

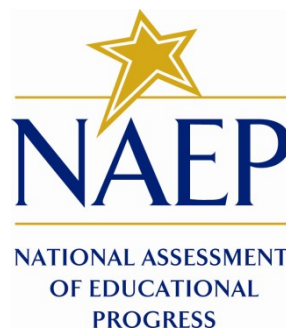
*NATIONAL CENTER FOR EDUCATION STATISTICS  
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS*

*Volume I  
Supporting Statement*

*Request for Clearance for  
NAEP Science Pretesting Activities*

*OMB# 1850-0803 v.73*

*Cognitive Interviews (Background Items and Science Interactive Computer  
Tasks [SICT] Cognitive Items)  
and  
Play Test Studies for Science Interactive Computer Tasks*



October 11, 2012

**Table of Contents**

1)	Submittal-Related Information.....	3
2)	Background and Study Rationale.....	3
3)	Sampling and Recruitment Plans.....	5
4)	Data Collection Process.....	8
5)	Consultations Outside the Agency.....	11
6)	Assurance of Confidentiality.....	12
7)	Justification for Sensitive Questions.....	12
8)	Estimate of Hourly Burden.....	12
9)	Estimate of Costs for Recruiting and Paying Respondents.....	15
10)	Costs to Federal Government for Cognitive Interviews and Play Testing.....	15
11)	Schedule.....	16

**Table of Figures**

Table 1. Sample Size: Science Background Questionnaire Items.....	6
Table 2. Sample Size: SICT Play Testing and Cognitive Interviews.....	8
Table 3. Estimate of Hourly Burden – Recruitment and Participation for Background Questionnaire Cognitive Interviews.....	13
Table 4. Specific Burden for Cognitive Interviewing for Science ICT.....	13
Table 5. Specific Burden for Play Testing for Science SICT.....	14
Table 6. Combined Burden for Background Questionnaire and SICT Science Pretest Activities	15
Table 7. Estimate of Costs.....	15
Table 8. Schedule of Milestones and Deliverables.....	16

## 1) Submittal-Related Information

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

## 2) 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 various subject areas and to also collect background questionnaire (i.e., non-cognitive) data to provide context for the reporting and interpretation of assessment results.

As part of NAEP's item development process, a portion of assessment items (cognitive and background) are pretested on a small number of students before they are administered to a larger sample through tryouts or pilot tests. These pretest activities can include cognitive interviews as well as play testing of items. This submittal requests clearance for various pretesting activities related to the upcoming science assessments that will be administered operationally in 2015 or 2019. Specifically, these pretest activities are

- Cognitive interviews for science background questions;
- Play testing for Science Interactive Computer Tasks (SICT) cognitive items; and
- Cognitive interviews for SICT cognitive items.

Included in the submittal are

- Volume I — supporting statement that describes the design, data collection, burden, cost, and schedules of the pretesting activities for the upcoming science assessments;
- Appendices — recruitment and communication materials; and
- Volume II — protocols and questions used in the pretesting sessions.

Educational Testing Service (ETS), the NAEP background and SICT item developer, is the lead contractor for these pretesting activities (see section 5). An overview of the background questionnaires and the SICTs is presented below, as well as a description of play testing and cognitive interviews.

### Background Questionnaires

In addition to assessing subject-area achievement, NAEP collects background questionnaire data to provide context for the reporting and interpretation of assessment results. These questionnaire data come from three respondent types: students, teachers, and school administrators. NAEP questionnaires serve to fulfill reporting requirements of federal legislation and to provide a context for reporting student performance.<sup>1</sup>

Periodically, NCES will add, revise, or delete questions from existing subject-specific background questionnaires. These modifications aim to improve questionnaire quality, replace or drop outdated questions, and collect data on new contextual factors that are expected to be associated with academic achievement. Questionnaires for the 2015 operational science assessment have undergone a systematic review process that has resulted in the addition and revision of several questionnaire items. These items have been reviewed by expert panels and NCES. In 2014, NCES will pilot test these new and revised

<sup>1</sup> Education Sciences Reform Act of 2002 (ESRA), National Assessment of Educational Progress (20 USC § 9622).

subject-specific background questions with a large national sample of students, teachers, and school administrators. Prior to the 2014 science pilot, these new and revised items will undergo cognitive interview testing. In addition, a subset of existing operational items will be included to enable comparison between the operational items and the proposed new or revised items.

### SICTs

The NAEP Science assessment covers content in three science disciplines: Life Science, Physical Science, and Earth and Space Sciences. There also are four science practices: *Identifying Science Principles* (which mainly involves the recall of principles), *Using Science Principles* (which mainly involves the application of science principles to construct explanations or make predictions), *Using Science Inquiry* (which mainly involves applying science principles to carry out scientific inquiry), and *Using Technological Design* (which involves using science principles to design solutions to real world problems or anticipate the effect of technological design decisions).

