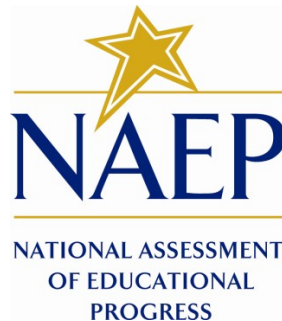# NATIONAL CENTER FOR EDUCATION STATISTICS
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

Volume I
Supporting Statement

NAEP Pretesting of Survey and Cognitive Items for Pilot in 2017 and 2018

OMB# 1850-0803 v.146

October 13, 2015

Table of Contents

1) **Submittal-Related Information**

This material is being submitted under the generic National Center for Education Statistics (NCES) clearance agreement (OMB #1850-0803). This generic clearance 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 assessment instruments. This request is to test new content for upcoming assessments through cognitive interviews, playtesting, and tryouts.

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 survey 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 survey) are pretested on a small number of respondents before they are administered to a larger sample through pilot or operational tests. These pretest activities can include playtesting and cognitive interviews, as well as tryouts of items, as defined later in this section. Pretesting helps us identify and eliminate, as much as possible, problems with items before those items are used in large-scale formal pilots. This, in turn, means fewer challenges in scoring and analysis, higher pilot item survival rates, less revisiting of test design, and therefore time efficiencies gained in operationalizing items.

This submittal requests clearance for various pretesting activities related to the upcoming assessments:

- Cognitive interviews for the 2018 core, civics, geography, and U.S. history, and 2019 reading and mathematics survey questions with students at grade 8, grade 8 teachers, and grade 8 school administrators (specifically principals), all to be pilot tested in 2017;

- Playtesting, cognitive interviews, and small-scale tryouts for the 2019 mathematics and reading/ELA selected cognitive items with students at grades 4 and 8, to be pilot tested in 2017; and

- Playtesting, cognitive interviews, and small-scale tryouts for the 2019 mathematics and reading/ELA selected cognitive items with students at grade 12, to be pilot tested in 2018.

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 aforementioned assessments;

- Appendices A-BV — recruitment and communication materials;

- Appendices BW-CQ — screeners and consent forms; and

- Volume II — protocols and questions used in the pretesting sessions.

Survey Questions

The 2018 core, civics, geography, and U.S. history, and 2019 reading and mathematics survey questionnaires aim to capture data related to important subject-specific (i.e., civics, geography, U.S. history, reading, and mathematics) and non-subject-specific (core) contextual factors for student achievement. Table 1 contains the possible areas of focus for questionnaire development for the upcoming NAEP survey questionnaires.

Table 1. Core Modules and Civics, Geography, U.S. History, Reading, and Mathematics Issues

	Core	**Civics, Geography, U.S. History**	**Reading**	**Mathematics**
Module 1/Issue 1	Socio-Economic Status (SES)	Resources for Learning and Instruction	Resources for Learning and Instruction	Resources for Learning and Instruction
Module 2/Issue 2	Technology Use	Organization and Instruction	Organization and Instruction	Organization and Instruction
Module 3/Issue 3	Grit	Teacher Preparation	Teacher Preparation	Teacher Preparation
Module 4/Issue 4	Desire for Learning	Student Factors	Student Factors	Student Factors
Module 5/Issue 5	School Climate	n/a	n/a	n/a

The main purposes of the cognitive interviews are to:

1. Identify problems with the items (i.e., ensure the item is understood by all participants, and confirm that items are not sensitive in nature or make the participant uncomfortable); and

2. Find ways to improve wording of existing items where possible.

Mathematics and Reading Cognitive Items

As NAEP moves forward with digital-based assessments, pretesting is especially important given unknown factors associated with innovative digitally-based items. A range of pre-pilot testing tools allows tailoring the selected approach to the specific question or purpose to be addressed. In NAEP, pretesting methods have often been used at different stages of development. For example, playtesting has mostly been conducted in early item development stages using mockups/wireframes, while cognitive interviews and tryouts, when needed, typically have occurred at the draft programmed stages. Selected technology-enhanced discrete items (such as drag-and-drop or graph creation) and dynamic reading passages with items will be pretested.

The primary focus of pretesting activities is on the measurement of content and cognitive skills and abilities, not software and interface usability. As such, the focus of this pretesting will be to determine whether items elicit the targeted knowledge and skills and whether any item content or presentation causes confusion or introduces construct-irrelevant variance.

