt well prepared to teach data analysis compared with 84 percent with teachers who felt well-prepared for algebra.

While these questions are not currently asked, NAEP 2013 did ask whether a teacher majored or minored in mathematics. As an overall

indicator of teacher knowledge this has the advantage of covering a range of mathematics coursework although the questions about particular mathematics coursework or preparation would yield useful additional information.

Exhibit IV-10. Average NAEP scores and percentages for mathematics, grade 8, by undergraduate major or minor in mathematics and percent eligible for National School Lunch Program: 2013

Percent eligible for National School Lunch Program	Major		Minor or special emphasis		No	
	Average scale score	Percentage	Average scale score	Percentage	Average scale score	Percentage
National	290	26	286	29	282	45
0-25%	304	29	302	29	299	43
26-50%	293	26	290	32	285	42
51-75%	283	23	278	30	275	47
76-100%	268	23	266	27	265	50

Source: NAEP Data Explorer

Exhibit IV-10 displays the percentages of grade 8 students in mathematics served by teachers with a major or minor in undergraduate mathematics. Note that NAEP also collects information on grade 8 teachers with a graduate math degree. However, a high percentage of these teachers would be expected to have an undergraduate math specialty and be counted as an undergraduate. An unduplicated count of undergraduate and graduate math majors is necessary, but is not given by the Data Explorer. It could be calculated through access to the raw data.

Overall in 2013, only 26 percent of grade 8 students had math teachers that majored as an undergraduate in math. Another 29 percent had teachers who minored in math, leaving 45 percent of the students with teachers lacking either a math major or minor. The group of schools with the lowest proportion of students in poverty had a somewhat lower proportion of teachers without any math specialty—43 percent compared with 50 percent for the highest-poverty school group. But the difference was not large, which was also the case for students across different racial/ethnic groups. The proportion of White students with a teacher who majored or minored in math was 5 percentage points higher than the proportion of Blacks and Hispanics. (Exhibit IV-11).

Exhibit IV-11. Average NAEP scores and percentages for mathematics, grade 8, by undergraduate major or minor in mathematics and race/ethnicity, school-reported: 2013

Race/ethnicity	Major		Minor or special emphasis		No	
	Average NAEP score	Percentage	Average NAEP score	Percentage	Average NAEP score	Percentage
National	290	26	286	29	282	45
White	299	27	294	29	292	43
Black	267	25	266	27	261	48
Hispanic	275	23	273	29	271	48
Asian/Pacific Islander	310	26	310	29	302	44
American Indian/Alaska Native	277	21	269	30	268	50

Source: NAEP Data Explorer
Note: Data for "Two or more races" are not included in the tables.

Sub-indicator 2. Teacher Experience

Teachers learn to be better teachers with experience. Several carefully designed studies (Hanushek, Kain, O'Brien and Rivkin, 2005; Cloftfelter, Ladd and Vigor, 2007) find a clear association between teacher experience and student achievement. While perhaps about half the gain from experience occurs during the first two years of teaching, studies have shown gains from teaching continue for up to 20 years

NAEP 2013 collects data on total years of experience in teaching and of experience in teaching a particular subject area. Total years of experience would be a useful indicator of experience for general classroom pedagogy, while years of experience teaching mathematics would be more relevant for a focus on the pedagogy for mathematics content. For developing an indicator of mathematics teacher quality, the NAEP 2013 question on experience in teaching mathematics is the most relevant:

> **Question:** Excluding student teaching, how many years have you taught mathematics in grades 6 through 12, counting this year?
> **Responses**: Less than 1 year, 1-2 years, 3-5 years, 6-10 years, 11-20 years, 21 or more years.

Exhibit IV-12. Average NAEP scores and percentages for mathematics, grade 8 by years taught mathematics in grades 6-12 and percent eligible for National School Lunch Program: 2013

Percent eligible for National School Lunch Program	5 years or less		6 years or more	
	Average scale score	Percentage	Average scale score	Percentage
National	**279**	**27**	**287**	**73**
0-25%	298	21	302	79
26-50%	283	24	290	76
51-75%	273	27	279	73
76-100%	262	36	268	64

Source: NAEP Data Explorer

Exhibit IV-12 displays grade 8 teachers of mathematics based on their experience in teaching math at grades 6-12 by 0-5 years or 5 or more years experience. The less experienced teachers are clearly more heavily concentrated in the highest-poverty schools, with 36 percent of students in these schools having teachers with 5 or less years of experience. This compares to just 21 percent of students with such inexperienced teachers in the lowest-poverty schools.

Less experienced teachers also tend disproportionately to teach Black and American Indian students compared to Whites or Asians (Exhibit IV-13). Only 21 percent of Asian students have a teacher with no more than 5 years of experience compared with 35 percent of Black students.

Exhibit IV-13. Average NAEP scores and percentages for mathematics, grade 8 by years taught mathematics in grades 6-12, students' race/ethnicity: 2013

Race/ethnicity	5 years or less		6 years or more	
	Average NAEP score	Percentage	Average NAEP score	Percentage
National	**279**	**27**	**287**	**73**
White	290	24	296	76
Black	260	35	266	65
Hispanic	268	30	274	70
Asian/Pacific Islander	300	21	308	79
American Indian/Alaska Native	266	33	272	67

Source: NAEP Data Explorer
Note: Data for "Two or more races" are not included in the tables.

Sub-indicator 3. Student perception of teacher quality

Research increasingly is supporting the predictive power of student responses in surveys on the quality of their teachers as indicators of teacher value-added or the impact of teachers on student achievement gains. Using a particularly rigorous experimental and longitudinal design,

Kane (2012) administered a student survey on the quality of teachers and compared the results with observations by trained teacher evaluators as predictors of student achievement-test growth.

