UCLA | CRESST NATIONAL CENTER FOR RESEARCH ON EVALUATION, STANDARDS, AND STUDENT TESTING

EXPLORING CAREER-READINESS FEATURES IN HIGH SCHOOL TEST ITEMS THROUGH COGNITIVE LABORATORY INTERVIEWS

Ayesha Madni, Jenny C. Kao, Nichole M. Rivera, Eva L. Baker, and Li Cai

DECEMBER 2018

CRESST REPORT **857**

The work reported herein was supported under an Enhanced Assessment Grant, PR/Award #S368A150011, with funding to the California Department of Education (CDE), and subaward to the National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

The findings and opinions expressed here do not necessarily reflect the positions or policies of the California Department of Education or the U.S. Department of Education.

To cite from this report, please use the following as your APA reference: Madni, A., Kao, J. C., Rivera, N. M., Baker, E. L., & Cai, L. (2018). *Exploring career-readiness features in high school test items through cognitive laboratory interviews* (CRESST Report 857). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Table of Contents

Exploring Career-Readiness Features in High School Test Items Through Cognitive Laboratory Interviews[1]

Ayesha Madni, Jenny C. Kao, Nichole M. Rivera, Eva L. Baker, and Li Cai

CRESST/University of California, Los Angeles

Abstract: This report is the first in a series of five reports considering career-readiness features within high school assessments. Utilizing feature analysis and cognitive lab interviews, the primary objective of this study was to verify and validate the existence of specific career-readiness features in select math and English language arts (ELA) test items. Feature analysis is a method to characterize items and tasks by components, so that item design, revision, and instruction may benefit. Seventeen students representing three high schools in Southern California participated. Results from the preliminary feature analysis on four math and two ELA items indicated that each test item contained between eight and 13 career-readiness features. Results from the cognitive labs supported the presence of these features. That is, these features were either part of the problem-solving process, and/or were necessary to solve the problem correctly.

Introduction

Many high school graduates lack the knowledge, skills, and attributes necessary to compete and succeed in fields such as advanced manufacturing, energy, health care, information technology, and science, technology, engineering, and mathematics (Council of Chief State School Officers, 2013, 2014; O*NET Online, n.d.). Many high school graduates may also be unprepared to pursue higher education. Of those who do seek additional education, a significant portion require remediation in their first semester. Roughly eight in 10 students entering community college in California require at least one developmental course in math, English, or both (Mejia, Rodriguez, & Johnson, 2016). Students enrolled in developmental classes spend additional time and money bringing their skills up to college level. These students are also more likely to drop out before completing their certificate or degree. This report is the first in a series considering career readiness indicators in high school assessments.

[1]We thank Eric Zilbert of the California Department of Education for his support and guidance. We are also grateful to Noelle Griffin for advice and support. Special thanks to Brettany Clemens and Joanne Michiuye for help with data collection.

The competencies students acquire in high school are foundational for college and career success. It is important that high school curricula meet the needs of all students and potentially, employers. While graduation rates are up overall, some traditionally underrepresented groups are not performing as well, with graduation rates for Hispanic/Latino, American Indian/Alaskan native, and African American/Black students all falling below 80% (The California Department of Education, 2016). As of 2014, there has been a substantial shortage of qualified job candidates in many high-demand, high-skill, and high-wage career categories (Executive Office of the President of the United States of America, 2014). By 2020, it is expected that 65% of all jobs will require postsecondary education and training (Carnevale, Smith, & Strohl, 2013). This shortage may be linked, in part, to an underlying deficit in college and career readiness in the K-12 and postsecondary educational systems.

The state of California is a part of a consortium implementing Smarter Balanced assessments for standardized testing, ensuring students are ready for college success. An inclusive assessment development process with substantial input from higher education partners has ensured that the assessments have built-in qualities that enable college-readiness inferences. As a result, many colleges and universities have been using scores from Smarter Balanced as evidence that students are ready for credit-bearing entry-level college courses without the need for remediation (Smarter Balanced Assessment Consortium, 2015). Going beyond measuring mathematics and English language arts (ELA) content knowledge is beneficial, especially given the amount of resources dedicated to completing these assessments. Research that investigates features of existing assessments can allow for improvements to be made without increasing the burden to students, teachers, and school staff.

This report considers the results of a cognitive lab study conducted with high school students across three California high schools and preliminary descriptive feature analysis of select Smarter Balanced math and ELA practice items in the context of career readiness. More specifically, the study focuses on whether career-readiness features are present in the selected test items, as represented through students' problem-solving process and results. Feature analysis is a method to characterize items and tasks by components, so that item design, revision, and instruction may benefit. The four main components of the feature analysis process include feature rating, step-by-step analysis, cognitive labs, and quantitative analysis (Baker, Madni, Michiuye, Choi, & Cai, 2015). Assessment items or tasks are tagged with these qualitative features. Next, quantitative analyses of student performance are used to determine the relationships among features and assessment performance. The study uses both qualitative and descriptive analyses to derive conclusions.

The report provides insight into aspects of career readiness and cognitive processes that may be required for students to successfully complete select assessment items. This report also provides a preliminary description and qualitative analysis of the career-readiness features present in select Smarter Balanced practice items in math and ELA for Grade 11. In particular,

we were interested in what specific career-readiness features are evident in these test items, as well as whether they can be confirmed through cognitive lab interviews conducted with high school students. This report is organized as follows: (a) theoretical background leading to the cognitive lab approach; (b) theoretical background leading to the feature analysis approach; (c) an overview and description of target career-readiness features; (d) a description and overview of item selection and item types; (e) summary of results; and (f) conclusion, including a brief description of other reports in this series.

Cognitive Laboratory Interviews: Theoretical Background

Verbal data have been frequently used to study cognitive processes in multiple disciplines such as psychology, cognitive science, and education. Cognitive laboratory interviews (henceforth, cognitive labs) have origins in cognitive interviewing and protocol analysis. Cognitive interviewing is "a psychologically-oriented method for empirically studying the way in which individuals mentally process and respond to survey questionnaires" (p. 106; Willis, 2009, as cited by Leighton, 2017). Protocol analysis, the technique for examining think-aloud data, was developed and refined by Ericsson and Simon (1980, 1984, 1993). Since development, many articles and books have been published on collecting and reporting verbal data (see for example, Chi, 1997; Leighton, 2017; van Someren, Barnard, & Sandberg, 1994). In educational assessment, think-aloud methods have been used to design, develop, and validate test items as well as evaluate items for special populations (Johnstone, Bottsford-Miller, & Thompson, 2006; Zucker, Sassman, & Case, 2004). Recent work in neuroimaging provides evidence for a think-aloud protocol as a valid measure of thinking. Comparing multiple-choice responses and think-aloud protocols, Durning et al. (2013) found differences in brain activation associated with working memory. Previous CRESST work has also utilized think-aloud methods related to assessment items (Baker et al., 2015; Wolf, Kim, Kao, & Rivera, 2009).

Verbal reports are a type of qualitative data that come from think-aloud interviews and cognitive labs. As Padilla and Leighton (2017) noted in their chapter devoted to providing guidance conducting validity studies using cognitive labs and think-aloud methods, there are no clear methodological best practices. While some consider think-aloud interviews and cognitive labs to be interchangeable, Leighton (2017) argued that they differ in objectives and methodology. In a think aloud, participants are encouraged to verbalize their thoughts out loud while completing a task. Participants are asked not to explain what they are doing, but to simply verbalize their thoughts and actions while attending to the task (Ericsson & Simon, 1993). The goal of a think aloud is to gather information on the problem-solving process, and it relies heavily on *concurrent* reporting. The interview is typically audio recorded to facilitate verbatim transcription to produce a verbal report which can then be coded to a set of themes. In contrast, a cognitive lab focuses on comprehension processes, and can use either *concurrent* or *retrospective* reporting, or both. In retrospective reporting, participants are asked probing questions after they have completed a task (Leighton, 2017). Cognitive labs tend to have more

flexibility; the rigor of the procedures can vary depending on the objectives of the interview (Leighton, 2017).