The framework states that SICTs should be used when the format offers advantages over other assessment methods (such as paper and pencil or hands-on tasks). The SICTs assess students' ability to use and apply their science knowledge and do science in a computer-based environment. In these tasks, students are presented with rich contexts and problems to solve on computer. Students may be asked to engage in a variety of simulated activities that are characteristic of science, such as designing scientific investigations or selecting and evaluating information relevant to the situations or problems they need to address. As such, numerous assessment items are included as part of each task.

SICTs were assessed as part of the 2009 science assessment. Development of SICTs for the 2015 and 2019 assessments has begun. These SICTs will target all three science content areas and four science practices, but most have a strong focus on students' knowledge, skills, and ability in science inquiry (i.e., the practice of *Using Science Inquiry*). To achieve efficiency with the development and rendering of the tasks, parallel versions or variants of some tasks will be created, thus reusing many of the elements of the initial variant. To view the 2009 SICTs, which are similar to those being developed for the current SICT project, please see: [http://nationsreportcard.gov/science\\_2009/ict\\_indepth.asp](http://nationsreportcard.gov/science_2009/ict_indepth.asp).

### Play Testing

Play testing is typically used in the development of computer games, to solicit user feedback on prototype versions of the games. Given that the SICTs feature simulations and other game-like features, play testing is a valuable way to gather feedback in the development process. For the SICTs, play testing will be done relatively early in the SICT task development process. It involves informal gatherings of students; in many ways, it is similar to a focus group approach. The purpose of play testing in the SICT project is to gather student views on early versions of the interactive computer tasks and begin to understand how students are thinking about those tasks. Students will be divided into small groups to look through paper or computer versions of PowerPoint mock-ups of the tasks (i.e., early prototype versions of the tasks). Assessment developers will give an overview of the tasks to students and provide guidance on what students should reflect on while looking at the tasks. These sessions are largely observational studies, although the facilitator may ask students some questions.

Through play testing, researchers will be able to identify construct-irrelevant features in tasks, such as inaccessible language in constructed-response stems or uninteresting or unfamiliar scenarios that result in poor student engagement with a task. Play testing will also help us identify usability and construct-irrelevant task features. Play testing early in the development cycle also allows us to refine tasks so that we can use subsequent and more intensive cognitive interviews (as described below) as more in-depth probes of how students think to collect important cognitive processing data as students engage and interact with the task.

### Cognitive Interviews

In cognitive interviews (often referred to as a *cognitive laboratory study* or *cog lab*), an interviewer uses a structured protocol in a one-on-one interview using two methods: think-aloud interviewing and verbal probing techniques. With think-aloud interviewing, respondents are explicitly instructed to "think aloud" (i.e., describe what they are thinking) as they figure out their answers to questions or tasks. The respondent reads each question, and then the interviewer records the cognitive processes that the respondent describes in arriving at an answer to the question. With verbal probing techniques, the interviewer asks probing questions, as necessary, to clarify points that are not evident from the "think-aloud" process. These probes might include, for example, asking the respondent to rephrase the question in his or her own words or assess whether the response categories for multiple-choice questions are relevant.

Cognitive interview studies are largely observational. The largely qualitative data collected will be mainly verbal reports in response to probes or from think-aloud tasks, in addition to volunteered comments. One objective is to explore how students are thinking and what reasoning processes they are using as they work through items. Another is to identify and correct problems of ambiguity or misunderstanding, or other difficulties respondents have answering questions.

For the background questions, the interviews will focus on how students, teachers, and school administrators answer questions about themselves, their classroom experiences, and their schools. For the SICT tasks, the objective of the cognitive interview study is to explore how students process information presented in the SICT tasks as they think through them.

## **3) Sampling and Recruitment Plans**

Existing research and practice have failed to offer a methodological or practical consensus regarding the minimum or optimal sample size necessary to provide valid results for cognitive interviews and similar small-scale activities.<sup>2</sup> Nonetheless, a sample size of five to fifteen individuals has become the standard. Several researchers have confirmed the standard of five as the minimum number of participants per subgroup for analysis for the purposes of exploratory cognitive interviewing.<sup>3</sup>

### **Background Questionnaires**

The recommended sample sizes for background questionnaire components of this study are based on an evaluation of the minimum number of participants in relation to the number of items and the complexity of each item. The new and revised items for the science subject-specific questionnaires are similar in