> "The student responses were more correlated with teachers' student-achievement gains in math and ELA than the observation scores were. (Just as we did with classroom observations, to avoid generating a spurious correlation between student survey responses and achievement scores for the same group of students, we estimated the correlation across different classrooms of students taught by the same teacher.) In other words, student responses were not only consistent across classrooms, they were predictive of student achievement gains across classrooms." (Kane, 2012)

NAEP asks two questions on the views of students related to the quality of their teachers and classes in grade 8 mathematics:

> Question 1a. How often do you feel the following way in your math class? I have a clear understanding of what my teacher is asking me to do.
> Response: Never or hardly ever; Sometimes; Often; Always or almost always
>
> Question 1b. How often do you feel the following way in your math class? I am learning
> Response: Never or hardly ever; Sometimes; Often; Always or almost always

No similar questions are asked at grade 4 or for reading.

The limitations of the Data Explorer require selecting only one of the questions in forming a three-variable composite index. Exhibit IV-14 displays a two-way table showing student achievement for a cross-tab of the questions "clearly understand the teacher" and " I am learning."

Exhibit IV-14. Average NAEP scores and percentages for mathematics, grade 8 by all students clearly understand what teacher asks and feel I am learning: 2013

Average NAEP scores and percentages for mathematics, grade 8 by all students clearly understand what teacher asks and feel I am learning: 2013

Feel I am learning	Clearly understand what teacher asks							
	Never or hardly ever		Sometimes		Often		Always or almost always	
	Average NAEP score	Percentage	Average NAEP score	Percentage	Average NAEP score	Percentage	Average NAEP score	Percentage
Never or hardly ever	263	42	262	30	269	11	291	17
Sometimes	267	10	265	53	277	19	305	18
Often	263	1	270	32	283	40	304	26
Always or almost always	264	1	264	13	284	33	296	53

Source: NAEP Data Explorer

Looking across the rows to hold constant the response to "I feel I am learning," produces achievement score differences of over 30 points (an estimated three years on the NAEP scale at grade 8) between student responses "never or sometimes understand the teacher" and "always or almost always understand the teacher." By contrast, controlling for clearly understand what the teacher asks, in three of the four columns yields little change in NAEP scores across student responses on "I feel I am learning." Based on its greater association with achievement, the question "I clearly understand what the teacher asks" has been selected as the teacher quality sub-indicator.

Exhibit IV-15. Average NAEP scores and percentages for mathematics, grade 8 by clearly understand what teacher asks and percent of school eligible for National School Lunch Program: 2013

Percent eligible for National School Lunch Program	Never or sometimes		Often		Always or almost always	
	Average NAEP score	Percentage	Average NAEP score	Percentage	Average NAEP score	Percentage
National	**266**	**24**	**283**	**33**	**297**	**43**
0-25%	283	18	297	33	311	50
26 to 50%	272	22	287	33	299	45
51-75%	262	27	277	33	289	40
76-100%	252	32	267	31	277	36

Source: NAEP Data Explorer

Nationally, nearly one-in-four students (24 percent) say they have a grade 8-math teacher whom they never or only sometimes understand (Exhibit IV-15). Students in schools with the greatest percentage of low-income children are much more likely to have such teachers than students in schools with the lowest percentage of students in poverty— 32 percent compared to 18 percent.

Exposure to a teacher who the student never or only sometimes understands is much more common among Black, Hispanic and American Indian students than among Whites and Asians (Exhibit IV-16). For example, Black students are more than twice as likely as Asian/Pacific Islanders to have a teacher of math who they do not understand—33 percent compared to 15 percent.

Exhibit IV-16. Average NAEP scores and percentages for mathematics, grade 8 by clearly understand what teacher asks and race/ethnicity: 2013

Race/ethnicity	Never or sometimes		Often		Always or almost always	
	Average NAEP score	Percentage	Average NAEP score	Percentage	Average NAEP score	Percentage
National	**266**	**24**	**283**	**33**	**297**	**43**
White	276	21	292	33	304	46
Black	251	33	265	32	274	35
Hispanic	258	29	272	33	283	38
Asian/Pacific Islander	279	15	301	30	318	55
American Indian/Alaska Native	255	31	268	33	285	36

Source: NAEP Data Explorer
Note: Data for "Two or more races" are not included in the tables.

Two- and Three-Variable Composite Index

Two- and three-variable composite indices are illustrated for teacher quality. The two-variable composite is composed of teachers who have a major/minor in math and the extent of student understanding of their teacher. Exhibit IV-17 displays results by school poverty for students in the highest teacher-quality category (students have a teacher with a math major or minor and almost always understand their teacher) compared with the lowest teacher-quality situation (students have a teacher without a major or minor in math and the math teacher is never or only sometimes understood).

Exhibit IV-17. Two-variable composite indicator of teacher quality for mathematics, grade 8, by percent of school eligible for school lunch; 2013

Percent eligible for National School Lunch Program	Percent students who have a teacher with a major/minor in math and is always/almost always understood		Percent students who have a teacher with no major/minor in math and is never or only sometimes understood	
National	**300**	**24%**	264	11%
0-25%	312	29%	283	8%
26-50%	302	26%	269	10%
51-75%	292	22%	259	13%
76-100%	277	19%	251	16%

source: NAEP Data Explorer

Key findings include:

- Within each category of schools, based on the percent of low-income children, students with teachers in the highest teacher-quality category have much higher achievement than those with teachers who have no math degree and are rarely understood.
- Students in the lowest-poverty schools are 50 percent more likely to have the better-trained and more understandable teachers than students in the highest-poverty schools — 29 percent compared with 19 percent.
- Students in the highest-poverty schools are twice as likely to have teachers without a math degree and who are rarely understood than students in the lowest-poverty schools — 16 percent compared with 8 percent.