Ericsson and Simon (1993) contended that think-aloud data collected concurrently allows for information to be gathered from a subject's short-term memory, which captures cognitive processes at the time they occur. Thoughts generated from long-term memory can be affected by perception and interpretation on the part of the subject. Verbalizations prompted by a question require the subject to comprehend the question, transform it to retrieval cues to select relevant information, and then put the retrieved information into a sequential form that allows the generation of a coherent series of verbalizations. The information from this process, Ericsson and Simon argued, might not be the same information they retrieved while actually performing the given task. Thus, they recommended encouraging subjects to verbalize with only neutral cues. However, they suggested that both concurrent and retrospective reports may be collected, as long as there is close correspondence between the concurrent and retrospective reports (between 2 and 10 seconds). Subjects can be expected to recall the actual sequence of their thoughts with high accuracy and completeness following a short duration.

One disadvantage to concurrent reporting, Branch (2000) argued, is that the cognitive load of problem solving while also speaking may be too challenging for some subjects. Concurrent verbal protocols may be less useful when the information is too difficult to verbalize or when the processes are automatic for the participants. Utilizing "Think Afters" can help mitigate these situations. However, since retrospective reporting may be influenced by forgetting and fabrication, Branch contended that both methods should be utilized to gather complete data.

Kuusela and Paul (2000) compared concurrent and retrospective data in order to reveal any qualitative differences. They noted that both types of data have advantages and disadvantages. Collecting concurrent data may both interfere with and facilitate information processing because of verbalizing choices during decision making. In addition, subjects have limited attentional capacity that may be directed away from the primary task during concurrent reporting, as described earlier. However, in retrospective reporting, subjects may verbalize thoughts and actions they perceive as more socially desirable. That is, these retrospective reports may contain a judgment or strategy that is more rational and refined. Results from Kuusela and Paul's study (2000) indicated that there were a greater number of concurrent protocol segments than retrospective protocol segments. Concurrent data provided more insights into decision-making steps. However, retrospective data provided more statements about final choice.

As Pressley and Afflerbach (1995) noted, "Spoken language is the data used in protocol analysis, and the richness and variability of language are the greatest assets and liabilities of the verbal reporting methodology" (p. 2). With this in mind, along with the primary objective of our study, we utilized both concurrent and retrospective reporting in the present study, as described in the Method section.

Cognitive Lab Study Goals

The primary purpose of conducting cognitive labs in the present study was to verify and validate potential career-readiness features that are necessary to solve math and ELA test items. There were three target goals for the cognitive lab study: (a) to determine the thought processes that students went through when solving each item; (b) to determine aspects/features of the items that might pose difficulties or facilitate students' comprehension of the items; and (c) to verify and validate potential career-readiness features that are part of students' problem-solving process. As such, the present study made use of both concurrent and retrospective reports in order to capture students' problem-solving steps as well as capture any decision making not immediately apparent in concurrent reporting. Since the primary purpose of the study was to verify features, a semistructured interview protocol was employed, which is described further in the Method section.

Feature Analysis: Theoretical Background

As summarized by Baker et al. (2015), one of the earliest references to the idea of feature analysis can be attributed to Gordon (1970) in the *Report of the Commission on Tests: II.* In this brief report, Gordon mentioned qualitative analysis of assessments to emphasize "description and prescription," that is, the qualitative *description* of cognitive functions leading to the *prescription* of the learning experiences required to more adequately ensure academic success. Gordon suggested that existing instruments can be examined with a view towards categorization and interpretation to determine whether data can be reported in qualitative ways, in addition to traditional quantitative ways. For instance, Gordon mentioned that response patterns can be reported differently for information recall or vocabulary. He also referred to features such as problem solving, expression, and information management, among many others.

The main rating framework underlying the feature analysis work is derived from Baker and O'Neil's (2002) approach to designing problem solving assessments, Jonassen's (2000) typology of problems, and CRESST's problem-solving ontology. Baker and O'Neil's approach first characterizes three types of problem-solving tasks: (a) a task in which an appropriate solution is known in advance, (b) a task in which there is no known solution to the problem, and (c) a task that requires the application of a given tool set to a broad range of topics. These are all features of problem solving that can be rated as part of a particular item and that require a specific associated cognitive demand or process on the part of the student.

Identifying the problem is often one of the most difficult aspects of problem solving (see Baker & O'Neil, 2002). The ambiguity of problem identification may be dependent on the prior knowledge of the learner and the purpose of the assessment. An assessment developer can adjust the difficulty of a task or item by stating the problem explicitly or obscuring it in context, such as within a narrative. The difficulty of an item or task can also be adjusted by either providing extraneous information or developing a task with missing information that needs to

be constructed. These types of adjustments might increase difficulty and also require an associated cognitive demand or cognitive process that might be more complex.

Similar to Baker and O'Neil (2002), Jonassen (2000) articulated different problem types with varying attributes that follow a continuum from well-structured to ill-structured tasks. Jonassen's problem typology assumes that there are similarities in the cognitive processes required to solve each problem, that the problem types are not mutually exclusive, and that each problem category varies with respect to abstractness and complexity. Baker and O'Neil's (2002) approach to developing problem-solving assessments is part of an overarching model developed by CRESST to determine the functional validity of assessments (Cai, Baker, Choi, & Buschang, 2014). Unlike previous models, this approach provides criterion-referenced evidence in support of assessment validity claims by integrating feature rating and step-by-step analysis with modern statistical techniques.

CRESST has previously used qualitative content analysis techniques to assess language demands in standardized assessments for math, ELA, and tests of English language proficiency (Wolf et al., 2008) as well as cognitive, grammatical, textual, and visual features on standardized assessments for students with disabilities (Abedi et al., 2011, 2012). A quantitative technique similar to feature analysis was previously introduced in Roberts, Chung, and Parks (2016) and used to categorize attributes of metadata created when children interacted with educational online games and media. This approach was previously used in work with PBS (Chung & Parks, 2015; Chung & Redman, 2015a, 2015b). This model ensures validity by going beyond simple task descriptions, and by yielding an explanation for possible areas of growth, identifying task elements that are suitable for instruction, and lastly, providing a method for comparability and prediction.

Feature Analysis Goals

Feature analysis is defined as the qualitative rating of tasks against a set of attributes, in this case, career- and college-readiness features, followed by a subsequent quantitative analysis to determine how these attributes determine task performance. This report focuses on a preliminary feature analysis aimed toward developing and refining the career-readiness feature set. Later reports delineate feature ratings of a larger set of Smarter Balanced math and ELA items, including quantitative analysis components. As aforementioned, the feature analysis process includes feature rating, step-by-step analysis, cognitive labs, and quantitative analysis (Baker et al., 2015). Feature analysis, in the context of career readiness, aims to address such questions as (a) What particular attributes/features does each item contain? (b) Which features appear more frequently across items? (c) Are there differences in feature representation across domain areas (i.e., math vs. ELA)? and (d) What particular attributes or features of items and tasks explain variations in item characteristics such as increased or reduced difficulty across items, tests, and tasks and why? This report provides preliminary answers to the first three

questions listed above through the use of feature rating and cognitive labs. Subsequent reports address these questions in further detail.

Feature Creation

The goal of the feature rating or scoring process is to determine what features and attributes are present or absent in a particular item and what steps need to be completed to solve an item or task correctly, and to perform descriptive analyses across features and items. This task is performed by content-area experts. The features are refined to target a particular content area prior to feature rating by the experts.

The feature creation process requires several steps. The initial step involves selecting and reviewing key resource materials. To create the current set of career skills and features, researchers reviewed the Bureau of Labor Statistics career data, the O*NET online databases, and previous CRESST ontologies and feature and rating schemes. These resource materials were studied to determine an initial set of career skills. These skills were refined by selecting those that were categorized as most important based on O*NET importance ratings. This process was repeated both across and within two exemplar careers that were chosen to guide the overall project. These two careers are emergency medical technician (EMT), representing the healthcare industry, and web developer, representing the technology industry. These careers were chosen as exemplars to represent two different industries, and because both careers, at the time of the study, were reported by the Bureau of Labor Statistics as having "Bright Outlook" and "rapid growth." In addition, these two careers did not require a four-year degree but did require some postsecondary training.

The feature set was further refined by utilizing Smarter Balanced items and blueprints as selection criteria (i.e., features that were not likely to be found within the items were taken out). Previous CRESST feature analysis results also informed the current feature set. Specifically, features that were found to contribute variance in previous CRESST studies were included as part of the current set. Finally, the feature set was refined by incorporating the expertise of select subject-matter experts (SMEs) in college and career readiness, business, pre-hospital care, and web development. These SMEs filled out a survey where they answered targeted questions about the career feature set. (See Appendix A for the survey.)

