National Assessment Governing Board
Committee on Standards, Design and Methodology

May 16, 2014
10:00 am – 12:30 pm

AGENDA

10:00 – 10:05 am	Introductions, Review of Agenda, and Welcome to Former Wyoming Governor James Geringer *Lou Fabrizio, COSDAM Chair*	Attachment A
10:05 – 10:45 am	NAEP Testing and Reporting on Students with Disabilities *Grady Wilburn, NCES*	Attachment B
10:45 – 11:25 am	Trends and the Transition to Technology-Based Assessments *Andreas Oranje, Educational Testing Service*	Attachment C
11:25 – 12:05 pm	Update on Academic Preparedness Research *Sharyn Rosenberg, NAGB Staff*	Attachment D
12:05 – 12:25 pm	Update on Development of TEL Achievement Levels Descriptions *Sharyn Rosenberg, NAGB Staff*	Attachment E
12:25 – 12:30 pm	Other Issues and Questions *COSDAM Members*	
	Information Items: • Update on Evaluation of NAEP Achievement Levels Procurement • Update on TEL Achievement Levels Setting Procurement	Attachment F Attachment G

Welcome to Governor James Geringer,
New Governing Board Member

The Committee on Standards, Design and Methodology welcomes Former Wyoming Governor James Geringer. Governor Geringer was appointed to the Board on March 4, 2014 by Secretary of Education Arne Duncan. Governor Geringer will be sworn in on May 16th and will serve in the category of Republican Governor.

Abbreviated Professional Biography for Governor James Geringer

The Honorable Geringer was the 30th governor of Wyoming, serving from 1995 to 2013. Governor Geringer, whose background includes mechanical engineering and agriculture production, also served in the Wyoming Legislature from 1983 to 1994, including six years each in the House and the Senate.

In his two terms as governor, he focused on improving education through standards, accountability and technology; modernizing Wyoming's economic base to extensively include technology; changing how natural resource agencies among state, federal and local governments worked together; establishing community-based health and family services programs; and implementing strategic planning and information technology systems.

Governor Geringer has focused on education policy and leadership through his involvement in such organizations as the National Commission on Mathematics and Science Teaching for the 21st Century, the Education Commission of the States, Western Governors University, and the National Research Council at the National Academies. He joined Environmental Systems Research Institute (ESRI) in 2003 as director of policy and public sector strategies to work with senior elected and corporate officials to use geospatial technology in business and government.

Currently Governor Geringer chairs the board of Complete College America, a national nonprofit focused on working with states to increase the number of Americans with career certificates or college degrees, and to close attainment gaps for underrepresented populations.

NAEP Testing and Reporting on Students with Disabilities

The March 2010 Governing Board policy on NAEP Testing and Reporting on Students with Disabilities (SD) and English Language Learners (ELL) was intended to reduce exclusion rates and provide more consistency across jurisdictions in which students are tested on NAEP to promote sound reporting of comparisons and trends (the policy statement is included as Attachment B2). The policy limits the grounds on which schools can exclude students from NAEP samples to two categories – for SD, only those with the most significant cognitive disabilities, and for ELL, only those who have been in U.S. schools for less than a year. Previously, schools excluded students with Individualized Education Programs (IEPs) that called for accommodations on state tests that NAEP does not allow, primarily the read-aloud accommodation on the Reading assessment. Under the current Board policy, schools should no longer decide to exclude students whose IEPs for state tests specify an accommodation not allowed on NAEP. Instead, such students should take NAEP with allowable accommodations. Additionally, parents and educators should be encouraged to permit them to do so, given that NAEP provides no scores and causes no consequences for individuals but needs fully representative samples to produce valid results for the groups on which it reports. By law, individual participation in NAEP is voluntary and parents may withdraw their children for any reason.

During the December 2013 Board meeting, the Committee on Standards, Design, and Methodology and the Reporting and Dissemination Committee met in joint session to discuss the 2013 student participation data for grades 4 and 8 Reading and Mathematics. There have been large increases in inclusion rates over the past ten years, and the Board's first inclusion rate goal—95 percent of all students in each sample—was met in almost all states in 2013. However, 11 states and eight districts failed to meet the Board's second goal of testing at least 85 percent of students identified as SD or ELL. Contrary to Board policy, NCES has continued to permit schools to exclude students whose IEPs called for accommodations that NAEP does not allow. NCES believes changing this practice could possibly be detrimental to students, increase refusals, change NAEP's target population, and be counter to current statistical procedures.

At the end of the December 2013 joint session, the Committees asked the staffs of NAGB and NCES to consider possible policy and operational changes and what their impact might be, as well as a timeline for possible Board action.

The staffs of NAGB and NCES have had several conversations about the implementation of the SD/ELL policy. The policy could be clarified by revising the language about converting excluded students to refusals. The fourth implementation guideline for students with disabilities states, *"Students refusing to take the assessment because a particular accommodation is not allowed should not be classified as exclusions but placed in the category of refusals under NAEP data analysis procedures."* NCES asserts that it is technically incorrect to apply a weight class

adjustment[1] that combines students who did not participate due to receiving accommodations on their state tests that are not allowed on NAEP with students who refused for other reasons. The former group cannot be assumed to be randomly missing, which is a necessary assumption for the current NAEP statistical procedures.

In the May 2014 COSDAM session, NCES will present alternative methods for estimating scores for such students, so that their lack of participation can be considered appropriately when calculating NAEP scores.

[1] This refers to a set of units (e.g., schools or students) that are grouped together for the purpose of calculating nonresponse adjustments. The units are homogeneous with respect to certain unit characteristics, such as school size, location, public/private, student's age, sex, and student disability status.

Reconciling NAEP and State Test Accommodation Differences for Estimating Scores

The National Assessment Governing Board created the NAEP Testing and Reporting on Students with Disabilities (SD) and English Language Learners (ELL) policy because the validity of NAEP results was being threatened by high and substantial variation in exclusion rates. Since this policy was adopted in 2010, the majority of jurisdictions have improved their inclusion rates of SD and ELL students and variation amongst the jurisdictions has significantly decreased. Of the students that are still excluded, the overwhelming majority require the read aloud accommodation on the NAEP reading assessment, which NAEP does not allow. Therefore, NCES is exploring ways to meaningfully represent this group of students in the estimates for the jurisdictions from which they were sampled.

In the absence of a read aloud accommodation for reading and, subsequently, direct estimates for this group of students, some indirect estimation has to be applied. The main question is: *how* indirect an estimate can be while still yielding reliable and valid NAEP estimates? Another way of phrasing this is: what type of data can be collected and used to yield appropriate estimates for students denoted as 'refusal' due to a requested accommodation not being available *and* the decision not to let the student participate? There are a number of options that could support accurate estimation:

1. As a fall-back, the Full Population Estimates (FPE) as currently developed and reported would not require additional data collection and could be used to represent this population specifically, or alongside excluded populations. NCES would explore cost-effective ways to include additional background data from school records and teacher and school questionnaires to the extent that those are not already available and used in the FPE process. These additional data could improve those estimates further.

2. Students selected for the reading assessment who cannot participate because they require a read-aloud accommodation could instead participate in an alternative subject (e.g., mathematics). The student questionnaire data collected from these students could be used to improve estimates. Note that these students would not be placed in a different sample (e.g., switched from reading to math), which would only exacerbate a lack of clarity about the population of inference, but rather the student questionnaire data would be collected to estimate appropriate reading scores for them.

3. As a variant of (2) above, subject performance data (e.g., mathematics) might be used to estimate these students' reading ability. This approach, if feasible, would be particularly improved if we would obtain some notion about the relationship between the two subjects and if the students that require the read-aloud accommodation but still participate in the reading assessment can be identified as a target distribution for estimation.