<sup>2</sup> See Almond, P. J., Cameto, R., Johnstone, C. J., Laitusis, C., Lazarus, S., Nagle, K., Parker, C. E., Roach, A. T., & Sato, E. (2009). White paper: Cognitive interview methods in reading test design and development for alternate assessments based on modified academic achievement standards (AA-MAS). Dover, NH: Measured Progress and Menlo Park, CA: SRI International. Available at: [http://alternateassessmentdesign.sri.com/documents/WhitePaper2009\\_Assessment\\_CognitiveInterview.pdf](http://alternateassessmentdesign.sri.com/documents/WhitePaper2009_Assessment_CognitiveInterview.pdf)

<sup>3</sup> See Van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). The think-aloud method: A practical guide to modeling cognitive processes. San Diego, CA: Academic Press. Available at: [ftp://akmc.biz/ShareSpace/ResMeth-IS-Spring2012/Zhora\\_el\\_Gauche/Reading%20Materials/Someren\\_et\\_al-The\\_Think\\_Aloud\\_Method.pdf](ftp://akmc.biz/ShareSpace/ResMeth-IS-Spring2012/Zhora_el_Gauche/Reading%20Materials/Someren_et_al-The_Think_Aloud_Method.pdf)

content and style to other existing NAEP science questionnaire items. Consequently, we will interview 10 student respondents per grade and 5 each for teachers and school administrators, per grade. (The grade levels comprise the subgroup of analysis for the science questionnaire items.) Therefore, to test these items adequately, there will be 55 cognitive interviews total across three respondent types (students, teachers, and school administrators) and three grade levels (4, 8, and 12), as summarized in table 1 below.

**Table 1. Sample Size: Science Background Questionnaire Items**

	<b>Grade 4</b>	<b>Grade 8</b>	<b>Grade 12</b>	<b>Total</b>
<b>Students</b>	10	10	10	30
<b>Teachers</b>	5	5	N/A	10
<b>School Administrators</b>	5	5	5	15
<b>Total</b>	<b>20</b>	<b>20</b>	<b>15</b>	<b>55</b>

Note: NAEP does not administer a teacher questionnaire at grade 12.

To ensure a diverse sample, students will be sampled to obtain the following criteria:

- Fourth-grade and eighth-grade students who are enrolled in a science course at some point during the current school year;
- An even mix of twelfth-grade students who are either currently enrolled in a science course or not currently enrolled in a science course;
- A mix of gender;
- A mix of race/ethnicity (Black, Asian, White, Hispanic); and
- A mix of socioeconomic background (at least four low-SES and at least four mid/high-SES per grade).

To ensure a diverse sample, teachers and school administrators will be sampled with the following criteria:

- A school population that includes fourth-grade students, or eighth-grade students, or twelfth-grade students, as appropriate;
- Teachers who are teaching fourth-grade or eighth-grade science at some point during the current school year;
- A mix of school sizes; and
- A mix of school socioeconomic demographics (at least two teachers and administrators from low-SES schools and at least two teachers and administrators at mid/high-SES schools).

Although the sample will include a mix of these characteristics, the results will not explicitly measure differences by those characteristics.

EurekaFacts, a subcontractor to ETS (see section 5), will recruit student, teacher, and school administrator participants for the background question-related cognitive interviews. They will use their Washington, DC/Baltimore metropolitan area participant database as well as leverage their contacts within organizations and groups (e.g., parent-teacher organizations, community organizations) that can serve as recruitment partners for the upcoming study (see appendices D through I for the sample recruitment scripts, e-mails, and letters). If needed, EurekaFacts will identify teachers and school administrators from targeted contact lists that may be purchased from reputable third-party vendors, such as Market Data Retrieval or the American Association of School Administrators. No more than two teachers or administrators will be recruited per school.

Potential participants will be screened using a screener script to ensure the mix of characteristics described above (see appendices D and G). For selected participants, EurekaFacts will confirm the interview date, time, and location with parents or legal guardians or students, if 18 years of age (see

appendices J and K). Teachers, school administrators, and the parents or legal guardians of the participating students (or students, if 18 years of age) will complete consent forms at the time of the interview (see appendices A through C for consent forms).

## ICTs

### Play Testing

For the SICT play testing, ETS will perform the recruitment, recruiting students from at least two demographic groups: students from urban districts and students from suburban districts. (The type of location, by grade level, comprises the subgroup of analysis for the SICT play testing.) Students will be recruited from districts that are located near the ETS Princeton, New Jersey, campus for scheduling efficiency and flexibility. Students may participate in play testing sessions only after we have received written consent forms from them (if over 18) or their parents (if under 18) (see appendices AB and AC) and only if they are from target demographic groups (in this case urban or suburban schools and from the targeted grades – Grades 4–5, 7–8, and 11–12).

With the above cited research in mind for play testing, we plan to convene five to six students per grade, per type of location, for each task. We believe five to six students should be sufficient at the play testing stage given that the key purpose is to identify usability errors and other construct irrelevant issues.<sup>4</sup> Each task will be tested at two sessions (defined by type of location – urban or suburban).