The SMEs were first asked to indicate the extent to which each career skill was not applicable, contributes to, or was essential to effectively and successfully complete daily job-related tasks. The SMEs then indicated which 30 skills were most important from those rated as essential. The SMEs were then asked to rank these 30 skills in order of importance and create an operational example of the 15 highest ranked skills. After this feature selection process, the SMEs were asked to review and verify the final set that would be utilized for feature rating. The goal was for the features to be action-oriented with adequate granularity to allow for implementation with low inference across domains and task types.

Target Career-Readiness Features

Table 1 includes the initial set of 36 career-readiness features created and reviewed by SMEs following the process delineated above. They are grouped broadly into three categories: Skills, Abilities, and Work Activities/Context for ease of presentation. “Skills” generally refer to developed capacities that facilitate learning or the more rapid acquisition of knowledge. “Abilities” generally refer to enduring attributes of the individual that influence performance. “Work Activities” generally refer to the types of job behaviors occurring on multiple jobs. “Work Context” refers to the physical and social factors that influence the nature of the work.

Table 1

Target Career-Readiness Features by Category

Feature	Description
Features related to skills	
Active learning	Understanding the implications of new information for both current and future problem solving and decision making.
Active listening	Giving full attention to what other people are saying and taking time to understand the points being made.
Complex problem solving	Identifying complex problems and reviewing related information to develop and evaluate options and implement solutions.
Critical thinking	Using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions, or approaches to problems.
Judgment and decision making	Considering the relative costs and benefits of potential actions to choose the most appropriate one.
Mathematics	Using mathematics to solve problems.
Monitoring	Monitoring/assessing performance of yourself, other individuals, or organizations to make improvements or take corrective action.
Reading comprehension	Understanding written sentences and paragraphs in work-related documents.
Features related to abilities	
Deductive reasoning	The ability to apply general rules to specific problems to produce answers that make sense.
Flexibility of closure	The ability to identify or detect a known pattern, figure, object, word, or sound that is hidden in other distracting material.
Fluency of ideas	The ability to come up with a number of ideas about a topic.
Inductive reasoning	The ability to combine pieces of information to form general rules or conclusions.
Information ordering	The ability to arrange things or actions in a certain order or pattern according to a specific rule or set of rules.
Mathematical reasoning	The ability to choose the right mathematical methods or formulas to solve a problem.
Memorization	The ability to remember information such as words, numbers, pictures, and procedures.
Number facility	The ability to add, subtract, multiply, or divide quickly and correctly.
Oral comprehension	The ability to listen to and understand information and ideas presented through spoken words and sentences.

Feature	Description
Problem sensitivity	The ability to tell when something is wrong or is likely wrong. It does not involve solving the problem, only recognizing there is a problem.
Selective attention	The ability to concentrate on a task over a period of time without being distracted.
Time sharing	The ability to shift back and forth between two or more activities or sources of information.
Written comprehension	The ability to read and understand information and ideas presented in writing.
Written expression	The ability to communicate information and ideas in writing so others will understand.
Visualization	The ability to imagine how something will look after it is moved around or when its parts are moved or rearranged.
Features related to work activities/context	
Analyzing data and information	Identifying the underlying principles, reasons, or facts of information by breaking down information or data into separate parts.
Documenting/recording information	Entering, transcribing, recording, storing, or maintaining information in written or electronic/magnetic form.
Estimating the quantifiable characteristics of products, events, or information	Estimating sizes, distances, and quantities; or determining time, costs, resources, or materials needed to perform a work activity.
Getting information	Observing, receiving, and otherwise obtaining information from all relevant sources.
Identifying objects, actions, and events	Identifying information by categorizing, estimating, recognizing differences or similarities, and detecting changes in circumstances or events.
Importance of being exact or accurate	Being very exact or highly accurate is important to performing this job.
Interacting with computers	Using computers and computer systems, including hardware and software, to program, write software, set up functions, enter data, or process information.
Judging the qualities of things, services, or people	Assessing the value, importance, or quality of things or people.
Making decisions and solving problems	Analyzing information and evaluating results to choose the best solution and solve problems.
Organizing, planning, and prioritizing work	Developing specific goals and plans to prioritize, organize, and accomplish your work.

Feature	Description
Processing information	Compiling, coding, categorizing, calculating, tabulating, auditing, or verifying information or data.
Thinking creatively	Developing, designing, or creating new applications, ideas, relationships, systems, or products, including artistic contributions.
Updating and using relevant knowledge	Keeping up-to-date technically and applying new knowledge.

Method

Participants

A total of 17 high school students (ranging in age from 16 to 18) participated in the study. Participants were recruited from three schools in Southern California with diverse school profiles. School A is a small charter school in a large city, serving Grades 6 through 12. School B is a large public high school in a small city, serving Grades 9 through 12. School C is a small magnet high school in a large city, serving Grades 9 through 12. Five participants were from School A, nine participants were from School B, and three were from School C.

Background information, based on student surveys, were available from 16 participants: nine students were in Grade 11 (56.3%) and seven students were in Grade 12 (43.7%). About 75% of the participants identified as male. Nine students identified as Hispanic (56.3%), three as White/Caucasian (31.3%), one as Asian/Pacific Islander (6.3%), and one as Black/African American (6.3%). Two students identified multiethnic backgrounds (12.6%). Zero students identified as American Indian or Alaska Native. Ten students (62.5%) identified a language other than English spoken at home at least half the time. Most participants reported coming from households which had at least one working laptop or computer (87.5%). See Appendix B for a breakdown of background variables.

Instruments

Background survey. An eight-item background survey, designed to be administered on paper following the cognitive lab, queried students' opinions on the difficulty of the test items and their perceived effort in completing the items. The survey also asked general demographic information (age, grade level, gender, frequency of languages other than English spoken at home, number of computers or laptops at home, and race/ethnicity). See Appendix C for the survey.

Item selection. Four math and two ELA items were selected from Smarter Balanced practice tests, which were available online at the time of the study. Items were selected to represent a variety of content specifications including varied claims, targets, standards, and response/item types with consideration for the limited administration time.

Table 2 shows the items and associated claims, targets, and depth of knowledge (DOK) levels. According to Smarter Balanced, claims are summary statements about the knowledge and skills students are expected to demonstrate on the assessment related to a particular aspect of the standards. A target is a statement that bridges the content standards and the assessment evidence that supports the claim. DOK is a measure of complexity considering the student's cognitive process in response to an item. There are four levels on the DOK scale, with 4 being the highest level. Math claims include Concepts and Procedures (Claim 1), Problem Solving (Claim 2), Communicating Reasoning (Claim 3), and Modeling and Data Analysis (Claim 4). ELA claims include Reading (Claim 1), Writing (Claim 2), Speaking and Listening (Claim 3), and Research/Inquiry (Claim 4).

Table 2

Claims, Targets, and Depth of Knowledge, by Item

Subject	Item	Claim	Target	Depth of knowledge	Evidence statement
Math	13	3	E	3	N/A
	16	1	J	1	N/A
	23	4	E	3	N/A
	29	2	D	2	N/A
ELA	28	4	3	2	The student will use reasoning, evaluation, and evidence to assess the credibility of multiple sources in order to select relevant information to support research.
	30	4	4	2	The student will cite evidence to support arguments or conjectures.

Note. For more information on claims, targets, and depth of knowledge, please go to www.smarterbalanced.org

An additional ELA task (listening comprehension) with three items and an additional math item were selected as bonus questions, given available time during administration. However, since few students reached these items they are not reported here.

Cognitive lab protocol. A protocol was developed to take notes during the cognitive lab (see Appendix D). The protocol included an image of each test item along with the additional bonus questions. On the form, each test item was followed by a form with questions, as well as space for the researcher to take open-ended notes as well as mark responses to probing questions with checkboxes, including "Did you understand what the question was asking you to do?" and "Was this problem easy, medium, or hard?" Researchers used the blank space to note participant utterances as well as participant actions. Appendix E provides the answer keys and rubrics for the items.

Feature Rating Procedures

Utilizing a preliminary set of career-readiness features, as described earlier, feature rating was performed on the six Smarter Balanced practice items selected for the cognitive lab. Three CRESST researchers trained together by coding one item and discussing disagreements to achieve consensus. Items were coded as 0 or 1 for each feature, with 0 meaning "No, the feature is not necessary to solve the problem correctly," and 1 meaning "Yes, the feature is necessary to solve the problem correctly." After the initial training, researchers then coded two items individually, meeting to discuss disagreements and achieve consensus. A final list of features for each item was produced, which is presented in the Results section. Interrater agreement (percent agreement) on the features for the latter five items was, on average, 88.8%.