We believe a three-variable composite would be most valuable as the KEI for teacher quality. It would be comprised of 6 or more years of experience in teaching math; having a major or minor in math; and being always or nearly always understood by students. In 2013, about 19 percent of grade 8 students were exposed to teachers in this highest quality category. Unfortunately, the limitations of the Data Explorer do not allow us to present the distribution of such teachers by student race/ethnicity or school poverty but data are available for this to be done with a more powerful analytical tool.

3. Technology KEI

We can say with considerable certainty, based on current trends, that over the next decade the use of information technology for educational purposes will increase dramatically both in schools and homes. Even if its impact is not disruptive, the new technology will surely touch and possibly alter many of the ways we teach and learn. Part of the impetus for this transformation will come from a tremendous decrease in the cost of connections and hardware; another from a great increase in the availability of useful and powerful technology-based teaching tools.

An important stimulus will be the use of technology in creative ways to assess and analyze the progress of students both for formative purposes to assist in teaching and for summative purposes to provide accountability. For example, the two Common Core assessment consortia are creating assessments that would not be possible without the use of information technology. And the assessments, in turn, may well provide a powerful impetus for schools and teachers to use technology for learning.

It will be important for NAEP to develop an indicator of the extent to which these new technologies are available to all students and of the support they give for learning of both basic and complex content and strategies. The indicator would assess the degree of *access, use, and impact* of technology in the classroom. Also, because teachers will assign homework that may require the use of technology NAEP should measure whether students have appropriate access at home to the technology they need.

We suggest one indicator for the *effective* use of technology, constructed from three sub-indicators: (1) student and school access to technology, (2) use of technology for instruction at school and at home, and (3) effectiveness—whether or not teachers and students believe that the technology adds value to learning beyond the impact of teachers and peers.   Each sub-indicator should be constructed of three or four questions (variables) asked of students, teachers and schools. We describe in detail below one way of developing a KEI for technology, but recognize that a group of more knowledgeable experts would surely have other approaches

Sub-Indicator 1. Access

Access to technology has two parts: access in school to an appropriate level of the Internet and hardware plus appropriate access at home.

Variable One: Broadband access in the school.
One question for principals: Does your school have sufficient broadband access for all students to have access to the Internet? 1= No broadband access at all; 2= Weak and uncertain access; 3= Regular access only for administrative functions and teacher use; 4= Good access for most uses by administration, teachers and students, but sometimes weak. 5=First-rate access for all administrative and instructional uses.

Variable Two: Hardware access in the school.
One question for teachers: Does your classroom have the technology hardware capacity for you to teach and for students to use technology alone or in teams?   1 = No technology at all; 2= Technology only for teacher to use for administrative functions and for presentations and demonstrations to class; 3= Appropriate technology available for teacher and for groups of students; 4 = Appropriate technology for all students to use in the classroom.

Variable Three: Technology access at home.
One question for students: Do you have access to the Internet and other technology at home that you need to do your schoolwork?   1= No; 2= Yes.

A composite variable (sub-indicator) might be created from these three questions. Of course, the individual questions would also be available to investigators.

One way of creating the composite (Sub-indicator 1) might be to divide responses into three groups: Group 1= Low Access (1 or 2 on variable one; 1 or 2 on variable two; 1 on variable three); Group 2= Medium Access (All students not in group one or three); Group 3= High Access (5 on variable one; 4 on variable two; 2 on variable three).

The result would be a sub-indicator for access to technology, which, could be tracked by itself as well as contributing to the overall KEI for educational technology.

Sub-indicator 2: Use of technology in classrooms by teachers and student.

We define use as having three components: purpose, control by teacher, and time spent using technology:

Variable One: Purpose for which technology is used.
One question for teachers: For what purposes do you use technology in your classroom? 1= For administrative purposes only; 2= For administrative work and teacher presentation only; 3= For administrative work, teacher presentations, and for students to work alone or in groups.

Variable Two: Level of teacher control in classroom use of technology.
One question for teachers: In your classroom do you maintain full control over the use of computers?   1= Yes, I almost always structure my lessons so that each student knows what to do; 2= Sometimes I give students the opportunity to work together and to explore; 3= I often give students the opportunity to work together or separately and to explore.

Variable three: Frequency of use.
One question for teachers: How often is technology involved in classroom instruction and learning? 1= Never; 2= Only occasionally when necessary; 3= Often; 4= Almost always.

Composite variable: Sub-indicator 2. Group 1= Low Use (1 or 2 on variable one; 1 on variable two; 1 or 2 on variable three): Group 2 Moderate Use (all other); Group 3 High Use: (3 on variable one, 3 on variable two, 4 on variable three).

Sub-indicator 3: Quality and effectiveness of technology use

This sub-indicator focuses on the question of whether teachers see the technology as an integral part of the instructional system and whether students and teachers think it adds value to learning.

Variable One: Teacher assessment of the quality and effectiveness of technology use.
One question for teachers: How important is the use of technology to the quality and depth of learning in your classroom? 1= Little importance; 2= Clearly supportive of my teaching; 3= Adds value to the learning that I

otherwise could not provide.

Variable Two: Student perception of how important technology is to her/his learning.
One question for students: Does the use of the technology help you learn?
1= No or very little; 2= Sometimes; 3= Yes, it really helps a lot.