Types of Pretesting

The following sections describe the different types of pretesting that will be used.

Playtesting (used in pretesting the cognitive items)

In playtesting, an innovation adapted from the game-design industry, a diverse set of students, individually or in small teams of two to four, will work through and discuss mockups of sets of

technology-enhanced items with the observer/facilitator or with one another and an observer/facilitator. Playtesting may take place early in the process using wireframes (somewhat functional storyboards for items) or programmed builds. Additionally, playtesting may be used in lieu of cognitive labs at a later stage of development, depending on item features and the questions that need to be answered. The main purposes of playtesting are to gather student reactions to early versions of interactive items and to begin to understand, in an informal way, how students are thinking about those items.

During playtesting, students will be encouraged to talk together about items and issues they confront, while observers note reactions to and potential problems with content or format. Observers will query students to draw them out, facilitate deeper reactions, or probe areas of possible confusion. Playtesting will allow identification of construct-irrelevant features in items (or stimuli such as passages), for example inaccessible language in item stems or uninteresting or unfamiliar scenarios that result in poor student engagement. Playtesting early in the development cycle allows for item refinements that can be tested in subsequent pretesting activities, for example tryouts.

Cognitive Interviews (used in pretesting the cognitive and survey items)

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 drawing on methods from cognitive science. The objective is to explore how students are thinking and what reasoning processes they are using as they work through tasks. In NAEP studies to date, two methods have been used, either separately or combined: 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 work through questions or tasks. With verbal probing techniques, the interviewer asks probing questions, as necessary, to clarify points that are not evident from the "think-aloud" process, or to explore issues that have been identified a priori as being of particular interest. In the current studies, verbal probing techniques only will be used for the survey items while the combination of think-aloud interviewing and verbal probing will be used for the cognitive items.

Cognitive interview studies produce largely qualitative data in the form of verbalizations made by students during the think-aloud phase or in response to the interviewer probes. Some informal observations of behavior are also gathered, since typically a second observer is present, in addition to the interviewer. Behavioral observations may include such things as nonverbal indicators of affect, suggesting emotional states such as frustration or engagement, and interactions with the task, such as ineffectual or repeated actions suggesting misunderstanding or usability issues.

In addition to think-aloud and verbal probing techniques, eye tracking methodology may be used during cognitive interviews for the cognitive items. Eye-trackers use an infrared video image of the eyes to calculate gaze location in real-time, so that it is possible to see where on the screen the student is looking at any given moment. Using this methodology, the student's gaze is tracked as he or she works through an activity. Eye tracking is useful for examining how students are focusing their attention when moving among texts and non-text stimuli for items across subjects. Eye tracking methods are also useful for examining patterns of students' attention to and processing of non-interactive stimuli, during which no information is being obtained from the student via button presses or other student-driven manipulations or actions in the environment (i.e., items during which the moment-by-moment logging of student actions will yield little direct evidence of students' cognition). For example, when viewing a text, video, or

photograph, only limited evidence of students' processing may be obtained via student action logs (e.g., overall time spent reading or viewing, responses to and timing data for associated questions). With the addition of eye tracking methodology, we can obtain an evidentiary trace of students' cognitive processing during these "non-interactive" components (e.g., to what extent do students attend to or process the most important or salient parts of the text, video, or photograph). This evidence could help to identify sources of construct-irrelevant variance or other issues with processing items, which might inform item revisions.

<u>Small-Scale Tryouts (used in pretesting the cognitive items)</u>

In tryouts, students will work uninterrupted through a selected set of draft programmed items. Tryouts provide a small-scale snapshot of the range of responses and actions items are meant to elicit, but which can be gathered much earlier in the assessment development process and with fewer resource implications than formal piloting.

Eye tracking may be used with a small subset of the tryout sample. For example, eye tracking may be desirable if we are testing dynamic passages requiring navigation and we wish to see how well students are negotiating such texts and how their attention is focused. Note that if eye tracking is used during tryouts, it would not change the test-taking conditions—students' gaze patterns would simply be tracked inconspicuously, while they complete items using exactly the same procedure as the remainder of the tryout sample, but in a separate room.

3) Sampling and Recruitment Plans

The sampling and recruitment plans, which differ by the type of testing, are described below.

Cognitive Interviews of Survey Questions

ETS is the survey questionnaire developer for NAEP and will be responsible for the overall conduct and management of the cognitive interview activity described in this package. EurekaFacts, subcontractor to ETS, will conduct the cognitive interviews (see Section 5).