In this session, NCES will discuss these options in more detail and provide a recommendation on which method would best serve NAEP.

Adopted: March 6, 2010

National Assessment Governing Board

NAEP Testing and Reporting on Students with Disabilities and English Language Learners

Policy Statement

INTRODUCTION

To serve as the Nation's Report Card, the National Assessment of Educational Progress (NAEP) must produce valid, comparable data on the academic achievement of American students. Public confidence in NAEP results must be high. But in recent years it has been threatened by continuing, substantial variations in exclusion rates for students with disabilities (SD) and English language learners (ELL) among the states and urban districts taking part.

Student participation in NAEP is voluntary, and the assessment is prohibited by law from providing results for individual children or schools. But NAEP's national, state, and district results are closely scrutinized, and the National Assessment Governing Board (NAGB) believes NAEP must act affirmatively to ensure that the samples reported are truly representative and that public confidence is maintained.

To ensure that NAEP is fully representative, a very high proportion of the students selected must participate in its samples, including students with disabilities and English language learners. Exclusion of such students must be minimized; they should be counted in the Nation's Report Card. Accommodations should be offered to make the assessment accessible, but these changes from standard test administration procedures should not alter the knowledge and skills being assessed.

The following policies and guidelines are based on recommendations by expert panels convened by the Governing Board to propose uniform national rules for NAEP testing of SD and ELL students. The Board has also taken into consideration the views expressed in a wide range of public comment and in detailed analyses provided by the National Center for Education Statistics, which is responsible for conducting the assessment under the policy guidance of the Board**.** The policies are presented not as statistically-derived standards but as policy guidelines intended to maximize student participation, minimize the potential for bias, promote fair comparisons, and maintain trends. They signify the Board's strong belief that NAEP must retain public confidence that it is fair and fully-representative of the jurisdictions and groups on which the assessment reports.

POLICY PRINCIPLES

1. As many students as possible should be encouraged to participate in the National Assessment. Accommodations should be offered, if necessary, to enable students with disabilities and English language learners to participate, but should not alter the constructs assessed, as defined in assessment frameworks approved by the National Assessment Governing Board.

2. To attain comparable inclusion rates across states and districts, special efforts should be made to inform and solicit the cooperation of state and local officials, including school personnel who decide upon the participation of individual students.

3. The proportion of all students excluded from any NAEP sample should not exceed 5 percent. Samples falling below this goal shall be prominently designated in reports as not attaining the desired inclusion rate of 95 percent.

4. Among students classified as either ELL or SD a goal of 85 percent inclusion shall be established. National, state, and district samples falling below this goal shall be identified in NAEP reporting.

5. In assessment frameworks adopted by the Board, the constructs to be tested should be carefully defined, and allowable accommodations should be identified.

6. All items and directions in NAEP assessments should be clearly written and free of linguistic complexity irrelevant to the constructs assessed.

7. Enhanced efforts should be made to provide a short clear description of the purpose and value of NAEP and of full student participation in the assessment. These materials should be aimed at school personnel, state officials, and the general public, including the parents of students with disabilities and English language learners. The materials should emphasize that NAEP provides important information on academic progress and that all groups of students should be counted in the Nation's Report Card. The materials should state clearly that NAEP gives no results for individual students or schools, and can have no impact on student status, grades, or placement decisions.

8. Before each state and district-level assessment NAEP program representatives should meet with testing directors and officials concerned with SD and ELL students to explain NAEP inclusion rules. The concerns of state and local decision makers should be discussed.

IMPLEMENTATION GUIDELINES

For Students with Disabilities

1. Students with disabilities should participate in the National Assessment with or without allowable accommodations, as needed. Allowable accommodations are any changes from standard test administration procedures, needed to provide fair access by students with disabilities that do not alter the constructs being measured and produce valid results. In cases where non-standard procedures are permitted on state tests but not allowed on NAEP, students will be urged to take NAEP without them, but these students may use other allowable accommodations that they need.

2. The decision tree for participation of students with disabilities in NAEP shall be as follows:

NAEP Decision Tree for Students with Disabilities

BACKGROUND CONTEXT

1. NAEP is designed to measure constructs carefully defined in assessment frameworks adopted by the National Assessment Governing Board.

2. NAEP provides a list of appropriate accommodations and non-allowed modifications in each subject. An appropriate accommodation changes the way NAEP is normally administered to enable a student to take the test but does not alter the construct being measured. An inappropriate modification changes the way NAEP is normally administered but does alter the construct being measured.

STEPS OF THE DECISION TREE

3. In deciding how a student will participate in NAEP:

   a. If the student has an Individualized Education Program (IEP) or Section 504 plan and is tested without accommodation, then he or she takes NAEP without accommodation.

   b. If the student's IEP or 504 plan specifies an accommodation permitted by NAEP, then the student takes NAEP with that accommodation.

   c. If the student's IEP or 504 plan specifies an accommodation or modification that is not allowed on NAEP, then the student is encouraged to take NAEP without that accommodation or modification.

3. Students should be considered for exclusion from NAEP only if they have previously been identified in an Individualized Education Program (IEP) as having the most significant cognitive disabilities, and are assessed by the state on an alternate assessment based on alternate achievement standards (AA-AAS). All students tested by the state on an alternate assessment with modified achievement standards (AA-MAS) should be included in the National Assessment.

4. Students refusing to take the assessment because a particular accommodation is not allowed should not be classified as exclusions but placed in the category of refusals under NAEP data analysis procedures.

5. NAEP should report separately on students with Individualized Education Programs (IEPs) and those with Section 504 plans, but (except to maintain trend) should only count the students with IEPs as students with disabilities. All 504 students should participate in NAEP.

   At present the National Assessment reports on students with disabilities by combining results for those with an individualized education program (who receive special education services under the Individuals with Disabilities Education Act [IDEA]) and students with Section 504 plans under the Rehabilitation Act of 1973 (a much smaller group with disabilities who are not receiving services under IDEA but may be allowed test accommodations).[*] Under the Elementary and Secondary Education Act, only those with an IEP are counted as students with disabilities in reporting state test results. NAEP should be consistent with this practice. However, to preserve trend, results for both categories should be combined for several more assessment years, but over time NAEP should report as students with disabilities only those who have an IEP.

6. Only students with an IEP or Section 504 plan are eligible for accommodations on NAEP. States are urged to adopt policies providing that such documents should address participation in the National Assessment.

For English Language Learners

1. All English language learners selected for the NAEP sample who have been in United States schools for one year or more should be included in the National Assessment. Those in U.S. schools for less than one year should take the assessment if it is available in the student's primary language.

[*] NOTE: The regulation implementing Section 504 defines a person with a disability as one who has a physical or mental impairment which substantially limits one or more major life activities, has a record of such an impairment, or is regarded as having such an impairment. 34 C.F.R. § 104.3(j)(1).

One year or more shall be defined as one full academic year before the year of the assessment.

2. Accommodations should be offered that maximize meaningful participation, are responsive to the student's level of English proficiency, and maintain the constructs in the NAEP framework. A list of allowable accommodations should be prepared by NAEP and furnished to participating schools. Such accommodations may be provided only to students who are not native speakers of English and are currently classified by their schools as English language learners or limited English proficient (LEP).

3. Bilingual versions of NAEP in Spanish and English should be prepared in all subjects, other than reading and writing, to the extent deemed feasible by the National Center for Education Statistics. The assessments of reading and writing should continue to be in English only, as provided for in the NAEP frameworks for these subjects.