Composite variable: Sub-indicator 3. Group 1= Technology not helpful for improving learning (1 on both variables one and two); Group 2= Somewhat helpful (all other); Group 3= Very helpful (3 on both variables one and two)

Overall KEI for Technology Use

To create the overall technology KEI we combine the three sub-indicators, using the same methodology that was used in creating each of them.

Thus, if we add together the three sub-indicators we have scores ranging from 3 to 9. A score of 3, for example, would result from a one on each sub-indicator. A score of 5 might result from several combinations: either one point on one sub-indicator and two points on both of the others OR one point on two of the sub-indicators and three points on one. There are seven possible groupings—for 3, 4, 5, 6, 7, 8, or 9 points.

A reasonably parsimonious way to combine these and thus aid public understanding would be to create a pooled group for scores of 3 and 4; a second pooled group for scores of 5, 6, and 7; and a third pooled group for scores of 8 and 9. We would number the three pooled groups as 1,2,and 3 with 1 being little effective use, 2 being some effective use, and 3 being very effective use. Thus, the composite Key Education Indicator (KEI) for technology use would have three values, 1, 2 and 3 ranging from little effective use to very effective use.

To actually create such an indicator would require a great deal of work, first on the individual items, and then on their relationships to each other and to NAEP assessment scores.

V. Illustrative Student Key Education Indicators

Two illustrative KEIs are explored. The first is a summary of the recommendations of the expert panel for a composite SES indicator. The second example illustrates a KEI for student engagement.

1. KEI for Socio-economic Status (SES)

We refer the reader to the first-rate expert panel report (Cowan, et.al., 2013) on preparing a NAEP indicator of socio-economic status (SES). The panel was convened by the National Center for Education Statistics (NCES) in response to a policy statement by NAGB.

Currently, NAEP uses as its prime indicator of poverty status whether a student is eligible for free or reduced-price lunch. For several decades this was determined for each child, based on family income as reported by parents each year. Increasingly, however, school-lunch eligibility is becoming less valid because of changes in federal law that permit whole-school or even whole-district eligibility in places where a substantial majority of students are eligible for the subsidized lunch program. While NAEP also collects other variables related to SES, such as parental educational attainment and reading materials in the home, these also are reported only as individual variables and not combined into a composite SES measure. The panel recommended use of a composite measure of SES and gave general guidance on how to create one.

The nature of the panel itself, as well as its recommendations, offers important elements to consider in constructing any KEI:

- *Create a panel of experts to develop the KEI.* This expert panel would make independent recommendations, based on the evidence, to NCES and NAGB.
- *Use composite measures.* As the Cowan panel wrote, "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."

- *Construct two composite options: a core and expanded SES*

measure.

- o The panel said the core measure would focus on "family conditions and consisted of family income and other indicators of home possessions and resources, parental educational attainment, and parental occupational status. …This should be the subject of immediate focus for NAEP reporting."
- o " Neighborhood and school SES could be used to construct an expanded SES measure, and measures of these variables could contribute to an expanded SES."

- *Consider linking NAEP measures with another data source, in this case data from the U.S. Census Bureau.* The expert panel gave as a rationale for such linking: "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."

Implementation of the expert panel recommendations is dependent upon follow-up by NCES and NAGB; significant steps are already underway

2. Student Engagement KEI: Reading

While mathematics learning is primarily dependent on school instruction, student achievement in reading is also strongly influenced by student engagement with reading material and oral language outside of school as well as in the classroom. A composite indicator of student engagement in reading would focus on student reading habits and perceptions, as measured by three variables:

- Reading is a favorite activity;
- Pages read in school and for homework;
- Learn a lot when reading books

The creation of this proposed KEI differs from the prior illustrative KEIs in two important ways. First, a regression analysis is used to assess whether these three factors empirically make an independent contribution to student achievement, thereby warranting inclusion in a composite index.

Second the data for these three sub-indicators are available for 2002 and 2013, which permits analyses of trends over a decade.

Regression analysis to estimate independent contributions of student-engagement sub-indicators to student outcomes

In addition to basing variable selection on research supporting the importance of a factor in student achievement, multiple regression analyses of several factors can estimate whether each makes an independent contribution. We acknowledge that multiple regression analysis with NAEP one-year cross-sectional data is not as strong methodologically as with longitudinal data. One weakness is that student achievement is cumulative and NAEP only measures contextual variables at one point in time. However, such analysis may be more appropriate with student characteristics, such as engagement, which should be stable over a number of years, than with the variables of teacher characteristics, which can change substantially each year.

Exhibit V-1. Multiple regression of the three composite sub-indicators for students' reading engagement on students' grade 8 reading scores, 2002

	Regression Coefficients	T Statistic
Multiple Correlation	0.30	
R Squared	0.09	

	Regression Coefficients	T Statistic
Intercept	239.02	229.45
Learn a lot when reading books		
Disagree	11.74	12.70
Agree	15.96	16.84
Strongly agree	16.02	11.49
Reading is a favorite activity		
Disagree	2.52	4.57
Agree	11.07	16.96
Strongly agree	22.67	23.40
Pages read in school and for homework		
6-10	5.21	9.51
11-15	8.57	10.77
16-20	10.61	11.93
More than 20	7.80	13.65

Exhibit V-1 displays the regression results. The coefficients within each variable group measure the effect on student reading scores relative to an omitted response. The omitted response is strongly agree for the first two sub-indicators (learn a lot when reading books and reading is a favorite activity) and 5 or less pages read for the last listed sub-indicator. The coefficients for the responses within each sub-indicator are statistically significant supporting the independent contribution of each variable to student reading achievement. All 12 variable coefficients are in the expected positive direction and eleven of the twelve are successively increasing. The only exception is the coefficient for "more than 20 pages read" which, while positive, is of an unexpected lower value than for the preceding two variables. This may reflect response error or teachers giving lower-achieving students more reading material.