Cognitive Lab Procedures

Recruitment. After approval from school districts and school site personnel, participants were recruited through flyers sent to school site coordinators. Within schools, participants were recruited from general education courses with no specific exclusion criteria. Students interested in participating were asked to bring home and return a signed parent permission form if they were under 18. The initial target recruitment was approximately three to five students per school site. However, since more than five students volunteered from School B, efforts were made to include all nine students who volunteered.

Researcher training. Five CRESST researchers were trained on administering the cognitive lab protocol. Two full-length trials were conducted internally to refine the cognitive lab protocol and demonstrate the interview process. Trainees took notes on the demonstration which were then compared for consistency. Timing was also refined and technology was also tested at this time.

Administration. Five CRESST researchers conducted the cognitive labs at the school sites. Each cognitive lab interview was conducted one-on-one and was designed to be conducted in less than one hour, either during the school day or after school. Each of the CRESST researchers conducted between one and five interviews. Researchers introduced themselves, explained the nature of the study, and gave participants the opportunity to complete the consent or assent process. All sessions were audio recorded with no participant objections. The purpose of the audio recording was to serve as a supplement to handwritten notes.

Additional information about *how* the student would participate in the cognitive lab was read at the beginning of the study. The script used at the beginning of the cognitive lab is on the first page of the protocol (see Appendix D). A brief demonstration of how to perform a think aloud was given using a simple arithmetic problem prior to administration. The demonstration item was conducted using paper and pencil.

Participants completed the cognitive lab items on laptops online via the Smarter Balanced website. The selected test items were designed to be administered in chronological order (i.e., Math Items 13, 16, 23, and 29, then ELA Items 28 and 30) but since not all test items from the practice test were used, researchers needed to click ahead to the selected items. Additionally, all participants were provided with pencils and scratch paper, which were collected at the end of administration. Participants were permitted to skip a problem if they did not want to or know how to answer it. In case of technological issues, paper copies of the test items were also available. Three of the participants completed a portion of the cognitive lab on paper after some issues with technology.

While completing test items, participants were instructed to think aloud, but were not interrupted if they chose to read or think silently in order for the researcher to remain neutral and maintain as much of a natural test setting as possible. Participants were gently prompted if it appeared they were not making progress or seemed stuck on a problem. After completing a problem, participants were prompted with "Tell me in words what you did to solve the problem." These retrospective reports were given immediately after completing a problem, which is consistent with Ericsson and Simon's (1993) recommendation to maintain close correspondence between concurrent and retrospective reports to reduce interference from long-term memory.

Following completion of the cognitive lab test items, the audio recorder was turned off and participants completed the background survey on paper. Participants were given a gift card in the amount of $25 as a thank you for their participation.

Analyses. Both quantitative and qualitative analyses were conducted on the collected data. Student performance on test items was scored according to both Smarter Balanced scoring guides as well as researcher-developed scoring (for more nuanced scoring for items with partial scoring of items). Checkbox items from the cognitive lab protocol were coded and entered. Background survey data were also entered. Descriptive statistics were computed for all of the above.

Using the audio recordings, participants' verbal utterances were electronically transcribed and integrated with handwritten notes from the cognitive lab protocol, so that each protocol report contained both participant utterances as well as actions. Using the features set list (see Table 1 earlier), reports from each participant were systematically examined for evidence of each of the features in the list (or lack thereof) by item. The feature list by item is presented in the Results section below.

Cognitive Lab and Feature Results

Participants completed four math and two ELA items, for a total of six test items. Three participants completed only one ELA item because of time constraints. Additionally, students were permitted to skip a problem if they did not want to or know how to answer, thus, some

data are incomplete. Individual cognitive lab sessions ranged in length from approximately 17 to 53 minutes with considerable variation (M = 38.6; SD = 9.5).

As part of the cognitive labs, students answered some questions regarding their effort, perceptions of item difficulty, and the degree to which they understood the items. Of the 16 participant responses, six reported that the cognitive lab problems were easier than other tests they had taken that year (37.5%), nine reported problems were about the same (56.3%), and one participant reported the test was harder (6.3%). Two participants reported putting in a little effort on the cognitive lab items (12.5%), seven participants reported putting in some effort (43.8%), and seven reported putting in a lot of effort (43.8%).

Key Features

Table 3 is a visual summary of the representation of key features, as rated by CRESST researchers, across the Smarter Balanced practice test items selected for the cognitive lab.

Table 3

Career Readiness Key Features by Test Item

Feature	Math item				ELA item		Total
	13	16	23	29	28	30	
Features related to skills							
Critical thinking					X		1
Mathematics	X	X	X	X			4
Monitoring						X	1
Reading comprehension			X	X	X	X	4
Features related to abilities							
Deductive reasoning			X	X	X	X	4
Flexibility of closure				X			1
Information ordering	X						1
Mathematical reasoning	X	X	X	X			4
Number facility	X	X	X	X			4
Problem sensitivity			X				1
Visualization				X			1
Written comprehension	X	X	X	X	X	X	6
Features related to work context							
Analyzing data or information	X	X	X	X	X	X	6
Documenting/recording information	X		X				2
Estimating the quantifiable characteristics of products, events, or information				X			
Getting information					X	X	2
Identifying objects, actions, and events			X	X	X	X	4
Importance of being exact or accurate	X						1
Interacting with computers		X					1
Judging the qualities of things, services, or people					X		1
Making decisions and solving problems	X	X		X	X	X	5
Processing information	X	X	X	X	X	X	6
Total number of features	10	8	11	13	10	9	62

As depicted in the table, feature representation ranged from eight to 13 features across items. Items 13, 23, 29, and 28 had the most feature representation. There were three main features that were represented across all of the selected items: analyzing data or information, processing information, and written comprehension.

The following features were least represented across the items: importance of being exact or accurate; judging the qualities of things, services, or people; interacting with computers; problem sensitivity; information ordering; visualization; flexibility of closure; critical thinking; and monitoring. The next section summarizes the findings from the cognitive lab.

Table 4 represents student responses across items. The "correct" column represents when a student completed all parts of an item correctly. The notation of correct and incorrect responses are based on Smarter Balanced item scoring criteria, with the exception of Math Item 13. Math Item 13 was worth two points but combined as one for ease of presentation. Researchers created a "partially correct" category to reflect students who may have solved one or multiple aspects of an item correctly. The attempted category represents how many students attempted to solve the item. If the attempted category does not reflect the complete sample of students (*N* = 17), such as with Math Item 14, it indicates that some students chose to skip the item. Overall, more students scored incorrect than correct across items. Appendix F contains a figure illustrating correct and incorrect items by participant.

Table 4

Correct and Incorrect Answer Attempts by Item, Frequency, and Percentage (N = 17)

Subject	Item	Attempted (%)	Correct (%)	Incorrect (%)	Partially correct (%)
Math	13[a]	17 /17 (100.0)	10/17 (58.8)	7/17 (41.2)	7/7 (100.0)
Math	16	14/17 (82.4)	7/14 (50.0)	7/14 (50.0)	N/A
Math	23	17 /17 (100.0)	5/17 (29.4)	12/17 (70.6)	N/A
Math	29	15 (88.2)	2/15 (13.3)	13/15 (86.7)	13/13 (100.0)
ELA	28	17/17 (100.0/)	5/17 (29.4)	12/17 (70.6)	11/12 (91.7)
ELA	30	14[b] (100.0)	2/14 (14.3)	12/14 (85.7)	12/12 (100.0)

Note. There were a total of 17 total participants, but not all participants answered all questions.

[a]For ease of presentation, Correct for Item 13 meant the student solved both parts of the problem correctly, and Incorrect meant at least one part was incorrect.

[b]Three students did not get to Item 30 because of time restrictions, not because they declined to attempt.

After completing each item, participants were asked if they understood what the problem was asking them to do. Participants responded with yes, no, or some combination. Additionally, participants were asked to rate the difficulty of individual problems as easy, medium, or hard. Table 5 includes information for students' self-reported understanding of each problem, as well as perceived difficulty rated from easy (coded as 1) to hard (coded as 3). In the paragraphs below, results related to correctness, student understanding and difficulty, and features are summarized and discussed by content area. For more detailed results by item, including the features rated by researchers and observed as part of students' item completion processes during the cognitive lab interviews, please see Appendix G.