4. Staff at each school should select from among appropriate ELL-responsive accommodations allowed by NAEP, including bilingual booklets, those that best meet the linguistic needs of each student. Decisions should be made by a qualified professional familiar with the student, using objective indicators of English proficiency (such as the English language proficiency assessments [ELPA] required by federal law), in accordance with guidance provided by NAEP and subject to review by the NAEP assessment coordinator.

5. Schools may provide word-to-word bilingual dictionaries (without definitions) between English and the student's primary language, except for NAEP reading and writing, which are assessments in English only.

6. NAEP results for ELL students should be disaggregated and reported by detailed information on students' level of English language proficiency, using the best available standardized assessment data. As soon as possible, NAEP should develop its own brief test of English language proficiency to bring consistency to reporting nationwide.

7. Data should be collected, disaggregated, and reported for former English language learners who have been reclassified as English proficient and exited from the ELL category. This should include data on the number of years since students exited ELL services or were reclassified.

8. English language learners who are also classified as students with disabilities should first be given linguistically-appropriate accommodations before determining which additional accommodations may be needed to address any disabilities they may have.

RESEARCH AND DEVELOPMENT

The Governing Board supports an aggressive schedule of research and development in the following areas:

1. The use of plain language and the principles of universal design, including a plain language review of new test items consistent with adopted frameworks.

2. Adaptive testing, either computer-based or paper-and-pencil. Such testing should provide more precise and accurate information than is available at present on low-performing and high-performing groups of students, and may include items appropriate for ELLs at low or intermediate levels of English proficiency. Data produced by such targeted testing should be placed on the common NAEP scale. Students assessed under any new procedures should be able to demonstrate fully their knowledge and skills on a range of material specified in NAEP frameworks.

3. A brief, easily-administered test of English language proficiency to be used for determining whether students should receive a translation, adaptive testing, or other accommodations because of limited English proficiency.

4. The validity and impact of commonly used testing accommodations, such as extended time and small group administration.

5. The identification, measurement, and reporting on academic achievement of students with the most significant cognitive disabilities. This should be done in order to make recommendations on how such students could be included in NAEP in the future.

6. A study of outlier states and districts with notably high or low exclusion rates for either SD or ELL students to identify the characteristics of state policies, the approach of decision makers, and other criteria associated with different inclusion levels.

The Governing Board requests NCES to prepare a research agenda on the topics above. A status report on this research should be presented at the November 2010 meeting of the Board.

Trends and the Transition to Technology-Based Assessments

The NAEP program is transitioning from paper-and-pencil based administrations to technology-based administration (TBA) over the next two to four years. In order for the program to fulfill one of its core charges, to provide trend information about what U.S. students know and can do in various subjects, this transition is not necessarily straightforward. Foremost, any changes due to differences in the mode of administration can conflate with differences due to changes in performance across years. In order to understand, and if possible, account for mode differences, several features have been designed and incorporated in the assessments during the transition:

1. Before any actual operational implementation, TBA-start-ups are conducted. These start-ups are designed to be the bridge between paper- and technology-based assessment modes, based on two randomly equivalent samples each taking one of the two modes. The paper-based component is the operational sample. while the TBA component is the bridge to future technology-based assessments
2. The TBA start-up makes use of existing paper-based items that are represented as technology-based items in order to impart fairly strict consistency in terms of content.
3. Item and task types that make broader use of the range of possibilities afforded by technology-based assessment will be introduced gradually into each subject area assessment.

This design has been developed as a way to increase the chances of maintaining trends (particularly through the use of existing paper-based items in the TBA start-up, as described by feature 2 above) while making a meaningful transition. Yet, it is important to realize that this is an empirical matter. Empirical criteria for evaluating the success of the bridge mostly center around the notion of student group comparisons. That is, would we reach substantially different conclusions between the two administration modes in terms of student group differences?

In this presentation, COSDAM will hear about the above features and questions in more detail. Maintaining trends will likely not be a single event nor a temporary concern. It is quite likely that the notion of mode changes will become a continuously present characteristic as technology develops at increasingly faster paces. Formats and standards of today in terms of device specifications and user interaction will be very different from those in the future. Therefore, the question about maintaining trends will have to be extended and generalized. The basic idea is that decisions of trend will always have to be data driven and that much can be done through experimentation within existing data collection paradigms to gather pertinent data in order to evaluate and estimate trends appropriately.

NAEP 12th Grade Academic Preparedness Research

Phase 1 Research

The first phase of the Governing Board's research on academic preparedness is now complete; results from more than 30 studies are available at: http://www.nagb.org/what-we-do/preparedness-research.html. During the August 2013 meeting, the Board voted on a motion to use the phase 1 research on academic preparedness for college in the reporting of the 2013 grade 12 national results for reading and mathematics, to be released on May 14, 2014. The motion, validity argument, and reports all emphasize the preliminary nature of the inferences for reporting purposes and specifically not as performance standards for academic preparedness for college. There is also repeated emphasis on the continued research being conducted during the second phase of the Board's academic preparedness research, which is intended to inform reporting of the 2015 grade 12 results for reading and mathematics.

May 2014 Update: During the May 2014 meeting, COSDAM will discuss the release of the 2013 grade 12 results in terms of academic preparedness for college, with a focus on how the technical issues from the research were conveyed and received.

Phase 2 Research

The second phase of the Governing Board's research on academic preparedness currently consists of the following studies that are planned or underway:

Study name	Sample	May 2014 Update
Statistical linking of NAEP and ACT	National; FL, IL, MA, MI, TN	Page 15
Longitudinal statistical relationships: Grade 12 NAEP	FL, IL, MA, MI, TN	Page 16
Statistical linking of NAEP and EXPLORE	KY, NC, TN	Page 17
Longitudinal statistical relationships: Grade 8 NAEP	KY, NC, TN	Page 18
Content alignment of NAEP and COMPASS		Page 19
Content alignment of NAEP and EXPLORE		Page 20
College Course Content Analysis		Pages 22-27
Evaluating Reading and Mathematics Frameworks and Item Pools as Measures of Academic Preparedness for College and Job Training (Research with Frameworks)		Pages 28-33

Brief overviews and informational updates are provided for each study.

May 2014 Update: We are preparing to convene a technical advisory group to make recommendations to the Board about the second phase of the research on academic preparedness for college, including:

- Conducting a standard setting procedure (e.g., evidence-based standard setting) to establish a grade 12 academic preparedness standard based on the research findings from phase one and phase two; and
- Evaluating the results from the grade 8 research questions to explore the feasibility of reporting future grade 8 NAEP results for reading and mathematics in terms of being on track for academic preparedness for college.

During the May 2014 meeting, COSDAM will be invited to provide any initial feedback on the above issues as well as additional issues and questions that COSDAM would like to see presented to the technical advisory group.

National and State Statistical Linking Studies with the ACT

The Governing Board is planning to partner with ACT, Inc. to conduct a statistical linking study at the national level between NAEP and the ACT in Reading and Mathematics. Through a procedure that protects student confidentiality, the ACT records of 12th grade NAEP test takers in 2013 will be matched, and through this match, the linking will be performed. A similar study at the national level was performed with the SAT in 2009. There will not be a national statistical linking study performed for NAEP and the SAT in 2013.

In addition, the state-level studies, begun in 2009 with Florida, will be expanded with 2013 NAEP. Again using a procedure that protects student confidentiality, ACT scores of NAEP 12th grade test takers in the state samples in partner states will be linked to NAEP scores. We are in the planning stages with five states to be partners in these studies at grade 12: Florida, Illinois, Massachusetts, Michigan, and Tennessee. In three of these states (IL, MI, TN), the ACT is administered to all students state-wide, regardless of students' intentions for postsecondary activities.