Because the NAEP series on school-poverty, based on eligibility for subsidized lunch, only goes back 2005, the results are shown disaggregated by student race/ethnicity.

Sub-indicator 1. Reading is a favorite activity

Student engagement in reading is strongly related to reading achievement (Guthrie and Wigfield, 2000). Engaged readers are more motivated to read. They are also more likely to read strategically and use multiple approaches to comprehend reading material.

The response to the question "reading is a favorite activity" is the NAEP contextual variable that approximates student engagement in reading.[2] The results for 2002 and 2013 (Exhibit V-2) show:

- There is a consistent positive relationship between the degree of agreement that reading is a favorite activity and student achievement scores. In 2013 the achievement difference between "strongly agree" and "strongly disagree" was 33 points—equivalent to about three grades on NAEP.

Exhibit V–2. Average NAEP scores and percentages for reading, grade 8 by reading is a favorite activity: 2013 and 2002

Year	Strongly disagree		Disagree		Agree		Strongly agree	
	Average NAEP score	Percentage	Average NAEP score	Percentage	Average NAEP score	Percentage	Average NAEP score	Percentage
2013	256	24	264	40	276	24	289	13
2002	255	25	262	41	271	23	283	11

Source: NAEP Data Explorer

- Quite a high percentage, nearly two-thirds of grade 8 students in 2013, either disagree or strongly disagree that reading is a favorite activity.
-
- The proportion of students in each category of agreement or disagreement with the statement did not change markedly in the decade of 2002 to 2013.

All racial/ethnic groups display similar percentages of how much students agree or disagree that reading is a favorite activity except for Asian/Pacific

---

[2] NAEP also currently asks questions about "read for fun on your own time" and "talk with friends about what you read," which could be used to approximate engagement. These were not examined because of the three-variable limit in Data Explorer tables.

Islanders students, who are 14 percentage points less likely than Whites to disagree or strongly disagree that reading is a favorite activity (Exhibit V-3). The changes between 2002 and 2013 indicate that:

- Within each racial/ethnic group, the largest improvement in NAEP

Exhibit V-3. Average NAEP scores and percentages for reading, grade 8 by reading is a favorite activity and race/ethnicity: 2013 and 2002

Race/ethnicity	Year	Strongly disagree		Disagree		Agree		Strongly agree	
		Average NAEP score	Percentage	Average NAEP score	Percentage	Average NAEP score	Percentage	Average NAEP score	Percentage
White	2013	262	25	273	38	286	23	295	14
	2002	261	27	271	40	282	22	291	11
Black	2013	244	23	247	41	257	25	268	11
	2002	240	21	244	42	250	27	260	10
Hispanic	2013	247	22	252	45	263	24	277	9
	2002	241	24	246	47	252	22	262	7
Asian/Pacific Islander	2013	264	13	273	36	284	34	303	18
	2002	261	15	262	39	270	33	283	13
American Indian/Alaska Native	2013	243	21	249	43	258	26	269	10
	2002	247	22	245	44	257	22	274	11

Source: NAEP Data Explorer
Note: Data for "Two or more races" are not included in the tables.

grade 8 reading scores between 2002 and 2013 occurred among students who agree or strongly agree that reading is a favorite activity. For example, among students who strongly agree, Blacks gained 8 points and Hispanics 15 points compared with only 4- and 6-point gains, respectively, for students who strongly disagree with the statement.

Sub-indicator 2. Pages read in school and for homework

School-related reading is different and complements home reading. School-related reading requires analyzing and evaluating what is read, but may lack the enjoyment of reading for pleasure. Research suggests both types of reading are beneficial. Students need "opportunities to practice reading for various purposes … lots of exposure to different kinds of reading materials" (Snow, 2001). The NAEP regression results in Exhibit V-1 are consistent with an independent contribution of reading in school and for homework to reading achievement.

The distribution of the typical number of pages read in school and for homework is little changed between 2002 and 2013.

Sub-indicator 3. Learn a lot when reading books

Student perceptions of whether they are learning are reasonably accurate barometers of whether they are learning (Kane, 2012). Moreover, current

Exhibit V-4. Average NAEP scores and percentages for reading, grade 8, by learn a lot when reading books: 2013 and 2002

Year	Strongly disagree		Disagree		Agree		Strongly agree	
	Average NAEP score	Percentage	Average NAEP score	Percentage	Average NAEP score	Percentage	Average NAEP score	Percentage
2013	246	3	259	18	268	64	281	14
2002	243	4	258	19	266	64	274	13

Source: NAEP Data Explorer

perceptions of own ability influence future behavior and hence future learning itself (Alexander, Entwisle and Horsey, 1997; Rhodes, 2007). OECD (2010) concludes, "Attitudes towards reading and learning, motivation, engagement in reading activities and reading proficiency are mutually reinforcing."[3]

NAEP correlational results reaffirm the research showing that student perceptions of whether they learn a lot when reading books is a strong correlate of NAEP reading scores (Exhibit V-4).

- The range in scores between strongly agree and strongly disagree with learning a lot when reading books is 35 points or about three and half years on the NAEP reading scale between grades 4 and 8.
-
- Interestingly, the strongly-agree group experienced by far the largest improvement in scores between 2002 and 2013.

- However, there was no significant change in the proportion of students in any of the agreement categories between these years.