Table 5

Student-Reported Understanding and Difficulty by Item, Frequency, and Percentage

Subject	Item	Understood question (%)	Perceived difficulty *M*
Math	13	93.3	1.2
Math	16	78.5	2.2
Math	23	88.2	1.8
Math	29	100.0	2.3
ELA	28	94.1	1.2
ELA	30	100.0	1.1

Note. n = 13 to 17 due to missing responses.

Math Item Results

For the four math items, 78.5% to 100% of the students indicated that they understood the items, and the students reported the math items as having a difficulty level ranging, on average, from 1.2 to 2.3 on a scale from 1 (easy) to 3 (hard). Feature representation across the math items specifically ranged from eight to 13. The most frequently occurring features across the math items were mathematics, mathematical reasoning, number facility, written comprehension, analyzing data or information, and processing information. The cognitive lab suggested that these features were salient to students' mathematical item completion processes. To successfully complete most of the math items, students needed to identify specific mathematical facts and/or principles (i.e., analyzing data and information), determine what mathematical formula or method to implement to solve the problem (i.e., mathematical reasoning), compute correctly (i.e., number facility), verify that the data or information provided was correct (i.e., processing information), and understand the information, instructions, and directions provided within the items (i.e., written comprehension). Two key features that were unique across two respective math items included problem sensitivity and interacting with computers. In one item, students needed to recognize that there was an issue and that key information was not explicitly stated in the item (i.e., problem sensitivity). In another item, students were required to use the computer to select a response, which was deemed as a unique interaction with the test item (i.e., interacting with computers). For further detail on each of the four math items, see Appendix G.

ELA Item Results

For the two ELA items, 94.1% and 100% of students reported that they understood the items, and the students reported that the ELA items had a difficulty level of 1.2 and 1.1, respectively, on a scale from 1 (easy) to 3 (hard). Generally, students perceived the ELA items to

be easier than the math items. Feature representation across the ELA items ranged from nine to 10, indicating that the ELA items had less feature representation than the math items overall. The most prevalent features across the ELA items were reading comprehension; deductive reasoning; getting information; identifying objects, actions, and events; making decisions and solving problems; and similar to math, written comprehension, analyzing data or information, and processing information. The three latter features were also important across the ELA items for similar reasons and processes as indicated for the math items. For the ELA items, the cognitive lab showed that the ability to apply specific rules to generate answers (i.e., deductive reasoning) was salient, as was obtaining information from all relevant sources (i.e., getting information) and analyzing the information to choose the best solution and solve the item (i.e., making decisions and solving problems). Moreover, judging the qualities of things, services, or people, and critical thinking were needed to solve one of the ELA items. For further detail on the two ELA items, see Appendix G.

Results Summary

Across both math and ELA items, 78.5% to 100% of the students indicated that they understood the items overall. Students also reported the items as having a difficulty level ranging, on average, of 1.1 to 2.3 on a scale from 1 (easy) to 3 (hard). Students reported math items as being more difficult than ELA items. Overall, the cognitive lab demonstrated that certain attributes related to career readiness were prevalent in students' item completion processes.

Feature representation ranged from eight to 13 features across the six Smarter Balanced practice items based on CRESST expert ratings. Three main features were represented across the selected items: analyzing data or information, processing information, and written comprehension. The cognitive lab revealed that these features were important to students' problem solving. Three of the math items (Items 13, 23, and 29) and one of the ELA items (Item 28) had the most feature representation (10 or more each). Written comprehension was essential for students to understand the item prompts, directions, instructions, and questions. Analyzing data and information was essential for students to understand key facts and data provided in the items. Processing information was important for students to ensure that the information and data that they were utilizing to solve the item was appropriate. Across both math and ELA, making decisions and solving problems; identifying objects, actions, and events; and deductive reasoning also surfaced as important features.

While many of the students in this study arrived at incorrect answers, their attempts to complete the items during the cognitive labs provided insight into the features, regardless of whether their answers were ultimately correct. For instance, students who solved math items incorrectly may have been limited in knowledge or ability to analyze data or information or have been lacking in number facility or mathematical reasoning. Students who solved the ELA items incorrectly showed issues with reading comprehension; written comprehension; or

identifying objects, actions, and events; among other features. Students' limitations in adequately engaging in some of these features led, in part, to their arriving at incorrect answers. Thus, students who solved problems incorrectly provided as much insight into features as those who solved problems correctly.

Conclusion

In today's global, knowledge-based economy, students need to be prepared to compete in a world that demands more than just basic skills (U.S. Department of Education, n.d.). In addition to increasing standards and curriculum reform, rigorous, high-quality assessments may reveal strengths and weaknesses that support student success in college and career. Identifying career-readiness features in existing Smarter Balanced assessments is a move in that direction.

The cognitive lab study provides preliminary evidence that career-readiness features can indeed be found in a small and targeted set of released Smarter Balanced math and ELA practice items. These findings are important as they provide insight into which career-readiness skills may be needed for students to solve high school math and ELA assessments. From a methodological standpoint, these findings support the approach used to draw additional inferences from content-area assessments. Career-readiness inferences drawn from the assessments support student success in college and career, and can suggest improvements in existing high school instruction and curriculum, which can potentially impact the workforce as a whole.

A limitation to this study is that the items selected for the cognitive lab were from practice test items, which are available publicly from the Smarter Balanced website. Active test items may have different attributes and offer fewer or greater numbers of features. Additionally, the small sample size, while demographically diverse, is limited, and suggests that findings may not be generalizable to other populations. However, the results validate our feature-rating process, with our rated features confirmed through the cognitive lab across all six items.

The next report in the series of reports on career readiness describes the rating of a larger number of active Smarter Balanced items utilizing a similar but refined process, with additional analyses devoted to psychometric modeling using these ratings. The third report in this series relates to conducting feature ratings of Korean assessment items for international benchmarking. The fourth report focuses on conducting a field study to validate assessment items across two different exemplar careers. Finally, the fifth report discusses the development of prototype innovative assessment items designed to directly measure specific career-readiness skills.

References

Abedi, J., Bayley, R., Ewers, N., Mundhenk, K., Leon, S., Kao, J., & Herman, J. (2012). Accessible reading assessments for students with disabilities. *International Journal of Disability, Development and Education*, *59*(1), 81-95.

Abedi, J., Leon, S., Kao, J., Bayley, R., Ewers, N., Herman, J., & Mundhenk, K. (2011). *Accessible reading assessments for students with disabilities: The role of cognitive, grammatical, lexical, and textual/visual features* (CRESST Report 785). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing.

Baker, E. L., Madni, A., Michiuye, J. K., Choi, K., & Cai, L. (2015). *Mathematical Reasoning Project quantitative analyses results: Grades 4, 8, and 11.* Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing.

Baker, E. L., & O'Neil, H. F., Jr. (2002). Measuring problem solving in computer environments: Current and future states. *Computers in Human Behavior, 18,* 609-622.

Branch, J. L. (2000). Investigating the information-seeking processes of adolescents: The value of using think alouds and think afters. *Library & Information Science Research*, *22*(4), 371-392.

Cai, L., Baker, E., Choi, K., & Buschang, R. (2014, April). *CRESST functional validity model: Deriving formative and summative information from Common Core assessments*. Presentation at the annual meeting of the American Educational Research Association, Philadelphia, PA.

The California Department of Education. (2016). *Dataquest*. Retrieved from: http://dq.cde.ca.gov/dataquest/

Carnevale, A. P., Smith, N., & Strohl, J. (2013). *Recovery: Job growth and education requirements through 2020.* Washington, DC: Georgetown University, Georgetown Public Policy Institute, Center on Education and the Workforce. Retrieved from https://repository.library.georgetown.edu/bitstream/handle/10822/559311/Recovery2020.FR.Web.pdf

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, *6*(3), 271-315.

Chung, G. K. W. K., & Parks, C. (2015). *Feature analysis validity report* (Deliverable to PBS KIDS). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing.

Chung, G. K. W. K., & Redman, E. H. (2015a). *Feature analysis framework – final* (Deliverable to PBS KIDS). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing.

Chung, G. K. W. K., & Redman, E. H. (2015b). *Feature analysis ratings best practices* (Deliverable to PBS KIDS). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing.

