2010 NAEP System Clearance 1st OMB Passback Responses 02-03-2010

1. **NAEP is in the midst of a 5 year (2008-2012) assessment cycle. Please provide an updated schedule of assessments and any special studies contemplated for the remainder of that cycle.**

Assessment and Special Study Schedule for 2011 and 2012 Assessments

The 2011 data collection consists of the following:

ASSESSMENTS	DATES
National, state, and urban district assessment in **reading** at grades 4 and 8 (including the NIES study)	Jan-Mar 2011
National, state, and urban district assessment in **mathematics** at grades 4 and 8 (including the NIES study)	Jan-Mar 2011
National assessments in **writing** at grades 8 and 12	Jan-Mar 2011
National and state assessment in **science** at grade 8	Jan-Mar 2011
Pilot assessments for 2013 **reading** and **mathematics** at grades 4 and 8	Jan-Mar 2011
Pilot assessment for 2012 economics at grade 12	Jan-Mar 2011
SPECIAL STUDIES	DATES
NAEP-TIMSS (Trends in International Mathematics and Science Study) linking study	Jan-Mar 2011
Multi-stage testing (MST) special study	Jan-Mar 2011
KaSA (Knowledge and Skills Appropriate) Study	Jan-Mar 2011
Texas Survey of 1st year college students	August 2010

The 2012 data collection consists of the following:

ASSESSMENTS	DATES
National Long-Term Trend (LTT) assessments in reading and mathematics at ages 9, 13, and 17	Oct 2011-May 2012
National assessment in economics at grade 12	Jan-Mar 2012
Pilot assessments for 2016 LTT in reading and mathematics at ages 9, 13, and 17	Jan-Mar 2012
Pilot assessment in science, including ICT and HOTs, at grades 4, 8,and 12	Jan-Mar 2012
Pilot assessments for 2013 reading and mathematics at grade 12	Jan-Mar 2012
SPECIAL STUDIES	DATES
Study in technological literacy at grade 8	April 2012

2. Please also provide key dates and activities for the kick off of planning for the next five-year cycle.

NAEP Key Dates and Activities

ACTIVITY	DATES	PURPOSE
Alliance Operations Kick-off Planning Meeting	April of each year	Plan logistics for the upcoming assessment cycle (for example, in April 2010, plan for the 2011 assessment)
Design Summit Meeting	December of each year	Plan and review design decisions and implications for the next three assessment cycles (for example, in December 2010, plan for the 2011, 2012, and 2013 assessments)
National Assessment Governing Board Meetings	Quarterly	Assessment schedules are reviewed and occasionally revised
NAEP Background Variable Standing Committee	Spring of each year	Review and decisions regarding background variables
NAEP 2012-2017 contract award	September 2012	Award contract for the design and administration of the 2013 – 2017 NAEP assessments, including details of the assessment components

2010 NAEP System Clearance 2nd OMB Passback Responses 03-22-2010

1. Thank you for providing the Assessment and Special Study Schedule for 2011 and 2012 Assessments. Please provide, for each of the five special studies listed, the following information:

   - Purpose
   - Origin (i.e., specifically from where is the interest or need for this study emanating)
   - Plans beyond the special study (i.e., what type of changes or new activities may result and in what time frame)
   - Cost of the special study and rough estimates of cost for the change or new activity it will inform

   NAEP-TIMSS Linking Study

   - *Purpose*: To compare states to international countries and provide international benchmarks for the states, in math and science at grade 8.

- *Origin*: Increased political interest in international benchmarks (both from the federal- and state-level) and the congruence of the two assessment schedules created the ideal opportunity to perform this study.
- *Plans beyond special study*: Plans to potentially provide state-level TIMSS estimates will be developed depending on the results of the study.
- *Cost*: Approximately $5M

