National Center for Education Statistics

National Assessment of Educational Progress

*Volume I*

*Supporting Statement*

***NAEP Technology Based Assessments (TBA)***

***Tools and Item Types Usability Study***

*OMB# 1850-0803 v.112*



August 6, 2014

Contents

[1 Submittal-Related Information 1](#_Toc393979952)

[2 Background and Study Rationale 1](#_Toc393979953)

[3 Recruitment and Sample Characteristics 1](#_Toc393979954)

[4 Study design and data collection 2](#_Toc393979955)

[5 Consultations outside the agency 3](#_Toc393979956)

[6 Justification for Sensitive Questions 3](#_Toc393979957)

[7 Paying Respondents 3](#_Toc393979958)

[8 Assurance of Confidentiality 4](#_Toc393979959)

[9 Estimate of Hourly burden 4](#_Toc393979960)

[10 Cost to federal government 4](#_Toc393979961)

[11 Project Schedule 5](#_Toc393979962)

1. Submittal-Related Information

This material is being submitted under the generic National Center for Education Statistics (NCES) clearance agreement (OMB #1850-0803), which provides for NCES to conduct various procedures (such as field tests, cognitive interviews, and usability studies) to test new methodologies, question types, or delivery methods to improve survey and assessment instruments.

1. Background and Study Rationale

The National Assessment of Educational Progress (NAEP) is a federally authorized survey of student achievement at grades 4, 8, and 12 in various subject areas such as mathematics, reading, writing, science, U.S. history, civics, geography, economics, and the arts. NAEP is administered by NCES, part of the Institute for Education Sciences, in the U.S. Department of Education. NAEP’s primary purpose is to assess student achievement in the different subject areas and collect survey questionnaire (i.e., non-cognitive) data to provide context for the reporting and interpretation of assessment results.

NCES is interested in studying the intuitiveness and the ease-of-use of interface tools and item types being developed for future NAEP assessments. Examples of recently developed tools include an on-screen scratchwork tool and an on-screen equation editor. Tools like these provide support to the student and allow greater flexibility in testing cognitive constructs. As part of the process of designing new tools and item types, it is important to study the user experience they provide. User testing of students is crucial for making design decisions that result in tools that are truly useful and intuitive to use. User testing can reveal usability problems that were unanticipated by developers and designers.

This study will present students with prototypes of new item types and tools on a touch-screen tablet. Students will be given assessment-related tasks to complete using the prototypes (see Volume 2). Results will consist of task completion success rates using each of the control and information elements studied. These task completion success rates will be combined with qualitative information from the ease-of-use survey and exit questions to compile recommendations for modifications to items and tools to ensure that their implementation in upcoming assessments does not present barriers to student performance. In addition, some of the information will be used to assess the usability of the hardware being used in the study. Screen size and keyboard and trackpad responsiveness are just some of the hardware properties that affect usability. Ease of use of the hardware will be used to inform future decisions regarding selection of appropriate systems for NAEP testing.

Volume I of this submittal contains descriptions as well as design, sampling, burden, cost, and schedule information for the study. Volume II contains examples of the type of tasks that will be included in the protocol as well as survey questions of the type that will be administered in the study. The appendices contain sample scripts, notifications, and thank you documents.

1. Recruitment and Sample Characteristics

NAEP State Coordinators (see section 5) will leverage their relationships within the states to recruit public schools. NCES representatives who work with private schools or the Bureau of Indian Affairs will leverage their relationships in a similar manner. Recruiting activities may include presentations about the study at NCES-sponsored workshops and meetings. Flyers will be handed out to interested parties at NCES-sponsored events (appendix A). Only schools that have not been selected to participate in NAEP assessments in the 2014-2015 school-year will be selected for this study, in order to eliminate any interference from participation in the usability study on participation in the NAEP assessment, and vice versa. Schools may participate in more than one user testing session if they wish, though students may participate in only one session, in order to prevent learning effects and to reduce individual burden. NAEP State Coordinators and other representatives will forward the contact information for willing schools to Fulcrum IT, the NAEP web technology contractor. Fulcrum IT will contact each school to make the arrangements for students from that school to participate.