Draft Research Questions for National and State Statistical Linking Studies with the ACT:

1. What are the correlations between the grade 12 NAEP and ACT student score distributions in Reading and Math?
2. What scores on the grade 12 NAEP Reading and Math scales correspond to the ACT college readiness benchmarks? (concordance and/or projection)
3. What are the average grade 12 NAEP Reading and Math scores and interquartile ranges (IQR) for students below, at, and at or above the ACT college readiness benchmarks?
4. Do the results differ by race/ethnicity or gender?

May 2014 Update: The data sharing agreements are still in the process of being finalized.

Longitudinal Statistical Relationships: Grade 12 NAEP

In addition to the linking of ACT scores to NAEP 12th grade test scores in partner states, the postsecondary activities of NAEP 12th grade test takers will be followed for up to six years using the state longitudinal databases in Florida, Illinois, Massachusetts, Michigan, and Tennessee. These studies will examine the relationship between 12th grade NAEP scores and scores on placement tests, placement into remedial versus credit-bearing courses, GPA, and persistence.

Draft Research Questions for Longitudinal Statistical Relationships, Grade 12 NAEP:

1. What is the relationship between grade 12 NAEP Reading and Math scores and grade 8 state test scores?
2. What are the average grade 12 NAEP Reading and Math scores and interquartile ranges (IQR) for students with placement in remedial and non-remedial courses?
3. What are the average grade 12 NAEP Reading and Math scores (and the IQR) for students with a first-year GPA of B- or above?
4. What are the average grade 12 NAEP Reading and Math scores (and the IQR) for students who remain in college after each year?
5. What are the average grade 12 NAEP Reading and Math scores (and the IQR) for students who graduate from college within 6 years?

May 2014 Update: The data sharing agreements are still in the process of being finalized.

State Statistical Linking Studies with EXPLORE

In 2013, linking studies between 8th grade NAEP in Reading and Mathematics and 8th grade EXPLORE, a test developed by ACT, Inc. that is linked to performance on the ACT, are planned with partners in three states: Kentucky, North Carolina, and Tennessee. In all three of these states, EXPLORE is administered to all students state-wide during grade 8.

Draft Research Questions for State Statistical Linking Studies with EXPLORE:

1. What are the correlations between the grade 8 NAEP and EXPLORE scores in Reading and Math?
2. What scores on the grade 8 NAEP Reading and Math scales correspond to the EXPLORE college readiness benchmarks (concordance and/or projection)?
3. What are the average grade 8 NAEP Reading and Math scores and the interquartile ranges (IQR) for students below, at, and at or above the EXPLORE college readiness benchmarks?

May 2014 Update: The data sharing agreements are still in the process of being finalized.

Longitudinal Statistical Relationships: Grade 8 NAEP

In 2013, the Governing Board will also expand the state-level studies by partnering with two states at grade 8. Again using a procedure that protects student confidentiality, secondary and postsecondary data for NAEP 8th grade test takers in the state samples in partner states will be linked to NAEP scores. These studies will examine the relationship between 8th grade NAEP scores and scores on state tests, future ACT scores, placement into remedial versus credit-bearing courses, and first-year college GPA.

Two states will be partners in these studies at grade 8: North Carolina and Tennessee.

Draft Research Questions for Longitudinal Statistical Relationships, Grade 8 NAEP:

1. What is the relationship between NAEP Reading and Math scores at grade 8 and state test scores at grade 4?
2. What are the average NAEP Reading and Math scores and the interquartile ranges (IQR) at grade 8 for students below the ACT benchmarks at grade 11/12? At or above the ACT benchmarks?
3. What are the average NAEP Reading and Math scores and the interquartile ranges (IQR) at grade 8 for students who are placed in remedial and non-remedial courses in college?
4. What are the average NAEP Reading and Math scores (and the IQR) at grade 8 for students who obtain a first-year college GPA of B- or above?
5. What is the relationship between grade 8 NAEP Reading and Math scores and grade 12 NAEP Reading and Math scores? (contingent on feasibility of sampling the same students in TN and NC)

May 2014 Update: The data sharing agreements are still in the process of being finalized. Kentucky will not be participating in the grade 8 longitudinal research.

Content Alignment Study of Grade 12 NAEP Reading and Mathematics and COMPASS

Content alignment studies are a foundation for the trail of evidence needed for establishing the validity of preparedness reporting, and are, therefore, considered a high priority in the Governing Board's Program of Preparedness Research. The alignment studies will inform the interpretations of preparedness research findings from statistical relationship studies and help to shape the statements that can be made about preparedness. Content alignment studies were recommended to evaluate the extent to which NAEP content overlaps with that of the other assessments to be used as indicators of preparedness in the research.

We plan to conduct an alignment study of grade 12 NAEP Reading and Mathematics and ACT COMPASS.

May 2014 Update: ACT has agreed to having a NAEP and COMPASS content alignment study performed by an independent third party. We are working on a procurement package and expect to release a Request for Proposals (RFP) in mid-June, with a target award date of mid-September.

Content Alignment Study of Grade 8 NAEP Reading and Mathematics and EXPLORE

Content alignment studies are a foundation for the trail of evidence needed for establishing the validity of preparedness reporting, and are, therefore, considered a high priority in the Governing Board's Program of Preparedness Research. The alignment studies will inform the interpretations of preparedness research findings from statistical relationship studies and help to shape the statements that can be made about preparedness. Content alignment studies were recommended to evaluate the extent to which NAEP content overlaps with that of the other assessments to be used as indicators of preparedness in the research.

We plan to conduct an alignment study of grade 8 NAEP Reading and Mathematics and ACT EXPLORE. Results from this content alignment study will be particularly important for interpreting the findings from the statistical linking studies of NAEP and EXPLORE.

May 2014 Update: ACT has agreed to having a NAEP and EXPLORE content alignment study performed by an independent third party. We are working on a procurement package and expect to release a Request for Proposals (RFP) in mid-June, with a target award date of mid-September.

OVERVIEW OF REFERENCED ASSESSMENTS

For additional background information, the following list presents a brief description of the assessments referenced in the phase two academic preparedness research studies. In each case, only the mathematics and reading portions of the assessments are the targets for analysis, although analyses with the composite scores may be conducted.

- ACT – The ACT assessment is a college admissions test used by colleges and universities to determine the level of knowledge and skills in applicant pools, including Reading, English, Mathematics, and Science tests. ACT has *College Readiness Standards* that connect reading or mathematics knowledge and skills and probabilities of a college course grade of "C" or higher (0.75) or "B" or higher (0.50) with particular score ranges on the ACT assessment.

- ACT EXPLORE – ACT EXPLORE assesses academic progress of eighth and ninth grade students. It is a component of the ACT College and Career Readiness System and includes assessments of English, Mathematics, Reading, and Science. ACT EXPLORE has *College Readiness Standards* that connect reading and mathematics knowledge and skills and probabilities of a college course grade of "C" or higher (0.75) or "B" or higher (0.50) by the time students graduate high school with particular score ranges on the EXPLORE assessment.

- ACT COMPASS – ACT Compass is a computer-adaptive college placement test. It is produced by ACT and includes assessments of Reading, Writing Skills, Writing Essay, Mathematics (Pre-algebra; Algebra; College Algebra; Geometry; Trigonometry), and English as a Second Language.

- SAT – The SAT reasoning test is a college admissions test produced by the College Board. It is used by colleges and universities to evaluate the knowledge and skills of applicant pools in critical reading, mathematics, and writing. The SAT has calculated preparedness benchmarks are defined as the SAT scores corresponding to a 0.65 probability of earning a first-year college grade-point average of 2.67 (B-) or better.