Multi-stage testing (MST) Study

- *Purpose*: To determine if adaptive testing could yield lower standard errors and better measurement precision in NAEP, particularly for low- and high-performing groups of students.
- *Origin*: The expansion of NAEP into lower-performing jurisdictions and groups of students over recent years (e.g., Puerto Rico, urban districts, and SD and ELL students) has suggested that the results may have lower measurement precision for some of these jurisdictions and groups. Adaptive testing would target the test to the individual's proficiency, thus yielding more precise estimates. In addition, the National Assessment Governing Board's Ad Hoc Committees on NAEP Testing and Reporting of Students with Disabilities and English Language Learners have suggested adaptive testing as a means to increase inclusion of these special populations.
- *Plans beyond special study*: If the study results indicate that adaptive testing is feasible and yields more precise estimates in NAEP, additional studies and possible changes to the NAEP operational procedures could be implemented, such as making the entire assessment adaptive. However, until the results are analyzed, no plans will be made.
- *Cost*: Approximately $3.1M

Knowledge and Skills Appropriate (KaSA) Study

- *Purpose*: To determine if the administration of blocks of items that target the knowledge and skills at the lower-end of the performance distribution would yield lower standard errors and better measurement precision for low-performing groups of students.
- *Origin*: The expansion of NAEP into lower-performing jurisdictions and groups of students over recent years (e.g., Puerto Rico, urban districts, and SD and ELL students) has suggested that the results may have lower measurement precision for these jurisdictions and groups. In order to potentially increase the measurement precision, items and blocks of items could be written to target students at the lower-end of the distribution, thus providing more information about what these students know and can do, resulting in improved measurement.
- *Plans beyond special study*: If the study results indicate that blocks of items that target the knowledge and skills at the lower-end of the performance distribution yield more precise estimates for certain jurisdictions and/or groups of students, then these blocks could be included in the operational item pool, thus expanding the universe of items that are assessed. In addition, if both the MST and KaSA studies yield favorable results, the

KaSA blocks could be used as part of an adaptive test. However, until the results are analyzed, no plans will be made.

- *Cost*: Approximately $400,000

Texas Survey of 1st year college students

- *Purpose*: To gather information regarding college freshmen performance on NAEP in 12th grade reading and mathematics.
- *Origin*: This study is one of approximately 20 studies planned or underway by the National Assessment Governing Board to enable NAEP to report on the academic preparedness of 12th grade students for entry level college credit coursework. The other studies include investigations to determine the content alignment and statistical relationship between NAEP and college admissions and placement tests, such as the SAT, ACT, Accuplacer, and COMPASS.
- *Plans beyond special study*: None, to date.
- *Cost*: TBD

Study in technological literacy at grade 8

- *Purpose*: To obtain information regarding the computer system, testing platform, and select items from a limited number of students prior to the technological literacy pilot in 2013.
- *Origin*: This study would be done as part of the NAEP development cycle for innovative and new computer-based assessments.
- *Plans beyond special study*: In 2013, all of the potential items will be pilot tested. In 2014, the technological literacy probe assessment will be conducted.
- *Cost*: TBD

2. OMB has been concerned that in many previous submissions new questionnaire items had not been cognitively tested prior to planned inclusion in a pilot test. At times the NAEP staff have indicated that the schedule did not permit such testing. This type of testing is required by OMB standards. OMB has worked with other NCES program areas to use NCES's generic clearance package to permit qualitative testing with a short lead time, so is eager to identify a plan to avoid these problems in the next NAEP cycle. Please describe how the NAEP program can meet OMB standards in the future using the NCES generic clearance or other processes.

New item development for non-cognitive (i.e., questionnaire) items most often involves less than 9 students to test out new questions and, thus, does not require OMB clearance. For upcoming pilots, cognitive labs will be performed and NAEP will inform OMB of the cognitive labs prior to their occurring. If the number of students used to tryout the questions is greater than 9, NAEP will first obtain OMB approval, likely utilizing NCES' generic clearance package.

For 2011, no new non-cognitive questions will be pilot tested. The economics pilot will use existing items from prior assessments. The writing pilot questions used in 2010 did employ cognitive lab testing prior to usage in the pilot.

3. Please clarify how the burden estimates for the items in the burden table were developed and validated.

The burden estimates are a function of the estimated time to complete each component and the estimated number of respondents.

Estimated time to complete each component

Currently, NAEP student background questions require 15 minutes to complete. This is a combined total for both the core and subject-specific questions. Previous cognitive laboratories with students and examination of the response rates have indicated that 15 minutes is appropriate for students to complete the questionnaire. Students are only allocated 15 minutes so that if they do not complete the questionnaire in that timeframe, they leave the remainder of the questions blank.