Given the number of tasks that can be tested during a single session and the number of total tasks, we anticipate 14 total sessions (7 for each type of location). Based on prior experience with similar studies, it is anticipated that the same students would return to participate in multiple sessions. Specifically, within type of location, we are assuming that each student will participate in approximately half of the sessions. Therefore, play testing is expected to involve a total of 72 students across the three grades (6 students \* 2 types of location \* 3 grades \* 2 “groups” of students [because each student will participate in approximately half of the sessions]).

ETS will recruit students using existing ETS contacts with teachers and staff at local schools and afterschool programs for students. E-mail or letters will be used to contact these teachers and staff at local schools and afterschool programs (see appendix AF). Paper flyers (appendix AG) and consent forms (appendices AH and AI) for students and parents will be distributed through these teachers and staff contacts. During this communication, the parent/guardian or student will be informed about the objectives, purpose, and participation requirements of the data collection effort, as well as the activities that it entails. Confirmation e-mails and/or letters will be sent to participants (appendices AJ and AK). Only after ETS has obtained written consent from the parent/guardian and/or student will the student be allowed to participate in the play testing session.

### Cognitive Interviews

Based on the research stated earlier, for the cognitive interviews for the SICTs, EurekaFacts plans to interview seven students per grade, per task. We believe that seven students per task should be sufficient at this stage given that the key purpose of the cognitive interview is to identify qualitative patterns in how students think at different points in the task and confirm the validity of the assessments. Based on the number of tasks that can be completed per session and the number of tasks to go through the cognitive interview process, cognitive interviewing is expected to involve a total of 140 students across the three grades.

For the SICT cognitive labs, EurekaFacts plans to recruit students from the following demographic populations:

<sup>4</sup> See Nielson, J. (1994). Estimating the number of subjects needed for a think aloud test. *Int J. Human-computer Studies*. 41, 385-397. Available at: <http://www.idemployee.id.tue.nl/g.w.m.rauterberg/lecturenotes/DG308%20DID/nielsen-1994.pdf>

- A mix of race/ethnicity (Black, Asian, White, Hispanic);
- A mix of socioeconomic background; and
- A mix of urban/suburban/rural

Although the sample will include a mix of student characteristics, the results will not explicitly measure differences by those characteristics.

EurekaFacts will perform the recruiting for the SICT cognitive interview sessions (see section 4 for recruitment areas). While they will use various outreach methods to recruit students to participate in the SICT cognitive interviews, the bulk of the recruitment will be conducted by telephone and based on their acquisition of targeted mailing lists containing residential address and land line telephone listings. EurekaFacts will also use a participant recruitment strategy that integrates multiple out-reach/contact methods and resources such as newspaper/internet ads, outreach to community organizations (e.g., Boys/Girls clubs, Parent-Teacher Associations), and limited on-site location-based (such as distribution of bookmarks) and mass media recruiting (such as postings on the EurekaFacts website).

Interested participants will be screened to ensure that students meet the criteria for participation in the study (e.g., they are from the targeted demographic groups outlined above and have given consent if they are over 18, or their parents/guardians have given consent if they are under 18). When recruiting participants, EurekaFacts staff will first speak to the parent/guardian of the interested minor before starting the screening process. During this communication, the parent/guardian will be informed about the objectives, purpose, and participation requirements of the data collection effort as well as the activities that it entails (see appendices O thru W for recruitment communication documents). After confirmation that participants are qualified, willing, and available to participate in the research project, they will receive a confirmation e-mail/letter and phone call (see appendices X, Y, Z, and AA). Informed parental consent will be obtained for all respondents under the age of 18 who are interested in participating in the data collection efforts.

Table 2 summarizes the number of students for both SICT activities.

**Table 2. Sample Size: SICT Play Testing and Cognitive Interviews**

	<b>Grade 4</b>	<b>Grade 8</b>	<b>Grade 12</b>	<b>Total</b>
Play Testing	24	24	24	72
Cognitive Interview	49	49	42	140
<b>Total</b>	<b>73</b>	<b>73</b>	<b>66</b>	<b>212</b>

## 4) Data Collection Process

### Cognitive Interview Process (Background Questionnaire and SICTs)

EurekaFacts will conduct the cognitive interviews, with oversight from ETS, for both the background items and the SICTs. EurekaFacts will ensure that qualified interviewers, trained on the cognitive interviewing techniques of the protocols, conduct the interviews. The interviews will be based on the protocol structures described in Volume II (Parts B and D).

