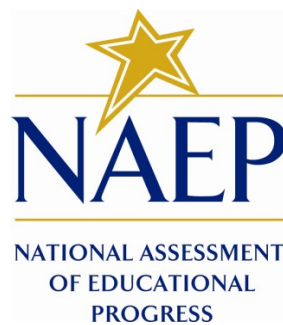# NATIONAL CENTER FOR EDUCATION STATISTICS
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

Volume I
Supporting Statement

*National Assessment of Educational Progress (NAEP)
Grade 8 Social Sciences Interactive Item Components
(IICs) Pretesting – Round 2*

OMB# 1850-0803 v.224

February 2018

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), which provides for NCES to conduct various procedures (such as pilot tests, cognitive interviews, and usability studies) to test new methodologies, question types, or delivery methods to improve survey and assessment instruments and procedures.

2) BACKGROUND AND STUDY RATIONALE

The National Assessment of Educational Progress (NAEP) is a federally authorized survey, by the National Assessment of Educational Progress Authorization Act (20 U.S.C. §9622), 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 conducted by NCES, which is part of the Institute of Education Sciences, within the U.S. Department of Education. NAEP's primary purpose is to assess student achievement in the different subject areas and collect survey questionnaire (i.e., non-cognitive) data to provide context for the reporting and interpretation of assessment results.

This request is to conduct, as part of the NAEP assessment development process, round 2 of pretesting activities including cognitive interviews and tryouts to collect data on recently developed interactive item components (IICs) for the 2020 grade 8 social science assessments (including civics, geography, and U.S. history). These IICs include an interactive timeline (U.S. History), a simulated web search (civics), a multimedia source container (civics and U.S. history), and a geographic information system (GIS) toolset (geography). In general, the focus of pretesting will be to investigate whether these IICs elicit the targeted knowledge and skills; whether any item content, interaction, or presentation causes confusion or introduces construct-irrelevant errors; and to gather information about how long students take to complete various IICs. A range of items using each new interactive item component will be included in the pretesting activities.

The initial round (round 1) of pretesting for the IICs was approved on May 22, 2017 (OMB #1850-0803, v. 197) and was conducted from August to September 2017. Cognitive interviews with 24 students from the Washington, D.C. metropolitan area were conducted to understand what reasoning processes students used as they worked through IICs and item sets. Tryouts were conducted with 75 students from the greater Washington, D.C. metropolitan area (including Washington, D.C., Maryland, and Northern Virginia) to collect information concerning students' thoughts about the broader IIC tasks and their experiences with the scenarios in the item sets, and to provide a reasonable sample of quantitative data on student performance, including timing data.

The pretesting findings (a) indicated that the IIC's elicit the targeted social science knowledge and skills and (b) provided valuable feedback on the effectiveness of item scoring guides. Findings also indicated that a number of software and user-interface updates were required to improve the student experience and reduce IIC load times in the assessment delivery system. Based on the pretesting findings, NAEP is implementing content, performance, and design changes for each IIC and, thus, an additional round of cognitive interviews and tryouts on the IICs is needed before piloting these items on a larger scale.

Consequently, this request is to conduct round 2 of pretesting activities on the grade 8 social science assessment IICs. The IIC revisions made based on the results of round 1 pretesting are not reflected in the ICR package materials because the assessment items, software, and interface are not subject to PRA. Given that round 2 will largely follow the recruitment and administration procedures used in round 1, the content of this request is very similar to that approved in May 2017 (OMB #1850-0803, v. 197) with only minor changes to reflect round 2 pretesting. An accompanying Changes Memo has been created to facilitate OMB's review and to provide a listing of the differences between the approved (v.197) ICR documents and those in this request, including an explanation of each revision.

