Using NAEP Data for Key Education Indicators

As authorized by the Governing Board Policy Statement on NAEP Background Data adopted in 2012, consultants have been preparing an exploratory analysis on using NAEP data for key education indicators. The purpose of this project is to illustrate the usefulness of NAEP in developing a limited number of indicators to represent crucial components of the education system and their interrelationships. The key idea is that instead of starting with contextual variables and looking for education issues they might address, there should first be a framework of important education policy issues and objectives that can be used to identify relevant contextual variables.

The work is being undertaken Marshall (Mike) S. Smith, former U.S. Under Secretary of Education and former Dean of the Stanford University Graduate School of Education, and Alan Ginsburg, former Director of Policy and Program Evaluation at the U.S. Department of Education. Smith chaired the Board's Expert Panel on Strengthening NAEP Background Questions, which presented its report in February 2012. Ginsburg served as a panel member and executive secretary, and has prepared several other exploratory analyses for the Board.

As explained in the statement of work for the project, an education indicator is an individual or composite statistic that measures progress toward an educational objective and is useful in a policy context. Such objectives are concerned not only with student performance but with the quality, equity, and efficiency of the education system in supporting academic achievement. One possible indicator might be the percentage of 8th grade science students with a teacher who majored or minored in science in college. Others might be the extent of severe absenteeism or the use of technology in science instruction.

At the working lunch Mike Smith will discuss the indicator project and the implications of using NAEP for this purpose. Board members will have a chance to ask questions and discuss the important issues involved.

Marshall S. (Mike) Smith

Marshall S. (Mike) Smith is retired and a Senior Fellow in Education Policy at the Carnegie Foundation for the Advancement of Teaching. He is a board member of a number of non-profit organizations in the San Francisco Bay Area. During the first two years of the Obama administration he served as Senior Counselor to Secretary of Education Arne Duncan and as Director of International Affairs. From 2001-2009 he directed the Education Program at the William and Flora Hewlett Foundation where he focused on developing the Open Education Resources movement, improving instruction, and reforming California's educational system.

Prior to that, in the Clinton Administration, he was the Undersecretary of Education for seven years responsible for all policy and budget matters. For the last four of those years he also served as the acting deputy secretary, the Education Department's second-ranked official under Secretary Richard Riley. During the Carter administration, he served as chief of staff to the first secretary for education, Shirley Hufstedler, and assistant commissioner for policy studies in the Office of Education. In the Ford administration he was the director of policy and budget for the National Institute of Education, the education research arm of the U.S. Government. While not in government, he was at different times an associate professor at Harvard University and a professor at the University of Wisconsin (at Madison) and at Stanford University. At Stanford, he was also the dean of the School of Education.

Smith has authored a large number of publications on topics varying from computer content analysis, evaluation and research methodology, social and educational inequality, early childhood education, open educational resources, federal policy, standards-based reforms and the use of technology in education in the developed and developing worlds. He is a member of the National Academy of Education and a fellow of the American Academy of Arts and Sciences. He holds bachelor's, master's, and doctorate degrees from Harvard.

ALAN GINSBURG

Alan Ginsburg was Director of Policy and Program Studies (retired) within the Office of Planning, Evaluation, and Policy Development at the U.S. Department of Education. He coordinated the Education Department's Government and Performance Results Act indicators and annual reports to Congress. Ginsburg's international work includes: Lead Shepherd (chair) of the Human Resources Development Working Group; and chair of the APEC Education Network (EDNET). His international mathematics work about Singapore and other Asian countries is extensively cited by the Common Core State Standards Initiative.

Ginsburg received his Ph.D. in economics from the University of Michigan. He received the Distinguished Presidential Rank Service Award, the federal government's highest award given to its civil service employees. He also received the American Evaluation Association's Gunnar Myrdal award for his contributions to the field of evaluation. He has been advisor to Education Week on their annual reports.