Participants will first be welcomed, introduced to the interviewer and the observer (if an in-room observer is present), and told they are there to help answer questions about how people answer survey or SICT items or tasks. Participants will be reassured that their participation is voluntary and that their answers may be used only for statistical 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]. Interviewers will explain the think-aloud process, conduct a practice question, and then participants will answer questions verbally.

Two methods will be employed for the think-aloud component of the cognitive interviews: (1) *concurrent* think-alouds (in which the participant verbalizes his or her thoughts *while* working through the item or task), and (2) *retrospective* think-alouds (in which the student verbalizes his or her thoughts *after* completing the item or task). The method also includes a verbal probing component conducted after completion of the think-aloud portion for that item or task. The amount of verbal probing done after each item or task varies, depending on the type of items or tasks and the goals of the study. For all background questions, the concurrent method will be used. For the SICTs, both concurrent and retrospective methods will be available and the specific approach used will be determined depending on the complexity of the task and the potential for a concurrent think-aloud to interfere with student performance.

Digital audio recording will capture students' verbal responses to the think-aloud interview. Interviewers may also record notes, including behaviors (e.g., the participant appeared confused) and if extra time was needed during a particular part of the task.

### Background Questionnaires

Student interviews will take place at the EurekaFacts cognitive interview laboratory site in Rockville, Maryland. Teachers and school administrators who agree to participate will be interviewed at their school locations. If needed (i.e., under circumstances in which highly interested participants are unable to meet in-person), a limited number of teacher and school administrator interviews may be conducted via telephone. Telephone interviews will be restricted to no more than two teachers and two administrators for the entire study. For these interviews, EurekaFacts will use viewing software such as *GoToMeeting*<sup>5</sup> to facilitate the educators' ability to view any necessary items during the cognitive testing experience.

After the participant reads each question and answers it while thinking aloud, the interviewer will ask both item-specific probes, as well as possibly asking some generic probes. The protocol which contains the welcome script, think-aloud instructions, hints for the interviewers, the specific background items included, and the generic and item-specific probes are contained in Volume II, Part B.

### SICTs

For the SICT cognitive interviews, EurekaFacts will also interview students at their Rockville, Maryland, site. In addition, should the need arise, additional recruitment/capture areas may include nearby school districts in West Virginia, Eastern Maryland, and the Baltimore metro area. In those instances, EurekaFacts will conduct off-site interviews in community centers or specialized (focus) group interview facilities.

As with the background questionnaire protocols, the protocols shown in Volume II, Part D contain the welcome script, think-aloud instructions, and hints for the interviewers. The SICT protocols contain a majority of generic interview questions that may apply to all tasks. For example: *What's going on in your head right now?* and *I see you're looking at the task [or screen/figure/chart/text]. What are you thinking?*

In addition, following the completion of the task, the interviewer will proceed with follow-up questions. In this verbal probing component, the interviewer asks students targeted questions about specific aspects of knowledge, skill, or ability that the task is attempting to measure, so that the interviewer can collect more information on the strategies and reasoning that students employed as they worked through a task. Verbal probing will include some targeted questions around particular aspects of the construct that SICTs

<sup>5</sup> See [www.gotomeeting.com](http://www.gotomeeting.com) for web conferencing and meeting tools description.

are measuring. For example, if part of a task targets a particular aspect of science inquiry that we want to explore in more depth, such as controlling variables when designing investigations, we might ask students to verbalize what they were thinking during the part of the task when they were asked to control variables. The targeted questions will be generated by ETS for each task and provided to EurekaFacts prior to testing (see Volume II, Part D, IV).

To facilitate retrospective think alouds, EurekaFacts will use *Morae Recorder* software (described in section 5). This system allows us to capture all of the screen events (e.g., where students click with their mouse and where they move their mouse on the computer screen) as well as simultaneous audio and video of the student participant, via a webcam (which captures their verbalizations and their nonverbal behaviors such as posture and facial expressions). This audio and video capture of students' behavior on the computer is especially helpful for retrospective think alouds because it allows both the interviewers and the students to review together all of the events that occurred during the task and thus enables interviewers to ask students to recall what was in their minds at each part of the task.

Where we are testing parallel versions or variants of tasks, we will not use a full task protocol, but rather will apply the think-aloud methodology to selected elements of the task variants. This will allow us to focus on, for example, those aspects that are different (such as specific content) while establishing that the overall structure of the task functions equally in both variants. Students will first perform the full cognitive interview think-aloud methodology on the initial version and then use the same methodology on the elements of focus in the parallel version.

### Analysis Plans

For the cognitive interview data collections, the key unit of analysis is the item. For the SICT cognitive interviews, although the analysis will be conducted at item level, for later reference, documentation will be grouped at the task level. Items will be analyzed across participants.