The results by race/ethnicity (Exhibit V-5) show:

- For each racial/ethnic group and for 2002 and 2013 students increasing agreement that they learn a lot when reading books is associated with an increase in NAEP scores.

---

[3] OECD (2010). *Pisa results: learning to learn-vol III, p. 27.*

- The percentages of students were largely unchanged in the different levels of agreement that they learn a lot when reading books.

Exhibit V-5. Average NAEP scores and percentages for reading, grade 8, by learn a lot when reading books and race/ethnicity: 2013 and 2002

Year	Race/ethnicity	Strongly disagree		Disagree		Agree		Strongly agree	
		Average NAEP score	Percentage	Average NAEP score	Percentage	Average NAEP score	Percentage	Average NAEP score	Percentage
2013	White	252	4	266	20	278	63	292	13
2002	White	248	5	263	21	275	63	285	12
2013	Black	236	3	246	14	250	66	258	18
2002	Black	227	3	242	13	246	66	251	17
2013	Hispanic	236	3	248	19	257	66	270	11
2002	Hispanic	232	4	239	20	249	65	257	11
2013	Asian	263	1	269	11	280	66	291	21
2002	Asian	242	2	266	14	266	63	273	21
2013	American Indian	228	3	242	18	254	68	265	11
2002	American Indian	220	5	243	23	254	60	264	13

Source: NAEP Data Explorer
Note: Data for "Two or more races" are not included in the tables.

Two and Three-Variable Composite Indicator

A two-variable composite indicator for student engagement in reading has been created (in Exhibit V-6) by the combination of student responses to the two questions: learn a lot when reading books and reading is a favorite activity. Exhibit V-6 compares the results for the least student engagement (most negative responses of strongly disagree or disagree on both questions) with the most positive responses (strongly agree or agree on both questions). The comparisons are displayed nationally and by racial/ethnic group, with the following results:

- In every comparison, the NAEP grade 8 reading scores are considerably higher for students with the most positive responses compared with the most negative responses within the same student group. Nationally, the difference of 25 NAEP points on grade 8 reading between the strongly agree/agree and the strongly disagree/disagree is equivalent to about two and half years on the NAEP reading scale.

Exhibit V-6. Average NAEP scores and percentages for a two-variable composite indicator for reading, grade 8, by learn a lot when reading books and reading is a favorite activity: race/ethnicity, 2013 and 2002

	Student responses: strongly disagree/disagree		Student responses: stongly agree/Agree	
	Average NAEP score	Percentage	Average NAEP score	Percentage
2013				
National	**256**	**19**	**281**	**34**
White	262	21	290	34
Black	243	14	261	33
Hispanic	246	20	268	30
Asian/Pacific Islander	267	10	292	49
American Indian/Alaska Native	239	19	262	33
2002				
National	**255**	**21**	**277**	**31**
White	260	23	287	30
Black	240	14	254	34
Hispanic	238	21	257	27
Asian/Pacific Islander	262	13	274	43
American Indian/Alaska Native	238	25	264	31

Source: NAEP Data Explorer
Note: Data for "Two or more races" are not included in the tables.

- The distribution of responses between strongly disagree/disagree and strongly agree/agree categories did not change much over the 11-year period with the exception of a relatively large increase in the Asian/Pacific Islander percentage of strongly positive responses.

- Between 2002 and 2013 NAEP scores increased somewhat more for students in the most positive response category compared with the most negative, especially for Black, Hispanic and Asian students.

A three-variable composite indicator for student reading habits and perceptions is also computed for the two polar cases of most negative and most positive students responses (Exhibit V-7). The results are similar to the two-variable composite in showing:

- Consistently higher NAEP grade 8 reading scores for students with strongly agree/agree responses compared with students who responded strongly disagree/disagree.
-
- Little change between 20012 and 2013 in the percentage of student responses in either the most positive or the most negative categories.

Exhibit V-7. Average national NAEP scores and percentages for a three-variable composite indicator for reading, grade 8, by learn a lot when reading books, reading is a favorite activity, and pages read in school and for homework: 2013 and 2002

Year	Student Responses: Least favorable toward reading*		Student Responses: Most favorable toward reading**	
	Average NAEP score	Percentage	Average NAEP score	Percentage
2013	253	12	285	19
2002	253	13	279	7

Source: NAEP Data Explorer
*Students disagree/strongly disagree that reading is a favorite activity and they learn when reading books and they typically read 10 or less pages in school or for homework.
** Students agree/strongly agree that reading is a favorite activity and they learn when reading books and they typically read 11 or pages in school or for homework.

VI. Recommendations to NAGB

This report discusses the importance of adopting a consistent set of priority contextual variables for regular NAEP data collection and reporting. Many of these variables should be components of Key Education Indicators, providing important composite data on factors affecting student achievement. Composite indicators are widely used in other fields, in education by international assessments, and by NAEP to develop achievement scales. They should now be applied to the NAEP contextual variables.

The report makes the following specific recommendations*:*

2. *Convene expert panels to develop frameworks for composite Key Education Indicators in several areas to be selected by the Governing Board. Each framework with accompanying specifications would provide the blueprint for preparing questions and methods of analysis and weighting. The process would be analogous to long-standing arrangements for preparing subject-matter frameworks and test item specifications for NAEP cognitive assessments. However, since each indicator framework would be more limited, the time and expense needed should be much less.*

   d. One of the KEIs should be an SES indicator based on the recommendations of the expert panel that reported to NCES. This indicator should be a composite of at least three factors—family

income and possessions, parental educational attainment, and parental occupational status.

e. Other indicators may be based on the illustrations in this report, as shown in the school and student groups in Exhibits III-1 and III-2. Consideration could be given to KEIs for specific assessment subjects and possibly grades. Development should start with a few areas of greatest value and interest.

f. Each KEI should be validated by research and theory. Before using in reports, each indicator must be tested in field studies along with the individual variables of which it is comprised.