The grade 4 teacher burden is now estimated to be 30 minutes rather than the 20 minutes used in prior submittals. Feedback from the administration has indicated that the 4th grade teacher survey requires 30 minutes due to each teacher responding to questions regarding multiple subjects. The 8th grade teacher questionnaires all require 20 minutes each for completion, based on prior years' data. Teacher questionnaires are distributed to teachers at the pre-assessment visit. They are asked to complete the questionnaires (either online or in hardcopy) so that they can be collected when the administrators return for the actual assessment (approximately two weeks later). The school questionnaire burden is estimated to be 30 minutes for completion, based on prior years' data.

Another component of the school burden is the pre-assessment visit and e-filing activities, which require school personnel time. The pre-assessment visit requires one hour, based on reports from the assessment administrators who conduct the pre-assessment visit. The e-filing burden is also estimated at one hour, based on feedback received from the school staff who complete the process.

Starting in 2010, in an effort to reduce burden, the SD and ELL data collection process has been revamped. Instead of requiring individual questionnaire booklets for each identified SD and /or ELL student, a worksheet allowing for the collection of multiple students (up to 10 per worksheet) has been developed. Thus, the burden for the school personnel to complete the SD and ELL worksheets is based on the number of students who will be identified as SD and ELL, and is decreased from previous assessment years. The average time to complete the worksheet per student is 10 minutes, based on feedback from the school personnel.

As part of the annual assessment process, a debriefing is held with school staff and the assessment administrators. One component discussed as part of the debriefing is the time needed to respond

to the teacher and school questionnaires. Improvements in the next years' assessment may be made as a result of that information. In addition, the burden estimates are reconsidered and potentially revised, based on the information collected during the debriefing. For example, based on the information collected during such debriefings, the burden estimate for 4th grade teachers has been increased to 30 minutes.

Estimated number of respondents

The number of students for each year's assessments is taken from the annual NAEP Design Plans. The design plans convey the scope of the assessment components based on the Governing Board assessment schedules and NCES plans. The design plans are working documents and, as such, they will change when subjects are added or deleted to a specific year's assessment schedule. The Governing Board establishes an overall assessment schedule for approximately the next decade and may decide to add or drop subjects closer to the scheduled assessment dates. NCES decides which pilot and special studies to include each year. Thus, an estimate has been provided in the system clearance submittal for the 2011-2013 assessments, but the burden estimates will be refined in the submittals that are done each year for the upcoming yearly assessments.

Averages based on prior assessment data are used to calculate the teacher, school, and SD and ELL burden estimates. For teachers, we estimate that 4 teachers per school will complete NAEP questionnaires. For the school questionnaire, the estimated burden is a function of the number of schools. In a large assessment year (e.g., 2011, 2013), the average number of students per school is approximately 54, while in smaller assessment years (e.g., 2012), the average is 40 students per school.

Another component of the school burden is the pre-assessment visit and e-filing activities, which require school personnel time. Survey sample information is collected from schools in the form of lists of potential students who may participate in NAEP. This sample information can be gathered manually or electronically at the school, district, or state level. If done at the school or district level, some burden will be incurred by school personnel. However, the e-filing process is only done in a subset of schools (approximately, 38% in recent years), which was taken into account when computing the e-filing burden.

The estimates stated above are based on prior assessment data and can change over time. For example, prior versions of NAEP system clearance packages used much lower percentages for the SD and ELL student participation rates. Over the past few years, NAEP has taken measures to include as many students as possible and expand the accommodations allowed. Based on the 2008 and 2009 data, the percentages of students identified as either SD or ELL has increased to 22 percent for grade 4, 18 percent for grade 8, and 14 percent for grade 12. Thus, the SD and ELL volumes and burdens have increased significantly over prior submittals.

4. As is common practice with many other NCES surveys, OMB would like to be invited to attend technical review panel meetings. Please indicate how this might be accomplished.