The types of data collected about the questions will include

- think-aloud verbal reports;
- behavioral data (e.g., errors in reading items or tasks; actions observable from screen-capture);
- responses to generic questions prompting students to think out loud;
- responses to targeted questions specific to the item or task;
- additional volunteered participant comments; and
- debriefing questions.

The general analysis approach will be to compile the different types of data in spreadsheets and other formats to facilitate identification of patterns of responses for specific items or tasks, for example, patterns of counts of verbal report codes and of responses to probes or debriefing questions. Each type of data for an item will be examined both independently and in conjunction with item-specific features in order to determine whether a feature or an effect of an item is observed across multiple measures and/or across administrations of the item. This approach will help ensure that the data are analyzed in a way that is thorough, systematic, and that will enhance identification of problems with items or tasks and provide recommendations for addressing those problems.

The deliverables from the analyses of the cognitive interviews will be a set of documents that pertain to each item (for background questionnaires) or task (for SICT). These will include information from EurekaFacts, such as observation data, student response data, interviewer notes, and post-task reports.

In addition, for the SICT, data from follow-up qualitative analyses of the screen capture of student actions and student verbalizations during think-aloud will be analyzed. These additional analyses will not be done

for every student in every task, due to the costly nature of this manual coding work; however, we feel it is important to have the facility for a more in-depth cognitive analysis to inform our understanding of the thinking occurring in specific places. The complete set of documents for each task, generated by the cognitive interview data collection and analysis procedures, will go into the record of development (the documentation that supports the revisions and decisions for the task throughout the development process).

### **Play Testing Process (SICTs only)**

As with the cognitive interviews, participants will first be welcomed and introduced to the interviewers. Participants will be reassured that their participation is voluntary and that their answers may be used only for statistical 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].

Assessment developers will then give an overview of the tasks to students and provide guidance on what students should reflect on while looking at the tasks. Assessment developers and cognitive scientists from ETS will act as facilitators and observers, taking notes on what students say and interjecting an occasional question aimed at eliciting students' reactions, places of confusion, and ways of thinking about the answers to the questions in the tasks. Each observer may choose to stay with one group of 2–3 students looking at and responding to tasks, or they may choose to move around to observe several groups of students looking at and responding to tasks.

For the most part, students will be allowed to explore and interact with the mocked-up task versions by themselves with little intrusion on the part of the interviewer. However, at a few strategic points, the interviewer may introduce questions meant to explore students' reactions to the task, such as:

*Did you find the problem in this task interesting – why or why not?*

*Are there any questions or words that seem confusing here? Did you understand that part?*

*How would you answer this question? [Ask different group members if their approaches would differ].*

*How could this task be improved? Could it be clearer, or more interesting, for example?*

Prior to each play testing session, ETS staff may identify some key focus areas for each task. If students do not provide sufficient comments on targeted parts, a staff member may ask a group of students if they had any thoughts about the particular sections, using questions such as those described above, but focused on specific places or issues in the task.

The deliverables from the play testing sessions will be aggregated notes from the observers in each session primarily on usability and design issues that are observed, and informal observations of student comments and feedback. One aggregate document will be produced for each task that is observed, with all observers contributing their observations to this common document. Since play testing is a more informal process that generates relatively unstructured information, we will not conduct formal analyses on these data. But the documents produced will also be part of the record of development.

## 5) Consultations Outside the Agency

Educational Testing Service (ETS) serves as the Item Development contractor (including all background items) and as the NAEP Science SICT Development contractor. As such, ETS will be responsible for the management of all activities described in this package. In addition, ETS will be responsible for the recruitment for and conduct of the SICT play testing sessions and guide and oversee the cognitive interview sessions (both background questionnaire and SICT) operated and conducted by EurekaFacts, LLC Associates.

EurekaFacts is a small, established for-profit research and consulting firm. EurekaFacts is working as a subcontractor for ETS on both the SICT and background questionnaire development projects to conduct the cognitive interviews. EurekaFacts offers facilities, tools, and staff to collect and analyze both qualitative and quantitative data. Located in Rockville, Maryland, EurekaFacts has a Cognitive Testing Laboratory facility that offers a range of cognitive interviewing and usability testing services. It also owns and operates *Morae software*, which allows for video and audio capture of students being interviewed along with technology to capture all of students' moves on the computer, including mouse clicks, mouse trails, time on task, and places for interviewers to insert comments and notes at any part of the task. The software also provides remote access to video so that NCES and ETS staff can observe the interviews from a distance in real time.

## 6) Assurance of Confidentiality

Participants will be notified that their participation is voluntary and that their answers may be used only for statistical 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 consent will be obtained from legal guardians (of minor students) or from students 18 years or older, teachers, and school administrators before interviews are conducted. Participants will be assigned a unique student 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. The consent forms, which include the participant name, will be separated from the participant interview files and secured for the duration of the study and will be destroyed after the final report is released.