4. *Identify questions previously used that could support developing trends over time for KEIs.*
   a. Consider reusing questions from old assessments, even if dropped more recently, to generate trends for variables likely to have a high priority in developing the KEIs. Examples include the questions on student, teacher, and principal perceptions incorporated in our illustrative KEIs that were last given in 2003. Repeating these questions would provide new information about trends that might help determine how best to create KEIs and effectively measure KEI changes over time.

   b. *Report results for currently administered NAEP contextual variables with trends of ten years or more.* The trend analyses will provide useful information on school, teacher and student changes over at least a decade while offering a better understanding of important areas for indicator development.

5. *Consider other actions to support KEI development.*
   c. Conduct psychometric studies on building composite indicators. Conduct exploratory analyses to determine preferred strategies for computing indicator weights.

   d. Examine possibilities for coordinating or linking with data from other federal data collections. An example is the SES expert panel recommendation for linking NAEP measures with U.S. Census collections.

5. *Build a repository of articles and publications that use NAEP variables and indicators, which would be readily available to scholars and the public. A possible model for this repository is the NCES Early Childhood Longitudinal Study Data Products and Publications (2013).*

6. *Improve the NAEP Data Explorer to allow users to focus readily on the most useful and timely variables and dramatically reduce the number routinely shown in searches.*
    c. Recent, useful variables should be placed in a prominent file; old, redundant, or useless variables in a secondary file.
    d. Enable the user to choose to see only those contextual variables available for selected years of interest.

Addendum on Long-Term NAEP

Long-term trend NAEP provides important national mathematics and reading results at ages 9, 13 and 17 dating back to 1970. Although an in-depth examination of contextual variables and possible KEIs for the long-term NAEP assessment was beyond the scope of this review, we believe that the underlying rationale for developing KEIs is equally applicable to long-term trend NAEP. Unfortunately, about half the contextual variables in long-term trend were eliminated in 2008 and 2012 without a clear rationale. Some of these should be restored to report on trends in important factors affecting academic achievement.

It is recommended that the Governing Board consider the following:

3. *Have the expert panels developing KEI frameworks and specifications for main NAEP also make recommendations for KEIs in the areas under consideration using contextual variables in the long-term trend assessments.*

4. *Restore useful questions that were eliminated in the 2008 and 2012 administrations of long-term NAEP by adding them to the next administration.*

References

Alexander, K., Entwisle, D. and Horsey, C. (1997). From first grade forward: Early foundations of high school dropout. Sociology of Education 70, 87-107.

Bill and Melinda Gates Foundation (2012). *Asking students about teaching.* Available online November 2013 at www.metproject.org.

Bill and Melinda Gates Foundation (2010). *Learning about teaching: initial findings from the measures of effective teaching project.* Available online November 2013: http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCsQFjAA&url=http%3A%2F%2Fwww.metproject.org%2Fdownloads%2FPreliminary_Findings-Research_Paper.pdf&ei=xX5zUqbcGqnisATf3oGgBA&usg=AFQjCNGeNTDMnAlHtRHEQN6rMeihIFORWg&bvm=bv.55819444,d.cWc.

Bryk, A. (2010). "Organizing schools for improvement." *Phi Delta Kappan.* V91.n7 pp.23-30. Available December 2013 online: http://www.carnegiefoundation.org/sites/default/files/elibrary/bryk_organizing-schools_pdk.pdf.

Center for Research in Mathematics and Science Education. (2010). Breaking the Cycle: An International Comparison of U.S. Mathematics Teacher Preparation. East Lansing: Michigan State University. Available November 2013 online: http://www.educ.msu.edu/content/sites/usteds/documents/Breaking-the-Cycle.pdf.

Center for Social and Emotional Education (CSEE) & Education Commission of the States (ECS) (2008). *The school climate challenge: narrowing the gap between school climate research and school climate policy, practice guidelines and teacher education policy.* Available January 2014 online: http://www.ecs.org/html/projectsPartners/nclc/docs/school-climate-challenge-web.pdf

Chen, C. & Stevenson, H. (1995). Motivation and mathematics achievement: a comparative study of Asian-American, Caucasian, and East Asian high school students. Child Development, 66, 1215-1234

Clotfelter, C.T., Ladd, H.F., Vigdor, J.L., "Teacher credentials and student achievement in high school: a cross-subject analysis with student fixed effects," Calder Center, October 2007. Available November 2013 online: http://www.caldercenter.org/PDF/1001104_Teacher_Credentials_Hig

[h_School.pdf](http://www.ccsso.org/Documents/h_School.pdf)

Council of Chief State School Officers. (2013, April). *Interstate teacher assessment and support consortium In TASC model core teaching standards and learning progressions for teachers 1.0: a resource for ongoing teacher development.* Available November 2013 online: http://www.ccsso.org/Documents/2013/2013_INTASC_Learning_Progressions_for_Teachers.pdf.

Council for Accreditation of Educator Preparation (2013). *CAEP 2013 standards for accreditation of educator preparation.* Available January 2014 online: http://caepnet.org/accreditation/standards/

Cowan, C., Hauser, R., Kominski, R., Levin, H., Lucas, S., Morgan, S., Spencer, M., & Chapman, C. (2012). *Improving the measurement of socioeconomic status for the national assessment of educational progress: A theoretical foundation.* National Center for Education Statistics. Available January 2014 online: http://nces.ed.gov/nationsreportcard/pdf/researchcenter/socioeconomic_factors.pdf

Dweck, C. (2006). *Mindset.* Ballantine Books, Random House Publishing Group, NY.