Council of Chief State School Officers. (2013). *Knowledge, skills, and dispositions: The Innovation Lab Network state framework for college, career, and citizenship readiness, and implications for state policy.* Washington, DC: Author.

Council of Chief State School Officers. (2014). *Opportunities and options: Making career preparation work for students*. Washington, DC: Author.

Durning, S. J., Artino, A. R., Jr., Beckman, T. J., Graner, J., van der Vleuten, C., Holmboe, E., & Schuwirth, L. (2013). Does the think-aloud protocol reflect thinking? Exploring functional neuroimaging differences with thinking (answering multiple choice questions) versus thinking aloud. *Medical Teacher, 35*(9), 720-726.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215-251.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. London, England: MIT Press.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). London, England: MIT Press.

Executive Office of the President of the United States of America. (2014). *Ready to work: Job-driven training and American opportunity.* Washington, DC: Author. Retrieved from https://obamawhitehouse.archives.gov/sites/default/files/docs/skills_report.pdf

Gordon, E. W. (1970). Toward a qualitative approach to assessment. *Report of the Commission on Tests, II. Briefs* (pp. 42-46). New York, NY: College Entrance Examination Board.

Johnstone, C. J., Bottsford-Miller, N. A., & Thompson, S. J. (2006). *Using the think aloud method (cognitive labs) to evaluate test design for students with disabilities and English language learners* (Technical Report 44). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology: Research & Development, 48*(4), 63-85.

Kuusela, H., & Paul, P. (2000). A comparison of concurrent and retrospective verbal protocol analysis. *The American Journal of Psychology*, *113*(3), 387-404.

Leighton, J. P. (2017). *Using think aloud interviews and cognitive labs in educational research.* New York, NY: Oxford University Press.

Mejia, M. C., Rodriguez, O., & Johnson, H. (2016). *Preparing students for success in California's community colleges.* Retrieved from http://www.ppic.org/content/pubs/report/R_1116MMR.pdf

O*NET Online. (n.d.). https://www.onetonline.org/

Padilla, J.-L., & Leighton, J. P. (2017). Cognitive interviewing and think aloud methods. In B. D. Zumbo & A. M. Hubley (Eds.), *Understanding and investigating response processes in validation research* (Social Indicators Research Series 69; pp. 211-228). Cham, Switzerland: Springer International Publishing.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Hillsdale, NJ: Erlbaum.

Roberts, J. D., Chung, G. K. W. K., & Parks, C. (2016). Supporting children's progress through the PBS KIDS learning analytics platform. *Journal of Children and Media, 10*(2). Retrieved from http://dx.doi.org/10.1080/17482798.2016.1140489

Smarter Balanced Assessment Consortium. (2015, January). *Close to 200 colleges and universities to use Smarter Balanced scores as part of placement* [press release]. Retrieved from http://www.smarterbalanced.org/close-to-200-colleges-and-universities-to-use-smarter-balanced-scores-as-part-of-placement/

U. S. Department of Education. (n. d.). *College- and career-ready standards.* Washington, DC: Author. Retrieved from https://www.ed.gov/k-12reforms/standards

van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical approach to modeling cognitive processes.* London, England: Academic Press.

Wolf, M. K., Herman, J. L., Kim, J., Abedi, J., Leon, S., Griffin, N., ... Shin, H. W. (2008). *Providing validity evidence to improve the assessment of English language learners* (CRESST Report 738). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing.

Wolf, M. K., Kim, J., Kao, J. C., & Rivera, N. M. (2009). *Examining the effectiveness and validity of glossary and read-aloud accommodations for English language learners in a math assessment* (CRESST Report 766). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing.

Zucker, S., Sassman, C., & Case, B. J. (2004). *Cognitive labs* (Pearson Technical Report). San Antonio, TX: Pearson Inc. Retrieved from http://images.pearsonassessments.com/images/tmrs/tmrs_rg/CognitiveLabs.pdf?WT.mc_id=TMRS_Cognitive_Labs

Appendix A:
Subject Matter Expert Survey

Indicate Career

Please indicate below what field you work in.

What is your career area?

What is your job title?

Introduction to Survey

Next is a list of skills. Please indicate on a scale from 1 to 3, whether the skill is "not applicable," "contributes to," or "essential" for career readiness. You can move back and forth between the skills/pages by utilizing the arrow buttons.

Skills

Please indicate on a scale from 1 to 3, whether the skill is "not applicable," "contributes to," or "essential" for career readiness.

Using electronic mail.

Writing letters and memos.

Being responsible for outcomes and results of other workers.

Decision making (making decisions that affect other people, the financial resources, and/or the image and reputation of the organization).

Importance of being exact or accurate.

Time pressure (the job requires the worker to meet strict deadlines).

Analyzing data and information (identifying the underlying principles, reasons, or facts of information by breaking down information or data into separate parts).

Making decisions and solving problems (analyzing information and evaluating results to choose the best solution and solve problems).

Organizing, planning, and prioritizing work (developing specific goals and plans to prioritize, organize, and accomplish your work).

Processing information (compiling, coding, categorizing, calculating, tabulating, auditing, or verifying information or data).

Updating and using relevant knowledge (keeping up-to-date technically and applying new knowledge to your job).

Evaluating information to determine compliance with standards (using relevant information and individual judgment to determine whether events or processes comply with laws, regulations, or standards).

Scheduling work and activities (scheduling events, programs, and activities, as well as the work of others).

Judging the qualities of things, services, or people (assessing the value, importance, or quality of things or people).

Thinking creatively (developing, designing, or creating new applications, ideas, relationships, systems, or products, including artistic contributions).

Getting information (observing, receiving, and otherwise obtaining information from all relevant sources).

Monitor processes, materials, or surroundings (monitoring and reviewing information from materials, events, or the environment, to detect or assess problems).

Estimating the quantifiable characteristics of products, events, or information (estimating sizes, distances, and quantities; or determining time, costs, resources, or materials needed to perform a work activity).

Identifying objects, actions, and events (identifying information by categorizing, estimating, recognizing differences or similarities, and detecting changes in circumstances or events).

Interpreting the meaning of information for others (translating or explaining what information means and how it can be used).

Documenting/recording information (entering, transcribing, recording, storing, or maintaining information in written or electronic/magnetic form).

Interacting with computers (using computers and computer systems, including hardware and software, to program, write software, set up functions, enter data, or process information).

Customer and personal service (knowledge of principles and processes for providing customer and personal services. This includes customer needs assessment, meeting quality standards for services, and evaluation of customer satisfaction).

Problem sensitivity (the ability to tell when something is wrong or is likely wrong. It does not involve solving the problem, only recognizing there is a problem).

Information ordering (the ability to arrange things or actions in a certain order or pattern according to a specific rule or set of rules).

Oral comprehension (the ability to listen to and understand information and ideas presented through spoken words and sentences).

Written comprehension (the ability to read and understand information and ideas presented in writing).

Written expression (the ability to communicate information and ideas in writing so others will understand).

Deductive reasoning (the ability to apply general rules to specific problems to produce answers that make sense).

Mathematical reasoning (the ability to choose the right mathematical methods or formulas to solve a problem).

Inductive reasoning (the ability to combine pieces of information to form general rules or conclusions).

Selective attention (the ability to concentrate on a task over a period of time without being distracted).

Fluency of ideas (the ability to come up with a number of ideas about a topic).

Memorization (the ability to remember information such as words, numbers, pictures, and procedures).

Category flexibility (the ability to generate or use different sets of rules for combining or grouping things in different ways).

Originality (the ability to come up with unusual or clever ideas about a given topic or situation, or to develop creative ways to solve a problem.

Time sharing (the ability to shift back and forth between two or more activities or sources of information).

Visualization (the ability to imagine how something will look after it is moved around or when its parts are moved or rearranged).

Flexibility of closure (the ability to identify or detect a known pattern, figure, object, word, or sound that is hidden in other distracting material).

Perceptual speed (the ability to quickly and accurately compare similarities and differences among sets of letters, numbers, objects, pictures, or patterns.

Speed of closure (the ability to quickly make sense of, combine, and organize information into meaningful patterns).

Number facility (the ability to add, subtract, multiply, or divide quickly and correctly).

Speech recognition (the ability to identify and understand the speech of another person).

Active learning (understanding the implications of new information for both current and future problem-solving and decision-making.

Active listening (giving full attention to what other people are saying and taking time to understand the points being made).

Critical thinking (using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions or approaches to problems).

Mathematics (using mathematics to solve problems).