The interviews will be recorded. The only identification included on the files will be the participant ID. The recorded files will be secured for the duration of the study and will be destroyed after the final report is submitted.

## 7) Justification for Sensitive Questions

Throughout the item, task, and process of developing interview protocols, effort has been made to avoid asking for information that might be considered sensitive or offensive. Reviewers have attempted to identify and minimize potential bias in questions.

## 8) Estimate of Hourly Burden

### Cognitive Interview Burden – Background Questions

The estimated burden for recruitment assumes attrition throughout the process.<sup>6</sup> Initial contact and response is estimated at 3 minutes or 0.05 hours. The follow-up phone call to screen participants and/or

<sup>6</sup> Assumptions for attrition rates, based on prior studies, are as follows.

answer any questions the participants (or their parents or legal guardians) have is estimated at 9 minutes or 0.15 hours per participant. The follow-up to confirm participation is estimated at 3 minutes or 0.05 hours. All interviews will be scheduled for no more than 60 minutes. Table 3 details the estimated burden for the background questionnaire cognitive interviews.

**Table 3. Estimate of Hourly Burden – Recruitment and Participation for Background Questionnaire Cognitive Interviews**

<b>Respondent</b>	<b>Hours per respondent</b>	<b>Number of respondents</b>	<b>Total hours</b>
<b>Parent or Legal Guardian, and Student Recruitment</b>			
Initial contact	0.05	350	17.5
Follow-up via phone	0.15	70*	10.5
Confirmation	0.05	35*	1.75
<b>Sub-Total</b>		<b>350</b>	<b>29.75</b>
<b>Teacher and School Administrator Recruitment</b>			
Initial contact	0.05	100	5
Follow-up via phone or e-mail/confirmation	0.15	50*	7.5
<b>Sub-Total</b>		<b>100</b>	<b>12.5</b>
<b>Participation (Interviews)</b>			
Grade 4 Students	1	10	10
Grade 8 Students	1	10	10
Grade 12 Students	1	10*	10
Teachers	1	10*	10
School Administrators	1	15*	15
<b>Sub-Total</b>		<b>55*</b>	<b>55</b>
<b>Total Burden</b>		<b>470</b>	<b>97.25</b>

\* Subset of initial contact group

#### Cognitive Interview Burden – SICT Items

Based on the proposed outreach and recruitment methods, we estimate initial respondent burden, regardless of the mode of initial interaction (e.g., a telephone recruiting call, receipt of a request to participate by postal mail, or receipt of an e-mailed message regarding the study) at 3 minutes or 0.05 hours. The follow-up phone calls to conduct participant screening and schedule the interviews are estimated at 9 minutes or 0.15 hours per family. The follow-up phone call and letter to confirm participation is estimated at 3 minutes or 0.05 hours. Student interviews with SICTs will be limited to 90 minutes for all students. The estimated burden for recruitment assumes attrition throughout the process, using the same rates as stated earlier. Table 4 details the estimated burden for the SICT cognitive interviews.

#### Play-Testing Burden – SICT Items

Teachers and staff officials will be contacted via e-mail and phone. Initial e-mail contact, response, and distribution of materials is estimated at 10 minutes or 0.167 hours. We anticipate distributing 200 flyers and consent forms via these contacts to parents and students, of which half will be flyers and half will be

- Students: 80 percent from initial contact to follow-up, 50 percent from follow up to confirmation and 15 percent from confirmation to participation.
- Teachers and School Administrators: 50 percent from initial contact to follow-up, 50 percent from follow up to participation.

consent forms for parents. Time to review flyers and consent forms is estimated at 5 minutes or 0.08 hours. For those choosing to fill out the consent form, the estimated time is 8 minutes or 0.13 hours. The follow-up e-mail or letter to confirm participation for each session is estimated at 3 minutes or 0.05 hours. Play testing sessions are expected to last 60 minutes for all students. Table 5 details the estimated burden for the SICT play testing.

**Table 4. Specific Burden for Cognitive Interviewing for Science ICT**

Respondent	Hours per respondent	Number of respondents	Total hours
<b>Parent and Student Recruitment</b>			
Initial contact	0.05	1,650	82.5
Follow-up via phone	0.15	330*	49.5
Confirmations	0.05	165*	8.25
<b>Sub-Total</b>		<b>1,650</b>	<b>140.25</b>
<b>Interviews</b>			
Grade 4	1.5	49	73.5
Grade 8	1.5	49	73.5
Grade 12	1.5	42*	63
<b>Sub-Total</b>		<b>98</b>	<b>210</b>
<b>Total Burden</b>		<b>1,748</b>	<b>350.25</b>