COLLEGE COURSE CONTENT ANALYSIS

Project Status Update
Contract ED-NAG- 12C-0003

The College Course Content Analysis (CCCA) study is one of a series of studies contributing to the National Assessment of Educational Progress (NAEP) Program of 12th Grade Preparedness Research conducted by the National Assessment Governing Board (NAGB). The purpose of the CCCA study is to identify a comprehensive list of the reading and mathematics knowledge, skills, and abilities (KSAs) that are pre-requisite to entry-level college mathematics courses and courses that require college level reading based on information from a representative sample of U.S. colleges. The Educational Policy Improvement Center (EPIC) is the contractor working for the Board to conduct this study.

Another goal of the CCCA study is to extend the work of the two previous preparedness studies—the Judgmental Standards Setting (JSS)[1] study, implemented in 2011 and the Job Training Program Curriculum (JTPC) study, implemented in 2012. The CCCA study is designed so the results can be compared to the JSS and JTPC studies, reporting on how this new information confirms or extends interpretations of those earlier studies. The design of the CCCA study is based on the JTPC study but with modifications based on the lessons learned.

May 2014 Update:
The final report was submitted on April 30, 2014. Board staff are conducting a final review of the report, before preparing it for dissemination on the Board's website.

1 National Assessment Governing Board. (2010). *Work Statement for Judgmental Standard Setting Workshops for the 2009 Grade 12 Reading and Mathematics National Assessment of Educational Progress to Reference Academic Preparedness for College Course Placement.* (Higher Education Solicitation number ED-R-10-0005).

COLLEGE COURSE CONTENT ANALYSIS
Final Report

Prepared by The Educational Policy Improvement Center

EXECUTIVE SUMMARY

The National Assessment Governing Board is an independent, bipartisan organization that sets policy for the National Assessment of Educational Progress (NAEP). The Governing Board established the NAEP Program of 12th Grade Preparedness Research to assess what NAEP can report on the academic preparedness of 12th grade students entering college and job training. The Governing Board commissioned the Educational Policy Improvement Center in October 2012 to conduct the College Course Content Analysis (CCCA) study, which used course artifacts to identify the prerequisite knowledge, skills, and abilities (KSAs) in mathematics and reading that are necessary for students to be prepared to qualify for entry-level, credit-bearing courses that satisfy general education requirements.

The CCCA study was designed to answer the following research questions:

- What are the prerequisite KSAs in mathematics and reading to qualify for entry-level, credit-bearing courses that satisfy general education requirements?
- How do these prerequisite KSAs compare with the 2009 and 2013 NAEP mathematics and reading frameworks and item pools?
- How do these prerequisite KSAs compare with previous NAEP preparedness research, i.e., descriptions of minimal academic preparedness requirements produced in the Judgmental Standard Setting (JSS) research?
- How can these prerequisites inform future NAEP preparedness research?

Study Method

Phase 1: Establishing the foundation for the study

The initial phase included selecting a representative sample of institutions, collecting course artifacts, organizing the artifacts into course packets, and recruiting content reviewers and NAEP content experts.

Selecting a representative sample. One hundred fifty-one (151) institutions contributed 184 course packets, which included a syllabus, textbook, and class assignment or assessment (hereafter referred to as course artifacts). Half of the 184 packets addressed mathematics courses, and the other half addressed courses with substantial reading demands. To generate a nationally representative sample of institutions, the target sample was stratified by

- program type (two-year and four-year)
- size (small < 4,999 students; medium 5,000–9,999; or large > 10,000)
- geographic region (East or West)
- institutional control (private or public)

Collecting course artifacts and assembling course packets. The mathematics course titles included precalculus/calculus, college algebra, finite mathematics, and statistics. The courses

with substantial reading demands included English literature, psychology, U.S. government, and U.S. history. Instructors who taught the included courses at the sampled institutions provided the course artifacts and verified that these artifacts met inclusion criteria. The course artifacts were compiled into course packets for the content analysis activities.

Recruiting content reviewers and NAEP content experts. Highly qualified college-level content experts, 16 in mathematics and 16 in reading, were recruited to conduct course content reviews. The content reviewers were postsecondary instructors who taught relevant introductory college courses. Another smaller group of doctoral-level content experts in mathematics and reading with specialized expertise and experience with the NAEP frameworks were engaged to provide guidance for the use of the NAEP frameworks within the overall study design.

Phase 2: Conducting content analysis activities

The next phase included training and qualifying the content experts to conduct two content reviews, independent and group, as part of the convergent consensus process; summarizing the data into detailed content maps of the KSAs; and completing a generalizability study, as a measure of the consistency of the coding process.

Training and qualifying reviewers. The training began by inviting content reviewers to conduct a "holistic" review of the course packets prior to being introduced to the NAEP framework. This provided reviewers with the opportunity to identify prerequisite KSAs independent of NAEP frameworks. Subsequent reviewer trainings were conducted to address the NAEP framework, the coding scheme, decision rules, and review processes. Training also included trial reviews of a subset of the course packets and an individualized feedback for each reviewer. Only reviewers who demonstrated an understanding of the materials and processes qualified to participate in the content reviews.

Independent and group content reviews. After training, the research project moved into the operational components of the study by facilitating two content expert reviews of the course packets: one independent review where reviewers individually applied a comprehensive coding scheme to the course packets, and one group review where the reviewers discussed their points of individual agreement and disagreement. Content reviewers also used a set of decision rules to support their identification of the prerequisite KSAs within each course packet. Decision rules are clarifying guidelines on pre-identified potential areas of ambiguity.

Summarizing the data and conducting reliability checks. The data collected from these reviews were summarized into content maps of the prerequisite KSAs for each course. Some prerequisite KSAs were not included within the NAEP framework objectives but were evident in the course packets. A generalizability study was conducted to evaluate coding reliability.

Phase 3: Conducting NAEP expert reviews

During the final phase, the NAEP experts used the content maps to develop narrative descriptions of the prerequisite KSAs necessary for students to be prepared for entry, without remediation, into credit-bearing entry-level courses. They then compared their narrative descriptions to NAEP item pools, borderline performance descriptions generated in previous research, and other content summaries captured in the Judgmental Standard Setting and Job Training Program Content studies. In these comparisons, NAEP experts analyzed

correspondence between the CCCA KSAs and content assessed in the NAEP. Through a facilitated discussion, they provided insight via their NAEP framework expertise.

Study Process Results

The validity of the results is based on institutional representativeness, artifact sufficiency, and coding reliability. Analysis indicates that the sample of institutions submitting artifacts for courses with substantial reading demands is generally representative of the population of institutions. Institutions submitting artifacts for mathematics courses underrepresent small institutions and private institutions and overrepresent public institutions and large institutions. Post hoc analyses suggest that deviations from representativeness do not pose a serious threat to validity. Content reviewers deemed the artifacts sufficient to allow them to focus on what students are expected to be able to learn, what students are expected to be able to do to demonstrate learning, and the kinds of content the students will reference in order to learn. Results from the generalizability study indicate the CCCA findings are reliable, i.e., rating was consistent within course packets during the independent and group reviews.

Prerequisite KSA Findings and Conclusions

Prerequisite KSAs

The results from the study indicate that most of the prerequisite KSAs for both mathematics courses and courses that require extensive college-level reading are reflected in the NAEP frameworks. A KSA was considered prerequisite if a student is either expected or required to possess this knowledge, skill, or ability to be prepared for entry into the course. The prerequisite KSAs were mapped to the NAEP frameworks or they were included as non-NAEP additional KSAs. Any KSA identified in 75% or more of course packets within a subject area was considered to be a common prerequisite. Prerequisite KSAs found less frequently within course packets demonstrated the range of prerequisites present within the sample of course packets, and, by extension, likely to be present within courses taught across U.S. institutions.