DEVELOPING A NAEP INDICATORS FRAMEWORK: LESSONS FROM MAJOR INTERNATIONAL AND DOMESTIC EDUCATION INDICATOR REPORTS

EXECUTIVE SUMMARY

By Alan Ginsburg and Marshall S. Smith

Introduction

This is the first of two reports exploring the use of background data collected by the National Assessment of Educational Progress (NAEP) to develop *key education indicators* at national, state, and urban district levels. *Key indicators are statistics that regularly measure an important condition of education.* For example, NAEP can tie to its achievement results the reporting of background conditions on: student attitudes toward learning, motivation and excessive absenteeism; measures of teacher quality; and indicators of the nature of reading and math instruction (e.g., instructional time).

The Government Accountability Office identified three broad purposes of indicators:

- Increase transparency and public awareness.
- Foster civic engagement and collaboration.
- Monitor progress, establish accountability for results, and aid decision-making.

In a NAEP context, indicators would also serve to:

- Identify for each subject assessed (e.g. reading) a set of key indicators, which are derived from the background variables and are continuously monitored.

Specifically, this first report is intended to develop a general indicators framework specifying an organizing structure, potential indicators, measurement criteria and reporting design. The report is based on a review of several major international and domestic data collections and reports produced by organizations other than NAEP:

International

- Organization for Economic Cooperation and Development, *Education At a Glance*
- International Education Association's 2011 TIMSS *Mathematics Assessment* covering grade 4 and 8.
- OECD's 2009 *PISA Report*

Domestic

- National Center for Education Statistics, *Condition of Education*
- Education Week's *Quality Counts*
- U.S. Department of Education's *Annual Priority Performance Goals*
- National Academy of Sciences' *Key National Education Indicators*

Potential Indicators by Organizing Structure

Exhibit EX-1

Potential NAEP Education Indicators From Which To Select Key Indicators For K-12			
Locus of Education Activity	**Key Drivers** - 21st Century Skills - Common Core Standards - Instructional Technology		
	Results	**Enablers**	**Context/Constraints**
Student	• Command of core content, using NAEP scores • College readiness levels by age and grade • Career readiness (21st century skills)	• Attended preprimary education • Chronic absenteeism • Student motivation and belief that hard work is more important than luck • Student positive attitudes toward subject • Student uses research-based approaches to learning subject • Student respect for teacher and visa versa • Participation in extra-curricular activities including community service	• Home learning environment • Formal and informal learning outside school – nature of the their neighborhood
Teacher	• Proportion of teacher evaluations that distinguish them from a basic standard • Quality of work that the students have • Teachers spend time supporting other teachers	• Teachers with less than 3-years experience • Teachers with mastery-level and current knowledge of content they are teaching. • Teachers with mastery-level and cotemporary knowledge of child and adolescent development • Teacher-student interactions that demonstrate high levels and qualities of involvement, language, stimulation, and expansion of thinking and cognition, and sensitivity to students' perspectives, individual experiences, and backgrounds • Teacher student interactions that indicate that teachers respect students.	• Teacher working conditions • Average district teacher salary • Time teachers spend teaching • Teacher has high quality professional development and comprehensive induction programs Quality of the principal • Teachers belong to professional learning communities
School/ Classroom	• School subject area assessment outcomes • School performance rating/ranking within their state • Parent satisfaction (on surveys) • Completion rates from each kind of school – elementary to middle, middle to high, high to graduate, graduate to college or job?	• Content of instruction aligned with standards • Effective use of technology to support instruction • School Climate – whether the school is a learning organization – do teachers work together? • Instructional time per subject • Engaged instruction in subject • Emphasis on continuous improvement on outcomes through both formative and summative assessments aligned with standards • Emphasis on continuous improvement of practices of teaching	• School SES Composition • Safe & orderly school climate • Teacher-student ratio • School resource shortages • School lacks key characteristics, coaches for teachers, support systems for students, technology, books
System district, state or nation)	- System core content outcomes	• Support for implementation of new content standards • Alignment of assessment with content standards • Accountability with emphasis on continuous improvement	• K-12 education spending as a share of gross domestic product • K-12 spending per student • Disparity in resources across districts within states

The indicator structure in Exhibit EX-1 is focused primarily around variables at student, teacher, school/classroom and system levels that support learning outcomes across three aspects of education conditions:

- *Results* indicators include student assessment outcomes (such as from NAEP), but also teacher evaluations that reflect student outcomes, and other outcomes such as secondary school completion and parent satisfaction with the school.