Monitoring (monitoring/assessing performance of yourself, other individuals, or organizations to make improvements or take corrective action).

Reading comprehension (understanding written sentences and paragraphs in work related documents).

Writing (communicating effectively in writing as appropriate for the needs of the audience).

Learning strategies (selecting and using training/instructional methods and procedures appropriate for the situation when learning or teaching new things).

Complex problem solving (identifying complex problems and reviewing related information to develop and evaluate options and implement solutions).

Persuasion (persuading others to change their minds or behavior).

Instructing (teaching others how to do something).

Time management (managing one's own time and the time of others).

Equipment selection (determining the kind of tools and equipment needed to do a job).

Negotiation (bringing others together and trying to reconcile differences).

Operations Analysis (analyzing needs and product requirements to create a design).

Quality control analysis (conducting tests and inspections of products, services, or processes to evaluate quality or performance).

Judgment and decision making (considering the relative costs and benefits of potential actions to choose the most appropriate one).

Systems analysis (determining how a system should work and how changes in conditions, operations, and the environment will affect outcomes).

Systems evaluation (identifying measures or indicators of system performance and the actions needed to improve or correct performance, relative to the goals of the system).

Introduction 2

The following list contains the skills that you indicated as "essential" previously. Please indicate from this list, the skills you believe are the 30 most important for career readiness by selecting "yes" for most important and "no" if not most important. You can move back and forth between the skills/pages by utilizing the arrow buttons.

Skills 2

Please indicate from this list, the skills you believe are the 30 most important for career readiness by selecting "yes" for most important and "no" if not most important. You can move back and forth between the skills/pages by utilizing the arrow buttons.

Using electronic mail.

Writing letters and memos.

Being responsible for outcomes and results of other workers.

Decision making (making decisions that affect other people, the financial resources, and/or the image and reputation of the organization).

Importance of being exact or accurate.

Time pressure (the job requires the worker to meet strict deadlines).

Analyzing data and information (identifying the underlying principles, reasons, or facts of information by breaking down information or data into separate parts).

Making decisions and solving problems (analyzing information and evaluating results to choose the best solution and solve problems).

Organizing, planning, and prioritizing work (developing specific goals and plans to prioritize, organize, and accomplish your work).

Processing information (compiling, coding, categorizing, calculating, tabulating, auditing, or verifying information or data).

Updating and using relevant knowledge (keeping up-to-date technically and applying new knowledge to your job).

Evaluating information to determine compliance with standards (using relevant information and individual judgment to determine whether events or processes comply with laws, regulations, or standards).

Scheduling work and activities (scheduling events, programs, and activities, as well as the work of others).

Judging the qualities of things, services, or people (assessing the value, importance, or quality of things or people).

Thinking creatively (developing, designing, or creating new applications, ideas, relationships, systems, or products, including artistic contributions).

Getting information (observing, receiving, and otherwise obtaining information from all relevant sources).

Monitor processes, materials, or surroundings (monitoring and reviewing information from materials, events, or the environment, to detect or assess problems).

Estimating the quantifiable characteristics of products, events, or information (estimating sizes, distances, and quantities; or determining time, costs, resources, or materials needed to perform a work activity).

Identifying objects, actions, and events (identifying information by categorizing, estimating, recognizing differences or similarities, and detecting changes in circumstances or events).

Interpreting the meaning of information for others (translating or explaining what information means and how it can be used).

Documenting/recording information (entering, transcribing, recording, storing, or maintaining information in written or electronic/magnetic form).

Interacting with computers (using computers and computer systems, including hardware and software, to program, write software, set up functions, enter data, or process information).

Customer and personal service (knowledge of principles and processes for providing customer and personal services. This includes customer needs assessment, meeting quality standards for services, and evaluation of customer satisfaction).

Problem sensitivity (the ability to tell when something is wrong or is likely wrong. It does not involve solving the problem, only recognizing there is a problem).

Information ordering (the ability to arrange things or actions in a certain order or pattern according to a specific rule or set of rules).

Oral comprehension (the ability to listen to and understand information and ideas presented through spoken words and sentences).

Written comprehension (the ability to read and understand information and ideas presented in writing).

Written expression (the ability to communicate information and ideas in writing so others will understand).

Deductive reasoning (the ability to apply general rules to specific problems to produce answers that make sense).

Mathematical reasoning (the ability to choose the right mathematical methods or formulas to solve a problem).

Inductive reasoning (the ability to combine pieces of information to form general rules or conclusions).

Selective attention (the ability to concentrate on a task over a period of time without being distracted).

Fluency of ideas (the ability to come up with a number of ideas about a topic).

Memorization (the ability to remember information such as words, numbers, pictures, and procedures).

Category flexibility (the ability to generate or use different sets of rules for combining or grouping things in different ways).

Originality (the ability to come up with unusual or clever ideas about a given topic or situation, or to develop creative ways to solve a problem.

Time sharing (the ability to shift back and forth between two or more activities or sources of information).

Visualization (the ability to imagine how something will look after it is moved around or when its parts are moved or rearranged).

Flexibility of closure (the ability to identify or detect a known pattern, figure, object, word, or sound that is hidden in other distracting material).

Perceptual speed (the ability to quickly and accurately compare similarities and differences among sets of letters, numbers, objects, pictures, or patterns.

Speed of closure (the ability to quickly make sense of, combine, and organize information into meaningful patterns).

Number facility (the ability to add, subtract, multiply, or divide quickly and correctly).

Speech recognition (the ability to identify and understand the speech of another person).

Active learning (understanding the implications of new information for both current and future problem-solving and decision-making.

Active listening (giving full attention to what other people are saying and taking time to understand the points being made).

Critical thinking (using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions or approaches to problems).

Mathematics (using mathematics to solve problems).

Monitoring (monitoring/assessing performance of yourself, other individuals, or organizations to make improvements or take corrective action).

Reading comprehension (understanding written sentences and paragraphs in work related documents).

Writing (communicating effectively in writing as appropriate for the needs of the audience).

Learning strategies (selecting and using training/instructional methods and procedures appropriate for the situation when learning or teaching new things).

Complex problem solving (identifying complex problems and reviewing related information to develop and evaluate options and implement solutions).

Persuasion (persuading others to change their minds or behavior).

Instructing (teaching others how to do something).

Time management (managing one's own time and the time of others).

Equipment selection (determining the kind of tools and equipment needed to do a job).

Negotiation (bringing others together and trying to reconcile differences).

Operations Analysis (analyzing needs and product requirements to create a design).

Quality control analysis (conducting tests and inspections of products, services, or processes to evaluate quality or performance).

Judgment and decision making (considering the relative costs and benefits of potential actions to choose the most appropriate one).

Systems analysis (determining how a system should work and how changes in conditions, operations, and the environment will affect outcomes).

Systems evaluation (identifying measures or indicators of system performance and the actions needed to improve or correct performance, relative to the goals of the system).

Introduction 3

The following list contains the 30 skills that you indicated as most important. Please review the skills and rank order them from 1 through 30, with 1 being the "most important" and 30 being the "least important." Please type the appropriate number in the writing space provided. After you have rank ordered the skills, please operationalize the top 15 skills by providing a career-related example. Please utilize the writing space provided. You can move back and forth between the skills/pages by utilizing the arrow buttons.

Skills 3

Please review the skills and rank order them from 1 through 30, with 1 being the most important and 30 being the least important. After you have rank ordered the skills, please operationalize the 15 top skills utilizing the writing space provided.

Using electronic mail.

Writing letters and memos.

Being responsible for outcomes and results of other workers.

Decision making (making decisions that affect other people, the financial resources, and/or the image and reputation of the organization).

Importance of being exact or accurate.

Time pressure (the job requires the worker to meet strict deadlines).

Analyzing data and information (identifying the underlying principles, reasons, or facts of information by breaking down information or data into separate parts).

Making decisions and solving problems (analyzing information and evaluating results to choose the best solution and solve problems).

Organizing, planning, and prioritizing work (developing specific goals and plans to prioritize, organize, and accomplish your work).

Processing information (compiling, coding, categorizing, calculating, tabulating, auditing, or verifying information or data).

Updating and using relevant knowledge (keeping up-to-date technically and applying new knowledge to your job).

Evaluating information to determine compliance with standards (using relevant information and individual judgment to determine whether events or processes comply with laws, regulations, or standards).

Scheduling work and activities (scheduling events, programs, and activities, as well as the work of others).