The range of pretesting methods allows tailoring the selected approach to the purpose to be addressed. Cognitive interviews allow for the gathering of qualitative data about how students work through item sets and offer opportunities to probe potential sources of construct irrelevance. The larger samples and timed testing conditions of tryouts are especially useful for gathering quantitative data about programmed IIC timing and item performance before piloting, and for investigating the possible effects of the different features of the IICs on students' performances.

Pretesting is intended to enhance the efficiency of the development of assessment instruments. Before piloting, it

helps to identify and eliminate problems with items and tasks. This can mean fewer challenges in scoring and analysis and higher pilot item survival rates. Results of this pretesting will be used to finalize NAEP social sciences assessments for 8th grade to be piloted in 2020 and administered nationally in 2022.

3) RECRUITMENT AND DATA COLLECTION

COGNITIVE INTERVIEWS

<u>Overview</u>

In cognitive interviews (often referred to as a cognitive laboratory study or cog labs), an interviewer uses a structured protocol in a one-on-one interview drawing on methods from cognitive science. In NAEP studies to date, two methods have been 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. With verbal probing techniques, the interviewer asks questions, as necessary, to clarify points that are not evident from the "think-aloud" process, or to explore additional issues that have been identified a priori or during the process as being of particular interest. This combination of allowing the students to verbalize their thought processes in an unconstrained way, supplemented by specific and targeted probes from the interviewer, has proven to be flexible and productive.

Cognitive interview studies produce largely qualitative data in the form of verbalizations made by students during the think-aloud phase and/or in response to interviewer probes. The main objective is to explore how students are thinking and what reasoning processes they are using as they work through items and tasks. Some informal observations of behavior and verbalizations are also gathered; behavioral observations may include nonverbal indicators of affect, suggesting emotional states such as frustration or engagement, and interactions with tasks, such as prolonged time on one item or ineffectual or repeated actions suggesting misunderstanding.

Cognitive interviews may be conducted for draft programmed builds, based on processes developed at ETS. EurekaFacts, under a subcontract to ETS, will carry out the interviews. The general approach will be to have a small number of participants work individually through tasks. Data will then be synthesized in the form of lessons learned about students' thinking, observed student behaviors, and whether the task items appear to be eliciting the constructs of interest. These lessons will then inform ongoing task development.

<u>Sampling and Recruitment Plan</u>

Existing research and practice do not 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[1]. Nonetheless, a sample size of five to 15 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[2].

Accordingly, approximately 10 students per task should be sufficient given that the key purpose of the cognitive interview is to identify qualitative patterns in how students are reasoning at different points when doing tasks. Assuming that each student will be able to engage with two or three of the four IICs (ideally the source container would be pretested with both civics and U.S. history content), cognitive interviewing is expected to involve approximately 20 students. The total duration of the student cognitive interview session, including administrative activities, will be 90 minutes.

For the cognitive interviews, students will be recruited by EurekaFacts staff from the following demographic populations:

- students who are enrolled in eighth grade for the 2017-2018 school year;
- students who completed eighth grade during the 2016-2017 school year;

---

[1] 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

[2] 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

- students who represent a mix of race/ethnicity (Black, Asian, White, Hispanic);
- students who represent a mix of socioeconomic background; and
- students who represent 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 cognitive interviews from the District of Columbia, Maryland, Virginia, Delaware, and Southern Pennsylvania. EurekaFacts will also administer interviews in other venues besides their Rockville, MD site, such as after-school activity organizations or community-based organizations. This allows them to accommodate participants recruited from areas other than Rockville, MD and will help to obtain a sample population including different geographical areas (urban, suburban, and rural).

While EurekaFacts will use various outreach methods to recruit students to participate, the bulk of the recruitment will be administered by telephone and will be based on their acquisition of targeted mailing lists containing residential address and landline telephone numbers. 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/legal guardians have given consent, and they are from the targeted demographic groups outlined above). When recruiting participants (see Appendix P), EurekaFacts staff will first speak to the parent/legal guardian of the interested minor before starting the screening process. During this communication, the parent/legal 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 email/letter and phone call. Informed consent from parents/legal guardians will be obtained for all respondents who are interested in participating in the data collection efforts (see Appendix Q).