Because NAEP assesses numerous content subject areas with complex designs, undertakes extensive instrument development, advanced analysis and reporting of NAEP results, NCES holds a large number of technical review panel meetings to ensure quality, relevance, and timeliness. The specific purposes of these meetings vary, but may include providing input on the content of cognitive and background questionnaire items at various stages of development, reviewing design and analysis plans, discussing and conducting validity studies, or planning quality assurance activities. Technical review panels also conduct item reviews by submitting written comments or meet remotely via a WebEX or conference call. The majority of the technical review panel meetings are not directly related to data collection, rather their purpose is to review the content of the cognitive item sections of the assessment instruments at various stages of development and thus may not be the most informative for OMB purposes (e.g. NAEP Design and Analysis Committee, State/TUDA Item Review, and NAEP Subject Area Standing Committee). There are, however, a few technical review panels that NCES believes would be quite informative to OMB, such as the National Validity Studies Panel, the Quality Assurance Technical Panel, and the meetings of the National Assessment Governing Board (see table below). If OMB is interested in attending specific meetings, NCES will be glad to include the designated OMB contact(s) on the meetings' distribution list, keeping OMB informed of the upcoming meeting dates, locations, and agendas (the only meetings for which NCES does not send out invitation because they are not organized by NCES are the NAGB meetings).

Name of Technical Review Panel	Purpose	Participants	Approximate Periodicity/Time
National Assessment Governing Board	Set policy for NAEP and review and approve cognitive and background questionnaire items	Educators, Policy Makers, Researchers, and General Public appointed by the Secretary of Education	Quarterly and remotely as needed March, May, August and November
NAEP Validity Study Panel	Provides a technical review of plans and products related to the validity of NAEP	Education and measurement researchers	Three times a year Winter, Summer and Fall
NAEP Quality Assurance Technical Panel	Advise NCES contractor on studies to improve NAEP processes and procedures	Education and measurement researchers	Two times a year January and July
NAEP Background Variable Standing Committee	Provide feedback on background questionnaire items and other background variables to NCES item development	Demography, education, and measurement researchers	One to two times per year

	contractor		
NAEP National Indian Education Study Technical Review Panel	Provide input to NCES contractors on the development of the NIES questionnaire items, data analysis and reporting	Experts in Native American and Alaska Native education and other education and measurement researchers	Prior to the administration of new questionnaire items and prior to the analysis and development of reports
Socio-economic status (SES) Expert Panel	Provide input to NCES contractors investigating the development of a new measure of socio-economic status for NAEP	Demography, education, and measurement researchers	Ad Hoc basis

5. To what does NCES attribute the low take-up rate by teachers of the on-line survey instrument and how does this compare to other NCES teacher surveys administered on-line?

In recent years, the on-line survey participation rate has been consistently around 10%. The paper-and-pencil versions are deemed more convenient for many teachers for various reasons, including: many teachers will complete the survey when they have time available, using multiple brief 5-minute timeframes to work on the survey; many teachers work on the surveys while students are working independently; not all schools have computer capabilities; and, teachers may need to reference information for the survey that is not available when they are on the computer. Teacher questionnaires are distributed to teachers at the pre-assessment visit. They are asked to complete the questionnaires (either online or in hardcopy) so that they can be collected when the administrators return for the actual assessment (approximately two weeks later).

Other NCES on-line teacher surveys are the Teacher Follow-up Survey (TFS) and the Beginning Teacher Longitudinal Study (BTLS). The 2008 - 2009 TFS featured an internet instrument with a paper-pencil response back-up. Official response rates will be released in August 2010, but preliminary un-weighted results indicate a response rate of over 60%. Additional responses were gathered with the support of an interviewer who typed in responses during a follow-up telephone interview. The BTLS, the first NCES survey being completed without paper-pencil back-up, is still in the field and final response rates are not yet available. Both surveys have a longer data collection window for teachers to respond than NAEP does. Similarly, the High School Longitudinal Survey (HSLS) has experienced a 60% online response rates for their teacher survey and uses CATI as a backup (HSLS does have a paper and pencil version), but, again, HSLS teachers have 5 months to complete the questionnaire as compared to the two weeks available to teacher filling out the NAEP survey.