\* Subset of initial contact group

**Table 5. Specific Burden for Play Testing for Science SICT**

Respondent	Hours per respondent	Number of respondents	Total hours	Average number of sessions per respondent	Total hours across sessions
<b>Student Recruitment via Teachers and Staff</b>					
Initial contact with staff: e-mail + flyer distribution	0.167	30	5		
<b>Sub-Total</b>		<b>30</b>	<b>5</b>		
<b>Parent and Student Recruitment</b>					
Flyer and consent form review	0.08	200	16		
Consent form completion and return	0.13	100*	13		
<b>Sub-Total</b>		<b>200</b>	<b>29</b>		
<b>Recruitment Totals</b>		<b>230</b>	<b>34</b>		
<b>Individual Interviews Sessions</b>					
Confirmation to parent via email or letter	0.05	72**	3.6	3.5***	12.6
Grade 4	1	24	24	3.5***	84
Grade 8	1	24	24	3.5***	84
Grade 12	1	24	24*	3.5***	84
<b>Interview Totals</b>		<b>72</b>	<b>75.6</b>	<b>14</b>	<b>264.6</b>
<b>Total Burden</b>		<b>Respondents: 302</b>	<b>Hours: 298.6</b>		

\* Subset of initial contact group

\*\* Subset of total recruitment

\*\*\* Each student would participate in half (3 or 4) of the total 7 sessions.

The combined totals for the background questionnaire and SICT science pretest activities are listed in table 6.

**Table 6. Combined Burden for Background Questionnaire and SICT Science Pretest Activities**

	<b>Number of respondents</b>	<b>Number of responses</b>	<b>Burden Hours</b>
Total Background Questionnaire Burden	470	660	97.25
Total SICT Burden for Cognitive Interviews	1,748	2,285	350.25
Total SICT Burden for Play Testing	302	834	298.60
<b>Overall Totals</b>	<b>2,520</b>	<b>3,779</b>	<b>746.10</b>

## 9) Estimate of Costs for Recruiting and Paying Respondents

Each participating student will receive a \$25 gift card for compensation for his or her time and effort. In addition, NCES is offering a gift card of \$25 for a parent or legal guardian to remunerate him or her for the time involved and to help offset the travel/transportation costs of taking the participating student to and from the cognitive laboratory or play testing site. Each participating teacher and school administrator being interviewed will receive a \$40 gift card for compensation for his or her time and effort. A generic gift card (e.g., Visa gift card) that can be used anywhere credit cards are accepted is the recommended incentive.

## 10) Costs to Federal Government for Cognitive Interviews and Play Testing

The estimated costs for the activities described in this package are described below.

**Table 7. Estimate of Costs**

<b>Activity</b>	<b>Provider</b>	<b>Estimated Cost</b>
<b>Background Questionnaire Cognitive Interview Activities</b> <ul style="list-style-type: none"> <li>• Design, preparation, and analysis for background questionnaire cognitive interviews</li> <li>• Preparation and conduct of background questionnaire cognitive interviews (including recruitment, incentive costs, data collection, analysis, and reporting) for science pretest activities</li> </ul>	ETS	\$ 35,000
	EurekaFactors	\$ 227,000
<b>SICT Play Testing Activities</b> <ul style="list-style-type: none"> <li>• Design, preparation, and conduct of play testing sessions (including recruitment, incentive costs, data collection, analysis, and reporting) for science pretest activities</li> </ul>	ETS	\$ 44,000
<b>SICT Cognitive Interview Activities</b> <ul style="list-style-type: none"> <li>• Design, preparation, and analysis for SICT cognitive interviews</li> <li>• Preparation and conduct of SICT cognitive interviews (including recruitment, incentive costs, data collection, analysis, and reporting) for science pretest activities</li> </ul>	ETS	\$ 25,000
	EurekaFactors	\$ 516,885
<b>Totals</b>		<b>\$ 847,885</b>

## 11) Schedule

Table 8 provides the schedule of milestones and deliverables for the science background questionnaire and SICT pretest activities. Note that the development of the SICTs is iterative. Therefore, the play testing and cognitive interview activities will take place at numerous intervals throughout the time period.

**Table 8. Schedule of Milestones and Deliverables**

<b>Activity</b>	<b>Dates</b>
<b>Background Questionnaire Cognitive Interviews</b>	
Recruit participants (subsequent to OMB clearance)	Nov 2012-April 2013
Data collection, preparation, and coding	Nov 2012-April 2013
Data analysis	Nov 2012-June 2013
Final study report	June 2013
<b>SICT Play Testing and Cognitive Interviews</b>	
Recruit participants (subsequent to OMB clearance)	Nov 2012-July 2014
Data collection, preparation, and coding	Nov 2012-July 2014
Data analysis	Nov 2012-July 2014
Final study report	August 2014