Data Collection Process

Cognitive interviews will take place at a range of suitable venues. In all cases, a suitable environment such as a quiet room will be used to administer the interviews, and there will be more than one adult present.

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 students do on social science tasks. Interviewers will explain the cognitive interview process and, to the extent that the think-aloud process is used, conduct a practice session with a sample question. An interviewer will ask students what they were thinking as they completed the questions and whether they believe the questions are clear and understandable. After the think-aloud process, students will be asked to answer a set of post-think-aloud follow-up questions (see Volume II). EurekaFacts staff may record audio and screen activity for analysis and take notes about students' reactions to these questions to revise and refine test content. No personal identifying information will be retained.

Protocols for cognitive interviews (see Volume II) will include probes to use as students work through item sets and after students finish answering items. Probes will include a combination of pre-planned questions 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 will be made to limit the number of verbal probes that can be used in any one session or in relation to any set of items. The welcome script, cognitive interview instructions, and hints for the interviewers are provided in Volume II.

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 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, such as behaviors (e.g., "the participant appeared confused"), questions posed by students, and observations of how long various items take to complete.

<u>Analysis Plan</u>

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

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

- student reactions and responses to items and components;
- behavioral data (e.g., observable actions recorded in interviewer notes, process data, and screen-captures);
- responses to generic questions;
- responses to targeted questions specific to the item(s);
- additional volunteered participant comments; and
- answers to debriefing questions.

The general analysis approach will be to compile the different types of data to facilitate identification of response patterns for specific items or item sets. Types of response patterns can include frequency counts of verbal report codes and responses to probes or debriefing questions, or student actions observed at specific points in a given item or item set. This overall approach will help to ensure that the data are analyzed in a way that is thorough, systematic, and will enhance identification of weaknesses with items and components and provide recommendations for addressing those problems.

SMALL-SCALE TRYOUTS

<u>Overview</u>

In tryouts, students will work uninterrupted through selected IIC draft item sets. These studies will be carried out by EurekaFacts, which will recruit participants, administer and observe the sessions, record interactions as appropriate, and report the results to ETS. Tryouts allow for pretesting of a wider range of content and the collection of more robust data on ranges of student responses, item difficulty, assessment timing, and alternative approaches to IIC functionality than is practical for cognitive interviews. Previous experience – for example with the social science discrete items, the Technology and Engineering Literacy assessment, and the reading assessment – shows that tryout-based insights are very informative.

<u>Sampling and Recruitment Plan</u>

EurekaFacts will use the same recruitment methods for tryouts as described in the cognitive interview section.

A maximum of 75 students will be recruited for small-scale tryouts across grade 8 civics, geography, and U. S. history. The IICs will be assembled into blocks and the number of students assigned per block will be based upon research goals currently under evaluation. The total duration of the student tryout session, including administrative activities, will be 90 minutes.

<u>Data Collection Process</u>

EurekaFacts will administer tryouts in small groups at their Rockville, MD site or another suitable venue (e.g., after-school activity organization or community-based organization). Tryout sessions will consist of small groups of approximately 10 students. Since students complete tryout session tasks on their own without any interruption, it is possible and most efficient to have several students complete tasks at the same time. Proctors will be present during the session and will follow a protocol to provide students with instructions. The proctor will take notes of any observations or issues that occur during the tryout session. After the tryout, students will be asked to answer a set of individual post-tryout follow-up questions and will participate in a brief post-tryout group discussion session regarding content and functionality (see Volume II).

<u>Analysis Plan</u>

The focus of tryout data may vary based on the IIC. However, score data and time to complete items will be captured and analyzed for all items. 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 to facilitate development, design, and programming decisions.

For the tryout data collection, documentation will be grouped at the item, item set, and block level. The items will be analyzed across participants.