Reading prerequisite KSAs. Some of the reading KSAs that are prerequisite to entering entry-level, credit-bearing college courses have substantial informational text reading demands, and some of the reading KSAs demand engagement with literary texts. KSAs common to both included the ability to locate or recall textually explicit information within and across texts and the ability to take different perspectives in relation to a text. Few differences were found in prerequisite KSAs among courses with substantial informational text reading demands. The key distinction between the courses with substantial informational text versus literary text reading demands was in the KSAs related to the ability to make complex inferences within and across texts; related KSAs were more commonly found to be prerequisite in courses dependent on informational texts rather than literary texts. However, courses within English literature were more likely than those with substantial informational text reading demands to have prerequisite KSAs relating to both informational and literary texts.

Mathematics prerequisite KSAs. The mathematics KSAs were mostly specific to a course title, i.e., there was not a single set of prerequisite KSAs that covered all entry-level, credit-bearing courses. The majority of prerequisite KSAs were represented in objectives within the NAEP framework's *Numbers and Operations* subject area. Prerequisite KSAs for precalculus/calculus

and college algebra were notably different from those for finite mathematics and statistics—they were represented by the objectives within the *Variables, Expressions, and Operations* and *Equations and Inequalities* subtopics of the *Algebra* subject area of the NAEP framework, whereas there were fewer prerequisite KSAs from these subject areas for statistics and finite mathematics. Additional KSAs that were not found within the NAEP framework were identified as prerequisites. These related to the use of technology (e.g., calculators, online resources) and the ability to read and communicate in various modes about mathematics.

Comparison of Prerequisite KSAs to 2009 and 2013 NAEP Mathematics and Reading Frameworks and Item Pools

Reading prerequisite KSA comparison. The NAEP reading framework aligned well with the prerequisite KSAs identified in this study, especially in the cognitive domains of *Locate/Recall* and *Integrate/ Interpret*. The prerequisite KSAs focused less on the ability to critique or evaluate, which may be what is covered within a college course and therefore not a prerequisite to entry into the course. The reading item pools represent fairly well the prerequisite KSAs noted in this study. The nature of the text excerpts in the 2009 and 2013 NAEP reading item pools did not, however, cover the range of texts that would be required in order to assess all of the prerequisite KSAs.

Mathematics prerequisite KSA comparison. The "big picture" of the mathematics frameworks was transparent to the NAEP experts when they envisioned the 4th, 8th, and 12th grade objectives side by side. After accounting for the pieces of NAEP objectives that were not considered relevant, some prerequisite KSAs are found in the 8th grade NAEP framework. Some additional KSAs identified by content reviewers were found by the NAEP experts to be implied within the NAEP frameworks and explicit in the frameworks for grades not referenced in this study. The CCCA prerequisite KSAs focused more heavily on application of mathematics (e.g., simulations, addressing real-world problems) than is evident in the item pools. The NAEP experts found the items to be skill-based and procedural in nature rather than applied. The CCCA content focused on application problems, which do match the NAEP framework, but the items do not reflect the same level of focus. Another main difference between NAEP mathematics and CCCA prerequisite KSAs is that many of the CCCA sampled courses did not have measurement or geometry objectives as prerequisite KSAs.

Comparison of Prerequisite KSAs to Previous NAEP Preparedness Research

Reading prerequisite KSA comparison. The CCCA prerequisite KSAs describe a depth of understanding and a level of cognitive demand that are beyond those described by the JSS minimum preparedness requirements. The NAEP experts portrayed the two sets of descriptions as concentric circles, with the JSS circle completely inside the CCCA circle. They concluded that the CCCA study, with the inclusion of the courses that relied heavily on informational texts as well as those relying primarily on literary texts, resulted in a narrative description that was broader and deeper than the JSS borderline performance descriptions (BPD).

The reading prerequisites identified in the CCCA study are more numerous and constitute more KSAs related to higher-order thinking skills than those described by the JSS descriptions of minimal preparedness. The specific NAEP objectives necessary for minimal preparedness from JSS were all subsumed by the prerequisites evident in CCCA course artifacts. For example, the JSS description included *offer evidence in support of*, but not judging, evaluating, and critiquing,

which is subsumed by *offering evidence to support a claim*.

Mathematics prerequisite KSA comparison. The prerequisites identified for precalculus/calculus and college algebra are similar to the JSS description of the KSAs necessary for minimal preparedness for entry-level college mathematics courses. The prerequisites for finite mathematics and statistics, however, require fewer and less cognitively complex KSAs than are described by the JSS description.

Given the differences in evidence, process, and unit of analysis, dissimilarities between study findings were expected. The CCCA study identified prerequisite KSAs from the evidence contained in college course packets of artifacts. As a standard-setting process, the JSS study identified what minimally prepared college students need to know and be able to do to succeed in entry-level college mathematics courses and courses with substantial reading demands.

Informing Future NAEP Preparedness Research

Studies focusing on the prerequisites from a broader range of entry-level college courses might be useful at identifying differences in prerequisites for courses that use a variety of texts and text types. Of particular interest is the inclusion of courses that contain both significant mathematics and reading demands (e.g., economics) and the examination of the prerequisites related to both content areas.

Exploring prerequisites related to writing, graphical representations of information, and technology use could complement the content analyses of reading and mathematics KSAs that have been conducted on college and job training courses.

EVALUATING READING AND MATHEMATICS FRAMEWORKS AND ITEM POOLS AS MEASURES OF ACADEMIC PREPAREDNESS FOR COLLEGE AND JOB TRAINING

Project Status Update
Contract ED-NAG-13C-0001

The National Assessment Governing Board contracted with the Human Resources Research Organization (HumRRO) in June 2013 to conduct three tasks related to research on 12th grade preparedness:

1. **Evaluation of the Alignment of Grade 8 and Grade 12 NAEP to an Established Measure of Job Preparedness:** In its June 2009 report, *Making New Links: 12th Grade and Beyond,* the Technical Panel on 12th Grade Preparedness Research recommended that content alignment studies be conducted to examine the structure and content of various assessments relative to NAEP. The purpose of such content alignment would be to determine whether the scores on NAEP and the other assessments convey similar meaning in terms of the knowledge and skills of examinees. In fact, the panel specifically recommended that content alignment studies be conducted between NAEP and WorkKeys to determine the correspondence between the content domain assessed by NAEP and that of WorkKeys. If the alignment is relatively high, or even moderately high in some cases, then statistical relations between NAEP and WorkKeys may allow for the interpretation of NAEP results in terms of how WorkKeys would typically be interpreted. Using WorkKeys as a measure of job training preparedness allows the comparison of findings from this research to findings from previous content alignment studies with WorkKeys. This would provide a cross-validity check with NAEP grade 12 and also expand the content alignment study by using NAEP grade 8 as well. This study will extend prior analysis of the relation of NAEP to WorkKeys by including the NAEP grade 8 assessments and by expanding the method for assessing content alignment. The study method will follow the Governing Board content alignment design document for preparedness research studies, with some modifications. The two-pronged approach includes alignment of: (a) WorkKeys to the NAEP frameworks, and (b) NAEP items to the framework from which WorkKeys was developed.

2. **O*NET Linkage Study:** This study: a) identified relevant linkages between the National Assessment of Educational Progress (NAEP) and training performance requirements for selected occupations, and b) compared the levels of knowledge, skills, and abilities (KSAs) required for the relevant NAEP content to the levels of KSAs required for the relevant job training content.