Various resources will be employed to recruit participants. For students, these will include:

- existing participant databases;
- targeted telephone and mail contact lists (i.e., lists that consist of individuals meeting basic criteria such as age or school grade);
- school system research/assessment directors;
- NAEP State Coordinators (see section 5) when possible, to recruit in schools;
- community organizations (e.g., Boys/Girls clubs, Parent-Teacher Associations, and limited on-site location-based and mass media recruiting); and
- outreach/contact methods and resources (e.g., internet ads, flyers/bookmarks, canvassing, and having representatives available to talk to parents, educators, and community members at appropriate local community events, school fairs, etc.)

Teachers and school administrators will be recruited using the following recruitment resources, in

addition to those mentioned above:

- national organizations' databases of administrators and faculty;

- NCES school database (e.g., Common Core of Data and Private School Universe Survey);

- contacts within organizations and groups that can serve as recruitment partners (e.g., Horton's Kids, Housing Authority of the City of Frederick); and, if needed

- targeted contact lists.

EurekaFacts will recruit participants so that a diverse sample is achieved. Specifically, grade 8 students will include a mix of gender, race/ethnicity, urban/suburban/rural, and socioeconomic (SES) backgrounds; and teachers and school administrators will be from a mix of school sizes and school socioeconomic demographics. Please note, while SES will be given a higher priority than other respondent characteristics when recruiting, sufficient balance of all other criteria will be ensured.

EurekaFacts will recruit potential participants in urban areas such Washington, D.C. and Baltimore, MD, as well as suburban and rural areas in Maryland and Virginia. In addition to the aforementioned areas, EurekaFacts may also recruit teachers and school administrators in other states represented in their database. No more than three students will be recruited per school. No more than one teacher or school administrator will be recruited per school.

To minimize the travel burden of students, parents/guardians, teachers, and school administrators, cognitive interviews will be conducted in nearby venues that are convenient for the participants, such as EurekaFacts offices in Rockville, MD, community centers, facilities of community-based organizations, and school building sites (after school only). Before conducting any interviews in school building sites, ETS, the school principal, and the NAEP State Coordinators will be notified to confirm approval. In addition, a limited number of teacher and school administrator interviews may be conducted via phone or over the internet, if needed.

One and a half hour (90 minutes)[1] cognitive interviews will be conducted with students, teachers, and school administrators. Participants will receive core, civics, geography, U.S. history, reading, and mathematics items. All student cognitive interviews and the majority of teacher and school administrator cognitive interviews will be conducted in-person.[2]

The recruitment process includes:

- EurekaFacts sends an email of introduction about the cognitive interview research, including flyers, an information brochure, and informational bookmarks (see Appendices A-H).

- After receiving a contact of interest, a EurekaFacts staff member will follow up with the parent/guardian, teacher, or school administrator via phone (see Appendices BW and BX), and ask them to provide demographic information to ensure that a diverse sample is selected as per the aforementioned criteria.

- If the parent/guardian allows their student to participate, and the teacher and school administrator agree to participate, EurekaFacts will follow up to confirm participation and the date and time of

---

[1] Please note that the 90 minutes includes time for introductions (maximum 15 minutes), conducting the interview (60 minutes), and debriefing and/or time for additional questions/feedback from the participants (maximum 15 minutes).
[2] A limited number of teacher and school administrator interviews may be conducted via phone or over the internet, if needed.

the cognitive interview session (see Appendices I and J).

- Parents/guardians, teachers, and school administrators will be required to sign informed consent forms prior to the cognitive interview session (see Appendices BY and BZ).

- Students, teachers, and school administrators with a signed consent will be asked to participate in cognitive interviews that may last up to 90 minutes. After participating in the cognitive interview, students, parents/guardians (only if they provided transportation to and from the cognitive interview), teachers, and school administrators will receive their incentive (see Section 9) and be sent a thank you letter/email (see Appendices K and L).

Playtesting of Cognitive Items

ETS will recruit students from a range of demographic groups. Students will be recruited from districts that are located near the ETS Princeton, New Jersey campus for scheduling efficiency and flexibility.

ETS will recruit students using existing ETS contacts with teachers and staff at local schools and afterschool programs. E-mail or letters will be used to contact these teachers/staff, and paper flyers and consent forms for students and parents will be distributed through these teachers/staff. 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. Confirmation e-mails and/or letters will be sent to participants. Only after ETS has obtained written consent from the parent/guardian will the student be allowed to participate in the playtesting session. After participating in the session, students and parents/guardians (only if they provided transportation to and from the playtesting session) will receive their incentive (see Section 9) and be sent a thank you letter/email (see Appendices M-W and CA-CE for materials used for the playtesting sessions).