6. Please provide some data on the response rates of private schools in past NAEP administrations.

Response rates of private schools for the 2009 administration were:

Grade 4 – 72%
Grade 8 – 72%
Grade 12 – 52%

7. Please provide estimates of the student response rates by grade in past NAEP administrations.

Response rates per grade for the 2009 administration were:

Grade 4 – 95%
Grade 8 – 93%
Grade 12 – 80%

8. There are a number of procedural edits to the Supporting Statement that we have discussed with NCES's liaison. We look forward to seeing revised documents in response to that conversation.

Procedural edits provided in a 2/24 email by the NCES-OMB liaison:

1. The issue we talked about (related words: cognitive vs. non-cognitive items, assessment vs. background, questionnaire vs. survey; this includes a need for consistent use throughout)

2. Assurance that untested items will go through cog lab testing before being tested on large numbers of people in a pilot. (p.4)

3. Taking out the word "annual" from "Details for selected special studies will be included in subsequent *annual* clearance requests." (p.8)

4. Talking about reporting six months after data collection elsewhere in the package and revising the paragraph: "In addition to the large volume of development and assessment that takes place, there is the factor of short turnaround times for production and reporting. NAEP is required to report mathematics and reading scores on the state and national levels every other year and results must be reported within six months of the administration. Due to the increased volumes, complex development, and six month reporting requirements, there are shorter time frames for data analysis, question reviews, assembly of questions for submittal to the Governing Board and OMB, and abbreviated windows for printing and distribution. The system clearance process shortens the timeframe for OMB approvals from 120 days to 45 days, which is critical for meeting printing and distribution deadlines." (p.9)
to something along the lines of: "In addition to the large volume of development and assessment that takes place, there is the factor of short turnaround times for production. Due to the increased volumes and complex development, there are shorter time frames for data analysis, question reviews, assembly of questions for submittal to the Governing Board and OMB, and abbreviated windows for printing and distribution. To meet these demanding timelines, NCES is requesting a waver of the 60-day federal register notice for individual clearances of studies described in this system clearance package."

5. Rephrase the paragraph on public comment to: One public comment was received in response to the 60-day Federal Register notice (Vol.74, page 57159, published on November 4, 2009). It pointed out that ELL student exclusion policies vary state-by-state and sometimes district-by-district. In response, NCES revised the ELL bullet point in the Supporting Statement Part B Section 1 so that it is not specific regarding which ELL students are excluded. The new bullet point reads: "Students are selected according to student sampling procedures with these possible exclusions: 1) The student is identified as an English Language Learner (ELL), preventing participation in NAEP, even with accommodations allowed in NAEP."

6. In Part A section 9 (Payments or Gifts to Respondents), add a statement to the effect that any incentives that might be used will be justified in future clearance packages and will be consistent with amounts approved in other studies with similar conditions.

7. Please revise Part A section 10 (Assurance of Confidentiality) to have one paragraph that discusses anonymity in NAEP – whose personally identifiable information (PII) is not collected (these are not subject to confidentiality laws and should be cited anonymity in their materials, not confidentiality). In another paragraph, please talk about those for whom identifiable information is collected (e.g. administrators, schools), where is it stored, with what precautions, and when is it destroyed and how. The confidentiality section should also cite FERPA and NCES statute. This section should also indicate under what laws is this collection authorized (e.g. ESEA and ESRA). OMB wants for Marilyn Seastrom to check this section before it is finalized.

8. Part B needs to include a brief paragraph that describes what the sampling frame for NAEP studies is (such as CCD, etc.).

9. Revise the sentence in Part B: "Plans also call for requesting NCES to provide letters to states and districts in support of the operational and filed tests" to "NCES will provide letters to states and districts in support of the operational and filed tests" – the idea is that the package is submitted by NCES and should not sound as if it was submitted by a contractor.

10. In Part B section 5 (Consultants…) please include government employees who contribute their expertise (especially from NCES).

We have incorporated the revisions referenced here in the final version of the System Clearance submittal. These revisions can be viewed in the track changes version of the documents (*Part A (Sys Cl) OMB files_Changes between 1-5 and 3-17 submissions.docx*) *and* (*Part B (Sys Cl) OMB files_Changes between 1-5 and 3-17 submissions.docx*), *included* as separate attachments with the final Part A and Part B files.