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

- student responses to items;
- process data; and
- responses to group discussion probes.

The general analysis approach will be to compile data on the functionality of item components, for example: Do items function as intended? Do student responses provide measurable results? Are students able to effectively use the item components? What are the reactions of students to the item components? This approach will help to ensure that the data are analyzed in a way that will identify weaknesses with the items and components and provide recommendations for addressing those problems.

4) CONSULTATIONS OUTSIDE THE AGENCY

Educational Testing Service (ETS) is the Item Development, Data Analysis, and Reporting contractor for NAEP and will develop the IICs, analyze results, and draft a report with results. EurekaFacts, a research and consulting firm based in Rockville, MD, a subcontractor for ETS, will recruit participants and administer the cognitive interviews and tryouts.

5) JUSTIFICATION FOR SENSITIVE QUESTIONS

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

6) PAYING RESPONDENTS

To encourage participation, a $25 gift card from a major credit card company will be offered to each student who participates in a pretesting session as a thank you for his or her time and effort. For sessions that take place in locations other than schools, a parent or legal guardian of each student will also be offered a $25 gift card from a major credit card company to thank them for bringing their participating student to and from the testing site.

7) ASSURANCE OF CONFIDENTIALITY

The study will not retain any personally identifiable information. Prior to the start of the study, students will be notified that their participation is voluntary. As part of the study, students will be notified that the information they provide 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 (20 U.S.C. §9573 and 6 U.S.C. §151).

For all participants, written consent will be obtained from parents/legal guardians (of minor students) before interviews are administered. 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, secured for the duration of the study, and will be destroyed after the final report is released. Pretesting activities may be recorded using audio or screen capture technology. 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 completed.

8) ESTIMATE OF HOURLY BURDEN

The estimated burden for recruitment assumes attrition throughout the process. Assumptions for approximate

attrition rates are 50 percent from initial contact to consent form completion and 25 percent from submission of consent form to participation. Cognitive interviews and tryout sessions are expected to take 90 minutes for grade eight students.

Table 1. Estimate of Hourly Burden for Pretesting Activities

Respondent	Number of respondents	Number of responses	Hours per respondent	Total hours
Student Recruitment via Teachers and Staff				
Initial contact with staff: e-mail, flyer distribution, & planning	13	13	0.33	5
Parent or Legal Guardian				
Flyer and consent form review	254	254	0.08	21
Consent form completion and return	127*	127	0.13	17
Confirmation to parent via email or letter	95*	95	0.05	5
Recruitment Totals	**267**	**489**		**48**
Student				
Grade 8 Cognitive Interviews	20	20	1.5	30
Grade 8 Tryouts	75	75	1.5	113
Interview Totals	**95**	**95**		**143**
Total Burden	**362**	**584**		**191**

*Subset of initial contact group

Note: numbers have been rounded and therefore may affect totals

9) COST TO FEDERAL GOVERNMENT

The total cost of the study is $493,193 as detailed in Table 2.

Table 2: Cost to the Federal Government

Activity	Provider	Total: Estimated Cost-
Cognitive Interviews		
Design and prepare for cognitive interviews; analyze findings & prepare report	ETS	$106,501
Prepare for and administer cognitive interviews (including recruitment, incentive costs, data collection, analysis, & reporting)	EurekaFacts	$115,135
Tryouts		
Design and prepare for task tryouts; analyze findings and prepare report	ETS	$131,538
Prepare for and administer task tryouts (including recruitment, incentive costs, data collection, & reporting)	EurekaFacts	$140,019

10) PROJECT SCHEDULE

Table 3 provides the overall schedule.

Table 3: Round 2 Schedule

Activity *Each activity includes recruitment, data collection, and analyses*	Dates
Cognitive interviews	April-June 2018
Small-scale tryouts	April- June 2018
Pretesting report submitted	July 2018