Five students will be convened per item set or draft block. Five students per grade should be sufficient at the playtesting stage given that the key purpose is to identify usability errors and other construct-irrelevant issues.[3] Playtesting group sizes are too small to reflect a nationally representative sample. We will make every effort, however, to include a diverse group representing a mix of gender, race/ethnicity, socioeconomic background, and urban and suburban students. A maximum of 60 students will participate in playtesting, across the three grades and two subjects. Playtesting session will be 60 minutes for grade 4 students and 90 minutes for grade 8 and 12 students. Note, based on prior experience with similar studies, it is anticipated that some students will participate in multiple sessions.

Cognitive Interviews of Cognitive Items

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.[4] 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

---

[3] 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.

[4] 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://www.measuredprogress.org/documents/10157/18820/cognitiveinterviewmethods.pdf

subgroup for analysis for the purposes of exploratory cognitive interviewing.[5]

Accordingly, five to seven students per set of items should be sufficient given that the key purpose of the cognitive interview is to identify qualitative patterns in how students think at different points when doing items. Given the number of items to be developed, a maximum of 84 students across grades 4, 8, and 12 for mathematics and reading will participate in cognitive interviews. The interviews will be 60 minutes for grade 4 and 90 minutes for grades 8 and 12.

Cognitive interviews may be conducted by ETS or EurekaFacts. Students will be recruited from the following demographic populations:

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

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

For any cognitive interviews conducted by ETS, students will be recruited from districts that are located near the ETS Princeton, New Jersey campus for scheduling efficiency and flexibility. For any cognitive interviews conducted by EurekaFacts, students will be recruited from the District of Columbia, Maryland, Virginia, Delaware, and Southern Pennsylvania. EurekaFacts will also conduct interviews in other venues beside their Rockville, Maryland site, such as after-school activities organizations or community-based organizations. This allows them to accommodate participants recruited from areas other than Rockville, MD and ensure that the sample population is representative of different geographical areas (i.e., urban, rural, and suburban). In all cases, a suitable environment (i.e., a quiet room) will be used to conduct the interviews and there will be more than one adult present.

As with playtesting, 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/staff at local schools and afterschool programs. Paper flyers and consent forms for students and parents will be distributed through these teachers and staff contacts. 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 W-AC and AG).

While EurekaFacts will use various outreach methods to recruit students to participate, 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 outreach/contact methods and resources such as newspaper/Internet ads, outreach to community organizations (e.g., Boys and Girls Clubs, Parent-Teacher Associations), social media, and mass media recruiting (such as postings on the EurekaFacts website). Interested participants will be screened to ensure that they meet the criteria for participation in the tryout (e.g., their parents/guardians have given consent and they are from the targeted demographic groups outlined above). When recruiting participants, EurekaFacts staff will first speak to the parent/guardian of

---

[5] 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

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 AH-AS, AW-BA, and CH-CJ).

For both sessions conducted by ETS and EurekaFacts, 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. Informed parental consent will be obtained for all respondents who are interested in participating in the data collection efforts. Eye tracking may be included for some of the interviews. After participating in the session, students and parents/guardians (only if they provided transportation to and from the interview session) will receive their incentive (see Section 9) and be sent a thank you letter/email (see Appendices AD-AF, AT-AV, CF-CG, and CK-CL).

Tryouts of Cognitive Items

EurekaFacts will perform the recruiting for tryouts from the District of Columbia, Maryland, Virginia, West Virginia, Delaware, and Southern Pennsylvania. This recruitment area will allow us to ensure the results are representative of various populations, and specifically, inclusion of students from rural areas. Students will be sampled to obtain the following criteria:

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

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

While EurekaFacts will use various outreach methods to recruit students to participate, 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 outreach/contact methods and resources such as newspaper/Internet ads, outreach to community organizations (e.g., Boys and Girls Clubs, Parent-Teacher Associations), social media, and mass media recruiting (such as postings on the EurekaFacts website).