- *Enabler indicators* reflect formal learning at different levels of education. These include student exposure to preschool, teachers' knowledge and skills and their ability to apply them to create a challenging and supportive classroom learning environment; and school instructional time and student engagement in the content areas. Enablers also include system policies and regulations at district, state and national levels regarding teacher certification, standards, assessment, and accountability.

- *Context/constraint indicators* reflect factors not readily manipulable by educators but may be changed by policy and funding shifts or proper interventions in the home learning environment. These factors include: learning at home and outside the school in formal and informal settings; factors influencing teacher quality, such as salaries and working conditions; and factors affecting the school learning environment including school safety, climate and class size.

Indicator Measurement

A sound measure for an indicator should meet criteria of validity, reliability, and consistency overtime.

Validity of Indicators. A valid measure is one that adequately captures the underlying education condition of interest. Combining responses from a number of questions around a topic into a larger comprehensive indicator scale produces richer indicator measures than reporting on a single question, but this approach currently is not used in NAEP background factor analyses. Exhibit EX-2 illustrates a scale developed from TIMSS at grade 4 measuring students' early numeracy activities before beginning primary school.

Exhibit EX-2 Development of Indicator Scales from Multiple Questions

Exhibit 4.9: Early Numeracy Activities Before Beginning Primary School*

TIMSS 2011 Mathematics 4th Grade

Reported by Parents

Students were scored according to their parents' frequency of doing the six activities on the *Early Numeracy Activities* scale. Students **Often** engaged in early numeracy activities had a score on the scale of at least 10.3, which corresponds to their parents "often" doing three of the six activities with them and "sometimes" doing the other three, on average. Students **Never or Almost Never** engaged in such activities had a score no higher than 6.9, which corresponds to parents "never or almost never" doing three of the six activities with them and "sometimes" doing the other three, on average. All other students had parents who **Sometimes** engaged them in early numeracy activities.

Country	Often		Sometimes		Never or Almost Never		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Hungary	75 (0.9)	528 (2.9)	23 (0.8)	495 (4.9)	1 (0.4)	~ ~	11.1 (0.04)
Czech Republic	75 (0.8)	514 (2.3)	25 (0.8)	508 (3.6)	0 (0.1)	~ ~	11.0 (0.03)
Slovak Republic	73 (1.0)	514 (3.3)	25 (0.8)	499 (5.2)	2 (0.5)	~ ~	11.1 (0.05)
Northern Ireland s	70 (1.2)	583 (3.5)	29 (1.2)	566 (4.9)	1 (0.2)	~ ~	11.2 (0.05)
Russian Federation	69 (1.1)	547 (3.7)	29 (1.0)	533 (4.7)	2 (0.3)	~ ~	10.9 (0.04)
Poland	68 (0.9)	488 (2.3)	31 (0.9)	471 (3.0)	1 (0.2)	~ ~	10.8 (0.03)
Ireland	66 (0.9)	539 (2.9)	33 (0.8)	517 (3.7)	2 (0.3)	~ ~	10.9 (0.04)
Australia s	61 (1.2)	540 (3.7)	36 (1.2)	520 (4.1)	3 (0.4)	488 (13.4)	10.7 (0.05)
Austria	61 (0.9)	515 (2.6)	38 (0.9)	502 (3.5)	2 (0.2)	~ ~	10.4 (0.03)
Croatia	60 (0.8)	496 (2.3)	39 (0.8)	482 (2.6)	1 (0.2)	~ ~	10.5 (0.03)
Germany r	59 (1.0)	538 (2.3)	40 (1.0)	528 (2.8)	2 (0.2)	~ ~	10.4 (0.03)

Source: IEA, *TIMSS,* 2011

Reliability of Indicators. A reliable indicator produces consistent results when repeatedly measuring the same underlying condition. Qualitative responses may be unreliable when sensitive to the position of the respondent. For example, Exhibit EX-3, taken from the NAEP background paper on science (by Alan Friedman and Alan Ginsburg), shows that teachers were more likely to indicate that resources within a school are "not at all available" than were principals in the same school. This is not surprising as it is principals who are responsible for school resource availability.