Given the number of interactions to be tested, up to 360 students will participate in the user testing, spread across grades 4, 8, and 12. These students will be selected from no fewer than 10 and no more than 20 schools, providing a balance between having a variety of participants and considering program staff travel and time costs. Students chosen for the study will not be included or excluded based on demographic criteria.

1. Study design and data collection

Prior to the user testing, Fulcrum IT staff will contact participating schools to make logistical arrangements. Initial contact will be made via email (appendix B), with a follow-up call to gather necessary logistical information (appendix C). A parental notification letter (appendix D) will be provided to the schools for their use to notify parents or guardians of students in the study. Fulcrum IT will bring all necessary materials, including computers and tablets, to the schools on the day of the study.

User testing will be conducted in small batches over the course of the school year, as part of an iterative process of design and testing of new and revised interactions developed over that period for tablet-based assessments.

Developers will submit prototypes designed to test specific interactions, and these prototypes will be used in a subsequent user testing group. User testing data will be reported back to the developers and NCES, as they are collected so that decisions can be made regarding design modifications. Modified features or items may then be included in a later user testing session to validate the usability of the changes.

A variety of subject areas will be included, not to test the subject content, but to test interactions that may be unique to items for that subject. For example, math items will be used to test an on-screen calculator or equation editor, as that subject area uses those two particular interactions. Reading items will be used to test different passage layouts and panel controls that are unique to reading items.

In addition to the multiple item types tested using prototypes, different participant groups may be tested using different touch-screen tablets, in order to test the impact of different hardware or operating systems on the usability and the interactions.

Each student will perform the study tasks during a one-on-one session with a facilitator. For some of the tasks, the facilitator will give verbal instructions, such as *“Imagine that you want to change the color of the tool bar up there [point] from black to white. Please show me how you would do that.”* For other tasks, students will be instructed to follow the written instructions on the screen, or to attend to a tutorial. For most tasks, participants will be asked to explain what they are doing and why as they perform the tasks.

User testing will take no more than 75 minutes per student. Students will be allowed to take breaks as needed. Screen capture will be used to document on-screen activity for later analysis. Although screen capture will include audio from the session, no video of the students or their surroundings will be recorded.

Students’ success or difficulty in completing assigned tasks will be analyzed to determine which information or control elements are missing or insufficient to allow successful completion of anticipated user tasks. While successful completion of tasks will be recorded, it is the tools and item types that are being evaluated rather than the students. All results will be used only to make recommendations regarding the design and development of tools and item types.

Results will be analyzed in terms of the percent of participants who successfully completed each task and task element. An example finding would be: “40 percent of participants found the volume control without assistance.” This finding would be used to determine that the volume control needed to be made more visible to users in order to be used successfully by 100 percent of the students.

In addition, ease-of-use ratings will be collected for the tasks, as well as any comments the participants have. The following instruments that will be used to gather data are included in the Volume II protocol.

1. *Participant ID and Welcome Script* – The welcome script introduces the interviewer and explains the study.
2. *Computer & Tablet Familiarity Survey* – This survey will be used to determine if differences in student performance can be attributable to previous experience using computers or touch-screen tablets.
3. *User Testing Scenarios* – This protocol contains the script for the facilitators to guide the interactions of the participants with the tablets. Data will be recorded on a tablet or laptop used by the facilitator.
4. *Exit Questions* – Sample questions that can be asked at the completion of the user tasks*.*
5. *Ease-of-Use Rating Survey* – This survey, completed by the facilitator, is used to record ease-of-use ratings for tasks.
6. Consultations outside the agency

Fulcrum IT LLC (Fulcrum IT)

Fulcrum IT is the NAEP contractor responsible for the development and ongoing support of NAEP computer-based assessments for NCES, including the system to be used for this usability study for tools and item types. Fulcrum IT will conduct the usability study and produce the final report.