Interested participants will be screened to ensure that they meet the criteria for participation in the tryout (e.g., their parents/guardians have given consent and they are from the targeted demographic groups outlined above). 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. 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. Informed parental consent will be obtained for all respondents who are interested in participating in the data collection efforts. After participating in the session, students and parents/guardians (only if they provided transportation to and from the interview session) will receive their incentive (see Section 9) and be sent a thank you letter/email (see Appendices BB-BV and CM-CQ).

A maximum of 300 students will be recruited for small-scale tryouts across grades 4, 8, and 12 for mathematics and reading. The tryouts will be limited to 60 minutes.

Table 2 summarizes the numbers and types of participants that are planned for all of the pretesting activities described in this package

Table 2. Sample Size: Pretesting Activities [6]

	Grade 4	Grade 8	Grade 12	Total
Cognitive Interviews – Survey Questions				
Students	N/A	30	N/A	30
Teachers	N/A	10	N/A	10
School Administrators	N/A	5	N/A	5
Pretesting – Cognitive Items				
Playtesting	20	20	20	60
Cognitive Interview	28	28	28	84
Tryouts	100	100	100	300
Total	**148**	**193**	**148**	**489**

4) Data Collection Process

The data collection process and analysis plans, which differ by the type of testing, are described below.

Cognitive Interviews of Survey Questions

Participants will first be welcomed, introduced to the interviewer and the observer (if an in-room observer is present), and told that they are there to help us ensure that students/teachers/administrators like them understand the newly developed core, civics, geography, U.S. history, reading, and mathematics items (see Volume II, Part B). Participants will be reassured that their participation is voluntary and that their responses will be used for research purposes only (see Section 6). As part of the introduction process, the interviewer will explain to participants that their responses will be audio recorded. For the phone/web-based teacher and school administrator cognitive interviews, the interviewer will explain the technology and describe the tools the participants may use, such as muting their phone and asking questions.

The interviewer will be tasked with keeping participants engaged by asking probe questions (see Volume II), soliciting responses from less talkative participants, and asking follow-up questions where appropriate (e.g., "*That's interesting, could you tell me a little bit more about that?*") Interviewers may also take additional notes during the in-person cognitive interviews, including behaviors (e.g., the participant's facial expressions indicated he/she is confused) and if extra time was needed to answer certain questions. See Volume II for the specific protocols and item probes for the survey questions being pretested.

Analysis Plan

After the session, the notes and audio recording will be summarized to report main findings and provide illustrative statements that will be analyzed by the NAEP questionnaire development team. The cognitive interview results will be used to help inform which items should be administered during the 2017 pilot test.

---

[6] For the cognitive items pretesting, this table represents the expected distribution across grades. Depending on the nature of the items and tasks and the specific recruitment challenges, the actual distribution may slightly vary. For burden purposes, the maximum number of students by pretesting activity will not exceed the total shown in the table.

Playtesting of Cognitive Items

Playtesting will take place in a range of locations so that staff can maximize opportunities to work with students. Depending on scheduling and participants, some could take place at ETS, some in schools, and some at organizations from which students will be drawn (e.g., Boys and Girls Clubs).

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 research purposes (see Section 6). Assessment specialists will then give an overview of the items to students and provide guidance on what they should reflect on while looking at the tasks. Assessment specialists and other staff (e.g., cognitive scientists or task designers) from ETS will act as facilitators and observers, taking notes on what students say and interjecting occasional questions aimed at eliciting students' reactions, places of confusion, and ways of thinking about the answers to the questions in the tasks and/or items. Each observer may choose to stay with an individual or one group of two to three students looking at and responding to items, or they may choose to move around to observe several groups or other individual students.

For the most part, students will be allowed to explore the mocked-up or programmed items 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, such as:

- *Did you find this question/passage 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?*
- *How could this item be improved? How could it be clearer?*

Prior to each playtesting session, ETS staff may identify some key focus areas for each item, set, or block of items. If students do not provide sufficient comments on targeted items or issues, a staff member may ask a group of students if they had any thoughts about the particular item or issue using questions such as those described above. Sessions will be recorded with audio so that observers are able to verify the accuracy of notes.

Analysis Plan

Student feedback from a playtesting session is immediate and can be evaluated after the session. Notes from the observers in each session will be aggregated; one aggregate document will be produced for each set of items that are observed, with all observers contributing their observations to this common document. Since playtesting is a more informal process that generates relatively unstructured information, no formal analyses of these data will be performed.