   For this study, tasks (i.e., performance requirements) for each occupation were extracted from O*NET. The O*NET, or Occupational Information Network, is the U.S. Department of Labor's occupational information database. The O*NET contains standardized descriptions of 974 occupations, including the five occupations that are the focus of the National Assessment Governing Board's (Governing Board) program of research on job preparedness. Because the O*NET descriptors provide a "common language" for describing similarities and differences across occupations, it is a very useful resource for this study. Occupational experts from each of the target occupations reviewed the O*NET task lists for their

appropriateness to job training. This review was necessary because the O*NET tasks describe *job* performance requirements, but not *training* performance requirements, and the focus of the Governing Board's research is preparedness for *job training*. Based on the feedback from the occupational experts, edits were made to the O*NET task lists to ensure their applicability to job training. Next, occupational experts used these lists to identify NAEP content that is relevant ("linked") to training performance requirements. The occupational experts also identified the training performance requirements that are relevant ("linked") to NAEP content. Irrelevant content was removed from further consideration. Finally, trained project analysts used academically-relevant KSAs from O*NET to systematically rate the levels of KSAs needed for the relevant NAEP content and the levels of KSAs needed for the relevant job training content. Disconnects between the levels of KSAs needed for NAEP and the levels needed for job training were flagged for discussion.

An executive summary of this report follows this Project Status overview.

3. **Technical Advisory Panel (TAP) Symposium:** HumRRO assembled a technical advisory panel (TAP) of five experts in educational measurement and five experts in industrial-organizational (I-O) psychology to review extant research and to generate ideas for commissioned papers on preparedness. The TAP met in Washington D.C. in late October 2013. This brainstorming session included presentations by Governing Board and HumRRO staff describing findings from previous studies and descriptions of other studies currently underway, followed by an open discussion of issues and possible additional areas of investigation. Each panelist was asked to use this information to propose a paper that he/she could develop. TAP members submitted nine proposals from which Governing Board staff commissioned five papers. Panelists have several months to develop the papers. The TAP will reconvene in a late summer 2014 symposium during which authors will present their papers and the entire panel will discuss implications for preparedness research. HumRRO will produce a proceedings document summarizing the commissioned papers and discussion. (A list of TAP members is included on the next page.)

In addition, HumRRO will produce a comprehensive project report at the conclusion of the contract in December 2014.

May 2014 Update:

Evaluation of Alignment of Grade 8 and 12 NAEP to an Established Measure of Job Preparedness: HumRRO has arranged with ACT to use operational WorkKeys forms in alignment workshops, and recruitment is currently underway for educators to serve on alignment workshops. The workshop dates have been set in June and July 2014.

O*NET Linkage: HumRRO staff analyzed project analysts' ratings (level of O*NET knowledge, skills, and abilities (KSAs) needed for NAEP and level of O*NET KSAs needed for job training) and submitted the final report to the Governing Board in early April 2014.

TAP Symposium: Governing Board staff reviewed proposals submitted by TAP panelists and commissioned five (5) papers to be completed by the panelists. Drafts are due in July 2014; final versions will be presented at the second TAP meeting in August 2014.

Technical Advisory Panel (TAP) Members

John Campbell
Professor of Psychology
University of Minnesota
(Member, NAGB Technical Panel on 12th Grade Preparedness Research, 2007-2008)

Michael Campion
Herman C. Krannert
Professor of Management
Purdue University

Gregory Cizek
Professor of Educational Measurement and Evaluation
University of North Carolina at Chapel Hill

Brian Gong
Executive Director of Center for Assessment
National Center for the Improvement of Educational Assessment, Inc.

Ronald Hambleton
Distinguished University Professor, Educational
Policy, Research, & Administration
Executive Director, Center for Educational Assessment
University of Massachusetts at Amherst

Suzanne Lane
Professor, Research Methodology
University of Pittsburgh School of Education

Barbara Plake
University Distinguished Professor, Emeritus
University of Nebraska-Lincoln

Ann Marie Ryan
Professor of Psychology
Michigan State University

Nancy Tippins
Senior Vice President
CEB Valtera

Comparisons Between NAEP and O*NET on Academic Preparedness for Job Training for Five Target Occupations

Prepared by Human Resources Research Organization

EXECUTIVE SUMMARY

This report describes the results from a study that a) identified relevant linkages between the National Assessment of Educational Progress (NAEP) in reading and mathematics, and training performance requirements for selected occupations, and b) compared the levels of knowledge, skills, and abilities (KSAs) required for the relevant NAEP content to the levels of KSAs required for the relevant job training content. The KSAs included in the current study come from the O*NET. The O*NET, or Occupational Information Network, is the U.S. Department of Labor's occupational information database. The O*NET contains standardized descriptions of 974 occupations, including the five occupations that are the focus of the National Assessment Governing Board's (Governing Board) program of research on academic preparedness for job training programs. The purpose of this program of research is collect evidence to determine the feasibility of using NAEP to report on the academic preparedness of U.S. 12th grade students for entry into job training. The five target occupations selected by the Governing Board for this program of research are: Automotive Master Technicians, Computer Support Specialists, Heating, Ventilation and Air Conditioning Technicians (HVAC), Licensed Practical Nurses (LPNs), and Pharmacy Technicians. Because the O*NET descriptors provide a "common language" for describing similarities and differences across occupations, it is a very useful resource for the present research.

For this study, tasks (i.e., performance requirements) for each occupation were extracted from O*NET. Occupational experts from each of the target occupations reviewed the O*NET task lists for their appropriateness to job training. This review was necessary because the O*NET tasks describe job performance requirements, but not training performance requirements, and the focus of the Governing Board's research is preparedness for job training. Based on the feedback from the occupational experts, edits were made to the O*NET task lists to ensure their applicability to job training. These lists of training performance requirements served as common content of job training programs. Next, occupational experts used these lists to identify NAEP reading and mathematics content that is relevant ("linked") to training performance requirements. The occupational experts also identified the training performance requirements that are relevant ("linked") to NAEP content. Irrelevant content was removed from further consideration. Finally, trained project analysts used academically-relevant KSAs from O*NET to systematically rate the levels of KSAs needed for the relevant NAEP content and the levels of KSAs needed for the relevant job training content. Disconnects between the levels of KSAs needed for NAEP reading and mathematics and the levels needed for job training were flagged for discussion.

An overview of the findings is as follows:

- The range of reading and mathematics skills required by NAEP (both grade 8 and grade 12) is broader than the range of reading and mathematics skills required by job training. This was

demonstrated by the finding that considerably more content on NAEP was rated as irrelevant to job training than was job training content rated as irrelevant to NAEP.

- The NAEP reading objectives most relevant to job training content are the objectives associated with the Locate/Recall cognitive target for NAEP informational reading.

- The NAEP reading objectives that were least relevant to job training content were the objectives associated with the Critique/Evaluate cognitive target.

- The NAEP mathematics objectives most relevant to job training content were the objectives associated with the Numbers Sense and Operations and Measurement (except for Computer Support Specialists) content areas. This was true for both grade 8 and grade 12 NAEP.

- The NAEP mathematics objectives that were least relevant to job training content were the objectives associated with Geometry (except for HVAC) and Algebra (except for LPNs). This was true for both grade 8 and grade 12 NAEP.

- The percentage of the mathematics objectives linked to occupations decreased considerably from grade 8 to grade 12, indicating that as the complexity of the objectives increased from grade 8 to grade 12 their relevance to job training decreased.

- Disconnects were found between the levels of KSAs required for proficiency on NAEP and the levels of KSAs required for entry into job training such that higher levels of the KSAs were required for NAEP than for job training. The largest disconnects occurred between grade 12 NAEP mathematics and job training. Disconnects also occurred between grade 12 reading and job training. The disconnects in required levels of KSAs tended to be smaller when comparing grade 8 content to job training content, particularly for grade 8 reading, which demonstrated several "matches" with KSA levels for training content (most notably with Written Comprehension, which was rated as "Moderate" for both the NAEP grade 8 reading content and for the job training content across all occupations).