NAEP State Coordinators

The NAEP State Coordinator serves as the liaison between the state education agency and NAEP, coordinating NAEP activities in his or her state. As previously noted, NAEP State Coordinators from selected states will work with public schools within their states to identify participating schools.

Educational Testing Service (ETS)

ETS serves as the Item Development (ID) and Design, Analysis, and Reporting (DAR) contractor on the NAEP project, developing cognitive and survey items for NAEP assessments. ETS staff will be involved in some item development activities and may assist in conducting and/or observing some user testing sessions.

1. Justification for Sensitive Questions

Throughout the item and protocol development processes, effort has been made to avoid asking for information that might be considered sensitive or offensive.

1. Paying Respondents

Student participants will not receive financial incentives for participation in this study. However, students will be permitted to keep the earbuds they used during the study. Schools participating in the usability study will each be offered a $50 gift card to a major office/school supply store as a token of appreciation. Gift cards will be provided at the end of the site visit, and a thank you email will be sent that evening to the school contact (appendix E).

1. Assurance of Confidentiality

The usability study will not retain any personally identifiable information. Prior to the start of the usability study, participants will be notified that their participation is voluntary and that their answers may be used only for research purposes and may not be disclosed, or used, in identifiable form for any other purpose except as required by law [Education Sciences Reform Act of 2002 (20 U.S.C. §9573)].

Written notification will be sent to legal guardians of students before user testing is conducted. Participants will be assigned a unique identifier (ID), which will be created solely for data file management and used to keep all participant materials together. The participant ID will not be linked to the participant name in any way or form.

Screen actions with audio will be captured from each session. The only identification included on the files will be the participant ID. The screen capture will be used for analysis after the session. Small portions of the screen capture for select sessions will be used in NCES briefings in order to demonstrate the methodology used for this study. No identifying information will be included in data analyses or study briefings.

1. Estimate of Hourly burden

School administrators provide pre-assessment information and help with the logistics of student and room coordination and other related duties. The school administrator burden is estimated at 12 minutes for the pre-assessment contact and up to one hour (total) for administration support. It is possible that up to 30 schools will be contacted initially in order to identify the 10-20 required for the study.

Parents of participating students will receive a letter explaining the study, for which the parent’s burden is estimated at three minutes. An additional burden (15 minutes) is estimated for a small portion of parents (up to 40) who may write refusing approval for their child or may research information related to the study. As such, it is possible that up to 400 parents will need to be notified in order to identify the 360 students required for the study. Estimated hourly burden for the participants is described in Table 1, below. Participants will be subsets of the initial contact group.

Table 1. Estimate of Hourly Burden

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Person** | **Task** | **Number of Individuals** | **Individual Participant Burden** | **Total Burden** |
| School Administrator | Initial Contact by NAEP State Coordinator | 30 | 12 minutes | 6 hours |
| School Personnel | Scheduling & Logistics | 20 | 60 minutes | 20 hours |
| Parents | Initial Notification | 400 | 3 minutes | 20 hours |
| Parents\* | Researching Study | 40\* | 15 minutes | 10 hours |
| Students | User Testing | 360 | 75 minutes | 450 hours |
| **Total (850 responses)** | | **810** | NA | **506 hours** |

\* These parents are a subset of those who were initially notified.

1. Cost to federal government

Table 2 provides the overall project cost estimates.

Table 2: Estimate of Costs

|  |  |
| --- | --- |
| **Activity** | **Estimated Cost** |
| Recruitment, User Testing, Analysis | $243,779 |
| School Incentive Payments | $1,000 |
| **Total** | **$244,779** |

1. Project Schedule

Table 3 provides the overall schedule.

Table 3: Schedule

|  |  |
| --- | --- |
| **Date** | **Event** |
| September 2014-May 2015 | Ongoing Recruiting for User Testing Sessions |
| September 2014-June 2015 | User Testing |
| August 2015 | Summary Report |