2010 NAEP System Clearance 3rd OMB Passback Responses 04-08-2010

1. Multi-Stage testing Study – Given the potentially large costs associated with this strategy for addressing estimate precision, we would like more information about estimated outyear costs as well as NCES's plans for comparing the costs and benefits of this strategies for addressing.

   At this time, NCES has no sound basis for estimating out-year cost or for doing a cost-benefit analysis of converting NAEP to an adaptive assessment. If the study results indicate that adaptive testing is feasible and yields more precise estimates in NAEP, additional studies and possible changes to the NAEP operational procedures could be implemented, such as making some assessments, like mathematics, adaptive. However, until the results are analyzed, no plans will be made. NCES' current focus primarily is on exploring the psychometric feasibility of the concept of adaptive testing in NAEP. If the concept proves to be feasible, NCES will have to determine how to apply the results to operational NAEP in a way that would be cost effective and efficient. Using the current delivery system for the writing assessment, which relies on administering the assessment on NAEP laptops which have to be transported to and from schools, might not be optimally cost effective or efficient if employed on a larger scale for assessments other than writing. At the time of researching and trying out delivery systems for the writing assessment, the unreliability and lack of commonality among computer systems in NAEP schools led NCES to rule out the possibility of using school-based computers for the assessment delivery. Current interest in computer based assessments, at the state and district levels, however, suggests that delivery via school-based

computers might be a future possibility for administering NAEP adaptive assessments. NCES would also want to explore other possibilities for delivering NAEP as a computer-based assessments. If this study is successful, NCES would have to explore delivery technologies before determining how to proceed. NCES, therefore, has no current basis for estimating out-year costs or doing a cost-benefit analysis for computer-based NAEP adaptive assessments.

2. In addition to the Texas Survey 1st year college students, please clarify the full scope of planned activities related to academic preparedness of 12th graders.

Proposed NAEP activities associated with investigating the preparedness of 12th graders, fall within NAGB's plans. NAGB has established a Technical Panel on 12th Grade Preparedness Research to advise on preparedness studies. A list of these studies can be found at http://www.nagb.org/publications/PreparednessFinalReport.pdf

3. 8th grade technology literacy – please explain the scope and what set of activities are proposed under "development cycle for innovative and new computer-based assessments."

We have attached an account of the expected development activities.

4. Thank you for clarifying which types of TRP meetings may be most productive for OMB to attend. Please add both Allison Cole and Shelly Wilkie Martinez to the distribution lists of the five panels mentioned.

The names have been added to the distribution lists for the five panels: NVS Panel, Quality Assurance Technical Panel, Background Variables Standing Committee, National Indian Education Study TRP, and the SES Expert Panel.

5. NAEP-TIMSS Linking Study -- Why does the NAEP program believe that there needs to be an increase from four to six states in this study (note: the TIMSS program provided no justification, rather indicating it was as a result of discussions with the NAEP program)? What is the total additional cost of increasing by approximately 100 schools and 4000 students, and what share of the increase is each program bearing?

   The decision to increase the number of states was made after considerable discussion with contractors and other experts about the design of the proposed NAEP-TIMSS Linking Study. It was decided that the sample should be increased to 8 states in order to provide adequate data to validate the outcome of the NAEP-TIMSS Linking Study, raising the number of resulting state estimates from 4 to 8. NCES has requested $3.45 million for the state benchmarking effort to be conducted during the TIMSS testing window and $5 million for the braided assessment printing, distribution, and administration [to be cleared under a separate OMB package]; scoring and linking of all of the NAEP blocks; and for contracting the NAEP state coordinators to recruit the state benchmarking schools and the additional 8th grade schools where braided assessment will be administered during the TIMSS and the NAEP testing windows. Based on the PIRLS/TIMSS field test and ensuing discussions with NCES staff and contractors, NCES designed the state samples so as to require fewer schools per state than originally estimated and arranged to coordinate school recruitment for the TIMSS/NAEP state benchmarking portion with the help of NAEP coordinators,

thereby reducing the expected per-school cost. As a result, NCES has determined that sampling 8 states is feasible within the requested budget. A future NAEP clearance package will seek OMB approval for the burden associated with the NAEP background questionnaire that will be given in association with the braided assessment during the NAEP window and will provide more details about the NAEP-TIMSS Linking Study as a whole.