Cognitive Interviews of Cognitive Items

Participants will first be welcomed by staff, introduced to the interviewer and the observer, and told they are there to help answer questions about how people answer mathematics and reading questions. Students will be reassured that their participation is voluntary and that their answers may be used only for research purposes (see Section 6). Interviewers will explain the think-aloud process and conduct a practice session with a sample question. The think-aloud component of the cognitive interviews will be a concurrent think-

aloud method, in which the student verbalizes his or her thoughts while working through the items. As reading entails verbal processing, thinking aloud can be challenging during reading of the text. Thus we will have students focus verbalization efforts when they are in the process of answering items, rather than reading passages. Eye tracking may be used along with concurrent thinking aloud, especially for reading.

The protocols (see Volume II, part C) for the think-aloud sections contain largely generic prompts to be applied flexibly by the interviewer to facilitate and encourage students in verbalizing their thoughts. 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?*" The think-aloud method also includes a verbal probing component (see Volume II) conducted after completion of the think-aloud portion for a given item or set of items. These verbal probes include a combination of pre-planned questions, identified before the session as important, and ad hoc questions that the interviewer identifies as important from observations during the interview, such as clarifications or expansions on points raised by the student. For example, if a student paused for a long time over a particular item, appeared to be frustrated at any point, or indicated an 'aha' moment, the interviewer might probe these kinds of observations further, to find out what was going on. To minimize the burden on the student, efforts are made to limit the number of verbal probes that can be used in any one session or in relation to any set of items.

Interactions and responses may be recorded via video screen-capture software (e.g., Morae® software by *TechSmith*). These recordings can be replayed for later analysis, to see how a given student progressed through the task. Digital audio recording will capture students' verbal responses to the think-aloud interview, using either the tablet's integral microphone or an external digital recorder, depending on the specific tablet platform used and compatibility with the screen-capture software. Interviewers will also record their own notes separately, including behaviors (e.g., the participant appeared confused) and whether extra time was needed during a particular part of the task.

Analysis Plan

For the cognitive interview data collections, documentation will be grouped at the discrete item, set, or block level. Items will be analyzed across participants.

The types of data collected about task items and components will include:

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

The general analysis approach will be to compile the different types of data to facilitate identification of patterns of responses for specific items or groups of items: for example, patterns of responses to probes or debriefing questions, common patterns of looking at particular screens (in cases where eye movements are captured), or types of actions observed from students at specific points in a given item or set of items.

This overall approach will help to ensure that the data are analyzed in a way that is thorough, systematic, and that will enhance identification of problems with items and provide recommendations for addressing those problems.

Tryouts of Cognitive Items

EurekaFacts will conduct tryouts at their Rockville, Maryland site or another suitable venue (e.g., after-school activities organization, or community-based organization). Tryout sessions will be conducted by EurekaFacts in small groups. Because tryouts are sessions where the students complete items on their own without any interruption, verbal probing, or think-aloud component, it is possible and most efficient to have several students complete items at the same time. A proctor will be present during the session and will follow a strict protocol (see Volume II) to provide students with general instructions, guide the group through the tryout, administer debriefing questions, and assist students in the case of any technical issues. In addition, the proctor will take notes of any potential observations or issues that occur during the tryout session.

Analysis Plan

Student responses to items will be compiled into spreadsheets to allow quantitative and descriptive analyses of the performance data. Completion times and non-completion rates will also be quantified and entered into the spreadsheets. These data sets will be used in item development, design, and programming decisions.

5) Consultations Outside the Agency

Educational Testing Service (ETS) serves as the NAEP Item Development contractor. As such, ETS will be responsible for the management of all activities described in this package, as well as conducting the cognitive item playtest and possibly some of the cognitive item cognitive interviews.

EurekaFacts is a small, established research and consulting firm in Rockville, Maryland. EurekaFacts offers facilities, tools, and staff to collect and analyze both qualitative and quantitative data. EurekaFacts is working as a subcontractor to ETS to conduct the cognitive item small-scale tryouts and possibly some of the cognitive item cognitive interviews.

The NAEP State Coordinator serves as the liaison between the state education agency and NAEP, coordinating NAEP activities in his or her state. NAEP State Coordinators from selected states may provide leads for potential participants for this study.

6) Assurance of Confidentiality

Participants are 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 consent will be obtained from participants who are over the age of 18 and from parents or legal guardians of students who are under the age of 18. 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. The consent forms,

which include the participant name, will be separated from the participant interview files, secured for the duration of the study, and will be destroyed after the final report is completed.