Dweck, C. (2008). *Mindsets and math/science achievement.* Prepared for the Carnegie Corp of New York-Institute for Advanced Study. http://dev.opeq.blenderbox.com/uploads/files/868cea31-5888-4e45-a832-62b4377dbbfb.pdf

Education Commission of the States (2012). "Teacher expectations of students. A self-fulfilling prophesy " in *Progress of Education Reform. December 2012.* Available January 2014 online http://www.ntp16.notlb.com/sites/default/files/AVATAR/The%20Progress%20of%20Education%20Reform,%20December%202012%20-%20Teacher%20Expectations.pdf

Epstein, M., Atkins, M., Cullinan, D., Kutash, K., and Weaver, R. (2008). Reducing Behavior Problems in the Elementary School Classroom: A Practice Guide (NCEE #2008-012). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides

Gottfried, M (2011), The detrimental effect of missing school, American Journal of Education, v. 117, no. 2, Feb. 2011, p. 147-182.

Gutherie, J & Wigfield, A (2000). "Contexts for engagement and motivation in reading" in *Handbook of reading research: vol III* (Kamil, Mosenthal, Pearson, and Barr). Available December 2013 online*:* http://www.readingonline.org/articles/handbook/guthrie/

Hanushek, E., Kain, J., O'Brien, D., Rivkin, S. (2005). *The market for teacher quality.* NBER Working Paper No. 11154.

International Association for the Evaluation of Educational Achievement (IEA) (2012). TIMSS 2011 international results in mathematics. Ina V.S. Mullis, Michael O. Martin, Pierre Foy, and Alka Arora. Available November 2013 online: http://timss.bc.edu/timss2011/downloads/T11_IR_Mathematics_FullBook.pdf .

Kane, Thomas (2012). "Capturing the dimensions of effective teaching." *Education Next.* Fall 2012/Vol. 12, No. 4

McKown, C. and Weinstein, R. (2008) Journal of School Psychology, Vol. 46, No. 3, 2008, pp. 235-261

National Research Council. (2010). *Preparing teachers: Building evidence for sound policy.* Committee on the Study of Teacher Preparation Programs in the United States, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press. Available November 2013 online: http://www.nap.edu/download.php?record_id=12882

National Research Council (2012). Key National Education Indicators: Workshop Summary. Steering Committee on Workshop on Key National Education Indicators, A. Beatty and J.A. Koenig, Rapporteurs. Board on Testing and Assessment and Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Council for Accreditation of Teacher Education (n.d.)/ *What makes a teacher effective: a summary of key research findings. Available November 2013 online: http://www.ncate.org/public/researchreports/teacherpreparationresearch/whatmakesateachereffective/tabid/361/default.aspx.*

National Center for Education Statistics (2013). *Early childhood longitudinal study data products and publications.* Available January 2013 online http://nces.ed.gov/ecls/pdf/bibliography.pdf

New York City Public Schools (2013). *Educator guide.* Available December 2013 online: http://schools.nyc.gov/NR/rdonlyres/E25F8B70-1C47-4212-9D01-94EC0C56993C/0/EducatorGuide_HS_2013_01_04.pdf

Nichols, J. (2003). Prediction indicators for students failing the state of Indiana high school graduation exam. Preventing School Failure, 47, 112-120.

Neild, R. C., & Balfanz, R. (2006). Unfulfilled promise: The dimensions and characteristics of Philadelphia's dropout crisis, 2000-2005. Baltimore: Center for Social Organization of Schools, Johns Hopkins University.

OECD (2010). *Pisa results: learning to learn-vol III.* Available December 2013 online http://www.oecd.org/pisa/pisaproducts/48852630.pdf.

Rhodes, H. (2007**).** *Confronting the challenges of student engagement*: *a case study of a school-based intervention.* Rand Corp. Available January 2014 online:
http://www.rand.org/pubs/rgs_dissertations/RGSD218.html

Rice, J. (2010). *The impact of teacher experience.* National Center for Analyses of Longitudinal Data in Education Research, Urban Institute. Available November 2013:
http://www.urban.org/uploadedpdf/1001455-impact-teacher-experience.pdf

Rosenthal, R. and Jacobson, L. (1968). Pygmalion in the Classroom. New York: Holt, Rinehart & Winston.

Rubie-Davies, C., Hattie, J. and Hamilton, R. British Journal of Educational Psychology, Vol. 76, No. 3, 2006. pp. 429-444.

Snow, C. (2001). *Improving reading outcomes: Getting beyond third grade. Aspen Institute.* Available December 2013 online:
http://www.gse.harvard.edu/~snow/Aspen_snow.html

 Strauss, V. (2012). "Should students evaluate teachers?" *Washington Post,* December 18. Available November 2013 online:
http://www.nap.edu/catalog.php?record_id=12882http://www.washingtonpost.com/blogs/answer-sheet/wp/2012/12/18/should-students-evaluate-teachers/.

Tenenbaum, H. and Ruck, M. (2007) Journal of Educational Psychology, Vol. 99, No. 2, 2007, pp. 253-273.

Van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M. and Holland, R. American Educational Research Journal, Vol. 47, No. 2, 2010, pp. 497-527

Wilson, S., Floden, R. and Ferrini-Mundy, J. (2001). *Teacher preparation research: current knowledge, gaps, and recommendations* Center for the Study of Teaching and Policy. Available November 2013 online http://depts.washington.edu/ctpmail/publications/repo.