The above set of findings call into question the validity of inferences that can be made about using NAEP to report on the preparedness of U.S. 12th grade students for entry into job training. Based on the findings from this study in conjunction with converging evidence from prior studies (ACT, 2010a; 2010b, WestEd & Measured Progress, 2011; 2012; WestEd & Educational Policy Improvement Center, 2013), we offer the following recommendations for the Governing Board's consideration:

- Given that there is converging evidence across studies that the Number Properties and Operations content area for mathematics and the Locate/Recall cognitive target from NAEP informational reading are most relevant to job training, consider the possibility of using subscores from these content areas to report on students' preparedness for job training.

- Given the greater correspondence between grade 8 content and job training content in reading and mathematics, consider the possibility of administering the grade 8 assessments to

12th grade students to make determinations about their academic preparedness for entry into job training.

- Consider the possibility of updating the working definition of job preparedness to include trainee outcomes, such as trainee performance in job training. Actual performance in job training is at a level that is somewhat beyond “just qualified” for placement into job training. Including training outcomes in the working definition of job preparedness might potentially lead to evidence that is more supportive of grade 12 NAEP as an indicator of job preparedness. Furthermore, including training outcomes as elements of the working definition of job preparedness would expand opportunities for future research investigations.

Update on Development of Technology and Engineering Literacy (TEL) Achievement Levels Descriptions (ALDs)

The first step in setting achievement levels for the 2014 grade 8 NAEP Technology and Engineering Literacy (TEL) Assessment is to finalize the achievement levels descriptions (ALDs) for grade 8. There are specific requirements for achievement level descriptions for NAEP. The achievement levels describe what students *should know and be able to do* to meet the requirements at each of three levels of achievement (Basic, Proficient, and Advanced). These TEL ALDs must be aligned with the content of the TEL Framework and the policy descriptions of each achievement level.

The preliminary achievement levels descriptions for each assessment area (Technology and Society; Design and Systems; and Information and Communication Technology) were developed as part of the Technology and Engineering Literacy Framework development project[1]. The Governing Board requires expert evaluation of these descriptions in order to develop final descriptions for use in the achievement levels setting process. The final descriptions will be for the overall Technology and Engineering Literacy Assessment rather than separately by assessment area.

A statement of objectives was developed, and a Request for Quotations was issued in early February. A contract was awarded to WestEd on April 4, 2014. The project staff works within WestEd's Science, Technology, Engineering, and Mathematics (STEM) program and has an extensive understanding of curriculum, instruction, and assessment issues in these content areas. Project staff have led framework development projects, conducted studies of educational technology programs, and developed assessments on the use of educational technology. Most significantly, they managed the development of the 2014 NAEP TEL Framework.

On May 1 – 2, experts in Design and Systems, Information and Communication Technology, and Technology and Society were convened to review, and revise as necessary, the descriptions included in the 2014 TEL Framework. Following the refinement of the TEL ALDs, WestEd will seek public comment from TEL education stakeholders and other key individuals and organizations from approximately May 12 – 30. The proposed TEL ALDs will be posted online at http://naeptelaldreview.com. The content experts will then reconvene a virtual meeting to review the public comment and develop finalized descriptions. These descriptions will be presented as recommendations to the Governing Board for approval at the July 31 – August 2, 2014 meeting.

At the May 2014 meeting, COSDAM will hear about highlights and challenges from the May 1 – 2 panel meeting, along with expected next steps for finalizing the TEL ALDs.

[1] http://www.nagb.org/publications/frameworks/technology/2014-technology-framework.html

Update on Evaluation of NAEP Achievement Levels Procurement

Objective To receive a brief informational update from NCES on the current status of the procurement being planned to evaluate NAEP achievement levels. Ongoing updates will be provided at each COSDAM meeting.

Background

The NAEP legislation states:

> The achievement levels shall be used on a trial basis until the Commissioner for Education Statistics determines, as a result of an evaluation under subsection (f), that such levels are reasonable, valid, and informative to the public.

In providing further detail, the aforementioned subsection (f) outlines:

(1) REVIEW-

A. IN GENERAL- The Secretary shall provide for continuing review of any assessment authorized under this section, and student achievement levels, by one or more professional assessment evaluation organizations.

B. ISSUES ADDRESSED- Such continuing review shall address--

(i) whether any authorized assessment is properly administered, produces high quality data that are valid and reliable, is consistent with relevant widely accepted professional assessment standards, and produces data on student achievement that are not otherwise available to the State (other than data comparing participating States to each other and the Nation);

(ii) whether student achievement levels are reasonable, valid, reliable, and informative to the public;-

(iii) whether any authorized assessment is being administered as a random sample and is reporting the trends in academic achievement in a valid and reliable manner in the subject areas being assessed;

(iv) whether any of the test questions are biased, as described in section 302(e)(4); and

(v) whether the appropriate authorized assessments are measuring, consistent with this section, reading ability and mathematical knowledge.

(2) REPORT- The Secretary shall report to the Committee on Education and the Workforce of the House of Representatives and the Committee on Health, Education, Labor, and Pensions of the Senate, the President, and the Nation on the findings and recommendations of such reviews.

(3) USE OF FINDINGS AND RECOMMENDATIONS- The Commissioner for Education Statistics and the National Assessment Governing Board shall consider the findings and recommendations of such reviews in designing the competition to select the organization, or organizations, through which the Commissioner for Education Statistics carries out the National Assessment.

Responsively, a procurement has been planned to administer an evaluation of NAEP achievement levels. The last update COSDAM reviewed on this topic was in February 2014.

In the following brief written update, NCES provides the Committee with a summary of the status of this procurement.

Evaluation of NAEP Achievement Levels

The National Center for Education Evaluation and Regional Assistance (NCEE), part of the Institute for Education Sciences (IES), will administer the Evaluation of the NAEP Achievement Levels. NCEE and the Department of Education's Contracts and Acquisitions Management (CAM) office began this procurement during fiscal year 2014. A Request for Comments (RFC) was released in late February, and a presolicitation announcement (which includes a draft statement of objectives) was released in late April (https://www.fbo.gov/index?s=opportunity&mode=form&id=d1593f47cbf554cae1fcd2f243610607&tab=core&_cview=0).

According to the draft statement of objectives, a report (produced within 18 months of contract award), *"…should provide sufficient information upon which the Commissioner of NCES can determine if the trial designation of the NAEP reading and mathematics achievement levels at grades 4, 8, and 12 should be removed or whether the trial designation should be continued"* (page 4). The draft statement of objectives also includes a 6-month option to extend the contract to 24 months; if this option is exercised, the contractor would plan and conduct dissemination events to communicate the conclusions of the final report to various groups of stakeholders.

The solicitation is expected to be released in early May, and the contract is expected to be awarded by the end of August.

Technology and Engineering Literacy (TEL) Achievement Levels Setting

A Request for Proposals (RFP) was issued on March 24, 2014 to set achievement levels for the 2014 grade 8 NAEP Technology and Engineering Literacy (TEL) Assessment. Proposals are due on May 9, 2014, with a target award date of June 30, 2014.

The Statement of Work (SOW) for the TEL achievement levels setting (ALS) includes the requirement for a pilot study and an operational achievement levels setting study. The Technical Advisory Committee on Standard Setting (TACSS) for the TEL ALS contract will need to include at least five members, one of whom must be a state assessment director or coordinator. The methodology for setting achievement levels for Technology and Engineering Literacy is not specified in the SOW, to allow for flexibility and innovation by the bidders. However, computers will need to be incorporated into the achievement levels setting process.

At the August 2014 meeting, COSDAM will receive an update on the status of the TEL achievement levels setting contract award, including project milestones.