The interviews will be recorded.[7] The only identification included on the files will be the unique ID assigned to each participant by the interviewer. 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 and task development process, as well as the process of developing interview protocols, effort has been made to avoid asking for information that might be considered sensitive or offensive. Reviewers have also attempted to identify and minimize potential bias in questions.

8) Estimate of Hourly Burden

The burden estimates, which differ by the type of testing, are described below.

Cognitive Interviews of Survey Questions

The estimated burden for recruitment assumes attrition throughout the process.[8] All cognitive interviews will be scheduled for no more than 90 minutes. Table 3 details the estimated burden for the survey questionnaire cognitive interview processes.

Playtesting of Cognitive Items

The estimated burden for recruitment assumes attrition throughout the process.[9] Play testing sessions are expected to last 60 minutes for grade 4 students and 90 minutes for grade 8 and 12 students. Table 4 details the estimated burden for the cognitive items playtesting.

Cognitive Interviews of Cognitive Items

The estimated burden for recruitment assumes attrition throughout the process.[10] All cognitive interviews will be scheduled for no more than 60 minutes for grade 4 students and 90 minutes for grade 8 and 12 students. Table 5 details the estimated burden for the cognitive item cognitive interview processes.

Tryouts of Cognitive Items

The estimated burden for recruitment assumes attrition throughout the process.[11] All tryouts will be scheduled for no more than 60 minutes. Table 6 details the estimated burden for the cognitive item cognitive interview processes.

---

[7] Details regarding the nature of the recordings are described in the specific interview sections.
[8] Assumptions for approximate attrition rates for direct participant recruitment are 33 percent from initial contact to follow-up, 50 percent from follow-up to confirmation. Note: The initial principal contact for student identification attrition rate is 33 percent from contact to follow-up.
[9] Assumptions for approximate attrition rates are 50 percent from initial contact (flyer from teacher) to consent form completion and 25 percent from submission of consent form to participation. Note: The initial principal contact for student identification attrition rate is 33 percent from contact to follow-up.
[10] Assumptions for approximate attrition rates are 50 percent from initial contact (flyer from teacher) to consent form completion and 25 percent from submission of consent form to participation.
[11] Assumptions for approximate attrition rates are 50 percent from initial contact (flyer from teacher) to consent form completion and 25 percent from submission of consent form to participation.

Table 3. Hourly Burden for Students, Teachers, and School Administrators for Core, Civics, Geography, U.S. History, Reading, and Mathematics Survey Questions Cognitive Interviews

Respondent	Hours per respondent	Number of respondents	Total hours
Principal/School Administrator or Point Person for Community Organizations for Student Recruitment			
Initial contact	0.05	10	1
Follow-up & identify students	1.0	7*	7
Sub-Total		**10**	**8**
Parent or Legal Guardian for Student Recruitment			
Initial contact	0.05	135	4
Follow-up via phone	0.15	90*	14
Consent & confirmation	0.15	45*	4
Sub-Total		**135**	**28**
Teacher and School Administrator Recruitment			
Initial contact	0.05	69	3
Follow-up via phone or e-mail	0.15	46*	7
Consent & confirmation	0.15	23*	3
Sub-Total		**69**	**15**
Participation (Cognitive Interviews)			
Students	1.5	30**	45
Teachers	1.5	10*,**	15
School Administrators	1.5	5*,**	8
Sub-Total		**45**	**68**
Total Burden		**244**	**119**

Note: All totals have been rounded. In addition, some totals may differ slightly from sum of subtotals, due to rounding.

* Subset of initial contact group (total number of responses = 470)

** Estimated number of actual participants will be somewhat less than confirmation numbers.

Table 4. Hourly Burden for Reading and Mathematics Cognitive Items Playtesting

Respondent	Hours per respondent	Number of respondents	Total hours
Student Recruitment via Teachers and Staff			
Initial contact with staff: e-mail, flyer distribution, and planning	0.33	8	3
Sub-Total		**8**	**3**
Parent or Legal Guardian, and Student (18 or older)			
Flyer and consent form review	0.08	160	13
Consent form completion and return	0.13	80*	11
Confirmation to parent via email or letter	0.05	60*	3
Sub-Total		**160**	**27**
Participation (Playtesting)			
Grade 4	1	20	20
Grade 8	1.5	20	30
Grade 12	1.5	20	30
Sub-Total		**60**	**80**
Total Burden		**228**	**110**

Note: All totals have been rounded. In addition, some totals may differ slightly from sum of subtotals, due to rounding.