Exhibit EX-3 Differences between teacher and school reported responses about science resource availability raise issues of response reliability

	Not at all	Small Extent	Moderate extent	Large extent
Science Kits are provided (teacher reported)	26	30	29	16
Science Kits are provided (school reported)	7	24	32	37
Science magazines and books are provided (teacher reported)	22	35	33	11
Science magazines and books are provided (school reported)	2	19	35	44

NAEP Data Explorer

Consistency over time. A consistent measure requires using the same measure for an indicator over time. When measures are changed from time period to time period it is unclear whether a change comes about because of a real change in the underlying condition or changes in the measure. The report by the Expert Panel on Strengthening NAEP Background Questions (2012) addressed this issue in its recommendation 1d:

> "Use consistency over time as a criterion to consider for question selection and wording. NAEP's inconsistent inclusion of background questions weakens its potential to track trends and improvements within a subject area and topic.

For example, the Expert Panel found that only one-third of the 2011 questions asking about course offerings provided at least a 6-year trend. No 2011 questions about curriculum or school resources were found on the 2005 or earlier questionnaires.

Sources of Indicator Data. The reports that were studied use two ways to obtain indicator data, which differentiate them from NAEP.

First, TIMSS and PISA both conduct a household survey to obtain information directly from parents or guardians about socio-economic status and the home learning environment. TIMSS innovatively combined with PIRLS to develop a joint sample household survey for grade 4 students. The household survey included questions about:

- Early numeracy activities in the home before beginning primary school (See Exhibit EX-2)
- Early literacy activities in the home before beginning primary school
- Amount of exposure to preschool
- Family perception about child's literacy and numeracy skills before entering primary school
- Family interaction with the child about school work
- Family perceptions about school
- Family literacy environment
- Family SES

A second source of data that is different from NAEP is the pooling of information across different surveys. The Condition of Education and Education at a Glance are drawn almost entirely from data series generated by other surveys. Quality Counts is a state-level amalgam of direct analyses of state policies by Education Week combined with data from other surveys, which prominently features NAEP assessment results.

A form of pooling could be the aligning of NAEP survey questions with international assessment items as illustrated in Exhibit EX-4. The exhibit suggests that at least for U.S. middle schools, only about 12 percent of U.S. principals are having at least some difficulty filling vacancies for mathematics teachers. This compares with other Western English-speaking countries of 41 percent of the principals having difficulty hiring math teachers in Australia, 37 percent in England, and 44 percent in New Zealand. Adding

the same question about vacancies to the NAEP principal survey for mathematics would yield U.S. state-by-state comparisons.

Exhibit EX-4

Schools Having Difficulties Filling Vacancies With Mathematics Teachers, Grade 8

Country	No Vacancies	Vacancies Are Easy To Fill	Vacancies Are Somewhat Difficult To Fill	Vacancies Are Very Difficult to Fill	Total of Vacancies Somewhat or Very Difficut To Fill
	Percent of students	Percent of students	Percent of students	Percent of students	Percent of students
Australia	25	34	31	10	41
Chinese Taipei	46	44	10	1	11
England	28	35	27	10	37
Finland	42	46	10	1	11
Hong Kong SAR	48	44	8	0	8
Japan	82	6	8	3	11
Korea, Rep. of	67	16	15	2	17
New Zealand	30	27	38	6	44
Norway	38	40	20	2	22
Russian Federatio	81	11	6	2	8
Singapore	59	38	2	0	2
United States	63	25	9	3	12

Source : 2011 TIMSS, Mathematics

Next Steps: Using the International and Domestic Indicator Framework as a Guide, Develop a NAEP Education Indicators Framework and Provide Examples with Current Data

A second report will be prepared for the Governing Board with a recommended set of *Key Indicators* and recommended improvements in NAEP data to strengthen indicator measurement or fill indicator gaps. This report will:

- Specify a NAEP Education Indicators Framework for the background variables applicable across cognitive assessments.
- Propose indicators that are research-based and estimable by:
  - offering examples using current NAEP data.
  - suggesting changes to the current NAEP questionnaires.
  - introducing a fundamentally new NAEP questionnaire or drawing data from education surveys other than NAEP.
- Explore opportunities for combining NAEP with international or other NCES indicator-supporting data.
- Explore how NAEP reports could best display a pyramid information approach along the lines of an indicator dashboard.