* Subset of initial contact group (total number of responses = 368)

Table 5. Hourly Burden for Reading and Mathematics Cognitive Items Cognitive Interviews

Respondent	Hours per respondent	Number of respondents	Total hours
Student Recruitment via Teachers and Staff			
Initial contact with staff: e-mail or phone, flyer distribution, and planning	0.33	12	4
Sub-Total		**12**	**4**
Parent or Legal Guardian, and Student (18 or older)			
Initial flyer review or phone screener	0.08	224	18
Consent form completion and return	0.13	112*	15
Confirmation to via email, letter, or phone	0.05	84*	4
Sub-Total		**224**	**37**
Participation (Cognitive Interviews)			
Grade 4	1	28	28
Grade 8	1.5	28	42
Grade 12	1.5	28	42
Sub-Total		**84**	**112**
Total Burden		**320**	**153**

Note: All totals have been rounded. In addition, some totals may differ slightly from sum of subtotals, due to rounding.

* Subset of initial contact group (total number of responses = 516)

Table 6. Hourly Burden for Reading and Mathematics Cognitive Items Tryouts

Respondent	Hours per respondent	Number of respondents	Total hours
Student Recruitment via Teachers and Staff			
Initial contact with staff: e-mail or phone, flyer distribution, and planning	0.33	40	13
Sub-Total		40	13
Parent or Legal Guardian, and Student (18 or older)			
Initial flyer review or phone screener	0.08	800	64
Consent form completion and return	0.13	400*	52
Confirmation to via email, letter, or phone	0.05	300*	15
Sub-Total		800	131
Recruitment Totals		840	145
Participation (Tryouts)			
Grade 4	1	100	100
Grade 8	1	100	100
Grade 12	1	100	100
Interview Totals		300	300
Total Burden		1,140	444

Note: All totals have been rounded. In addition, some totals may differ slightly from sum of subtotals, due to rounding.
* Subset of initial contact group (total number of responses = 1,840)

Total for All Pretesting Activities

The combined totals for all of pretesting activities are listed in Table 7.

Table 7. Combined Burden for Pretesting Activities

	Number of respondents	Number of responses	Burden Hours
Survey Questions			
Total Cognitive Interview Burden	244	470	119
Cognitive Items			
Total Playtesting Burden	228	368	110
Total Cognitive Interview Burden	320	516	153
Total Tryout Burden	1,140	1,840	444
Overall Totals	**1,932**	**3,194**	**826**

9) Estimate of Costs for Recruiting and Paying Respondents

For all student pretesting activities held outside of school hours, a $25 gift card (from a major credit card company) will be given to each student. If transportation is provided, the parent or legal guardian of each student will receive a gift card of $25 to thank them for their time and effort. Teacher and school administrator participants will be offered a $40 gift card (from a major credit card company) to thank them for taking part in the study.

10)   Costs to Federal Government

The estimated costs for the pretesting activities in this submittal are described in Table 8.

Table 8. Estimate of Costs

Activity	Provider	Estimated Cost
Design, preparation, analysis, and reporting for survey questionnaire cognitive interviews	ETS	$ 20,000
Preparation and conduct for survey questionnaire cognitive interviews (including recruitment, incentive costs, data collection and documentation)	EurekaFacts	$ 168,000
Design, prepare for, and conduct cognitive item playtesting sessions (including recruitment, incentive costs, data collection, and summary of findings)	ETS	$ 504,114
Design, prepare for, and conduct cognitive item cognitive interviews (including recruitment, incentive costs, data collection, analysis, and reporting)	ETS	$ 380,320
Prepare for, and conduct cognitive item cognitive interviews (including recruitment, incentive costs, data collection, analysis, and reporting).	EurekaFacts	$ 390,701
Design, prepare for, and conduct scoring and analysis of cognitive item tryouts.	ETS	$ 404,593
Prepare for and conduct cognitive item tryouts (including recruitment, incentive costs, data collection, reporting).	EurekaFacts	$ 392,587
Total		**$ 2,260,315**

11)   Schedule

Table 9 depicts the high-level schedule for the various activities. Each activity includes recruitment, data collection, analyses, and reports.

Table 9. High-Level Schedule of Milestones

Activity	Dates
Cognitive interviews for survey questions	November 2015 – February 2016
Pretesting for mathematics and reading cognitive items at grades 4 and 8	November 2015 – February 2016
Pretesting for mathematics and reading cognitive items at grade 12	October 2016 – February 2017