Judging the qualities of things, services, or people (assessing the value, importance, or quality of things or people).

Thinking creatively (developing, designing, or creating new applications, ideas, relationships, systems, or products, including artistic contributions).

Getting information (observing, receiving, and otherwise obtaining information from all relevant sources).

Monitor processes, materials, or surroundings (monitoring and reviewing information from materials, events, or the environment, to detect or assess problems).

Estimating the quantifiable characteristics of products, events, or information (estimating sizes, distances, and quantities; or determining time, costs, resources, or materials needed to perform a work activity).

Identifying objects, actions, and events (identifying information by categorizing, estimating, recognizing differences or similarities, and detecting changes in circumstances or events).

Interpreting the meaning of information for others (translating or explaining what information means and how it can be used).

Documenting/recording information (entering, transcribing, recording, storing, or maintaining information in written or electronic/magnetic form).

Interacting with computers (using computers and computer systems, including hardware and software, to program, write software, set up functions, enter data, or process information).

Customer and personal service (knowledge of principles and processes for providing customer and personal services. This includes customer needs assessment, meeting quality standards for services, and evaluation of customer satisfaction).

Problem sensitivity (the ability to tell when something is wrong or is likely wrong. It does not involve solving the problem, only recognizing there is a problem).

Information ordering (the ability to arrange things or actions in a certain order or pattern according to a specific rule or set of rules).

Oral comprehension (the ability to listen to and understand information and ideas presented through spoken words and sentences).

Written comprehension (the ability to read and understand information and ideas presented in writing).

Written expression (the ability to communicate information and ideas in writing so others will understand).

Deductive reasoning (the ability to apply general rules to specific problems to produce answers that make sense).

Mathematical reasoning (the ability to choose the right mathematical methods or formulas to solve a problem).

Inductive reasoning (the ability to combine pieces of information to form general rules or conclusions).

Selective attention (the ability to concentrate on a task over a period of time without being distracted).

Fluency of ideas (the ability to come up with a number of ideas about a topic).

Memorization (the ability to remember information such as words, numbers, pictures, and procedures).

Category flexibility (the ability to generate or use different sets of rules for combining or grouping things in different ways).

Originality (the ability to come up with unusual or clever ideas about a given topic or situation, or to develop creative ways to solve a problem.

Time sharing (the ability to shift back and forth between two or more activities or sources of information).

Visualization (the ability to imagine how something will look after it is moved around or when its parts are moved or rearranged).

Flexibility of closure (the ability to identify or detect a known pattern, figure, object, word, or sound that is hidden in other distracting material).

Perceptual speed (the ability to quickly and accurately compare similarities and differences among sets of letters, numbers, objects, pictures, orpatterns.

Speed of closure (the ability to quickly make sense of, combine, and organize information into meaningful patterns).

Number facility (the ability to add, subtract, multiply, or divide quickly and correctly).

Speech recognition (the ability to identify and understand the speech of another person).

Active learning (understanding the implications of new information for both current and future problem-solving and decision-making.

Active listening (giving full attention to what other people are saying and taking time to understand the points being made).

Critical thinking (using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions or approaches to problems).

Mathematics (using mathematics to solve problems).

Monitoring (monitoring/assessing performance of yourself, other individuals, or organizations to make improvements or take corrective action).

Reading comprehension (understanding written sentences and paragraphs in work related documents).

Writing (communicating effectively in writing as appropriate for the needs of the audience).

Learning strategies (selecting and using training/instructional methods and procedures appropriate for the situation when learning or teaching new things).

Complex problem solving (identifying complex problems and reviewing related information to develop and evaluate options and implement solutions).

Persuasion (persuading others to change their minds or behavior).

Instructing (teaching others how to do something).

Time management (managing one's own time and the time of others).

Equipment selection (determining the kind of tools and equipment needed to do a job).

Negotiation (bringing others together and trying to reconcile differences).

Operations Analysis (analyzing needs and product requirements to create a design).

Quality control analysis (conducting tests and inspections of products, services, or processes to evaluate quality or performance).

Judgment and decision making (considering the relative costs and benefits of potential actions to choose the most appropriate one).

Systems analysis (determining how a system should work and how changes in conditions, operations, and the environment will affect outcomes).

Systems evaluation (identifying measures or indicators of system performance and the actions needed to improve or correct performance, relative to the goals of the system).

Appendix B:
Participant Background Data

Background variables	Frequency (%)
Age	
16	6 (37.5)
17	4 (25.0)
18	6 (37.5)
Gender	
Female	4 (25.0)
Male	12 (75.0)
Grade	
11	9 (56.3)
12	7 (43.7)
Other home language spoken	
Never	3 (18.7)
Once in while	3 (18.7)
About half the time	2 (12.5)
Almost all the time	8 (50.0)
Working computers/laptops at home	
0	2 (12.5)
1	5 (31.2)
2	4 (25.0)
3	3 (18.7)
4 or more	2 (12.5)
Ethnicity	
American Indian or Alaska Native	0 (0.0)
Asian or Pacific Islander	1 (6.2)
Black or African American	1 (6.2)
Hispanic/Latina/o	9 (50.0)
White or Caucasian	3 (18.7)
Other	2 (12.5)

Note. One student did not complete the background survey (*n* = 16).

Appendix C:
Background Survey

1. How **hard** was this test compared to other tests you have taken this year?

☐ Easier

☐ About the same

☐ Harder

☐ Much Harder

2. How much **effort** did you put into this test compared to other tests you have taken this year?

☐ Hardly any effort

☐ A little effort

☐ Some effort

☐ A lot of effort

Student Background

1. How old are you?

☐ 16 ☐ 17 ☐ 18 ☐ 19 or older

2. Grade: ☐ 11th ☐ 12th

3. Gender:_____

4. How often do people in your home talk to each other in a language other than English?

☐ Never ☐ Once in a while ☐ About half of the time ☐ All or most of the time

5. How many computers or laptops are in your home? (Only count the ones that are working)

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 or more

6. With which of the following groups do you most strongly identify? (It's okay to pick more than one!)

- ☐ American Indian or Alaska Native
- ☐ Asian or Pacific Islander
- ☐ Black or African American
- ☐ Hispanic / Latino/a
- ☐ White or Caucasian
- ☐ Other, please describe. _____

Appendix D:
Cognitive Lab Protocol

UCLA Cog Lab Study Participant Code_____

EAG/CDE: Cognitive Lab Study Protocol

Date:	Start time:
Researcher Initials:	Recorder Number & Folder:

DIRECTIONS/SCRIPT

Thanks for volunteering to be in our study!

Today, you are going to be asked to solve a few math and English language arts problems. The reason we are doing this study is to test out the problems to see how different students solve problems.

For each problem, we want you to try your best, but don't worry if you don't think you know the right answer. Some problems are harder than others and some of them may not be familiar to you depending on what you've done in class this year.

As you solve the problems, you should say out loud what you are doing while you do it. Here's how it works:

- You'll answer the problem and talk out loud as you go, so I can hear what you're thinking.
- I'll ask you questions like: "Tell me what you did to solve the problem" or "Was that problem easy, medium, or hard?"
- Try your best to tell me what you're thinking when it crosses your mind or why you chose to do a procedure or strategy. If you need a few seconds to think, that's okay. Just let me know.
- Then you'll move onto the next problem. Unfortunately, we won't be able to tell you if your answers are correct or incorrect.

I'm going to take notes on what you say. But just in case I miss something, I am going to audio record you. You can ask me to stop recording at any time. After you've completed the study, I will stop the audio recorder, and you'll complete a brief survey about your background.

Now I'm going to show you what I mean by "think aloud."

[Example of think aloud process: Researcher solves math problem below while thinking aloud.]

$$\begin{array}{r} 2357 \\ +\ 652 \\ \hline \end{array}$$

Item	Claim	Domain	Target	DOK	CCSS-MC	CCSS-MP
#13	3	A-REI	E	3	A-REI.B	3

Emily is solving the equation $2(x + 9) = 4(x + 7) + 2$. Her steps are shown.

Part A
Click on the first step in which Emily made an error.

Part B
Click on the solution to Emily's original equation.

Part A

Step 1: $2(x + 9) = 4(x + 7) + 2$

Step 2: $2x + 18 = 4x + 28 + 2$

Step 3: $2x + 18 = 4x + 26$

Step 4: $-8 = 2x$

Step 5: $-4 = x$

Part B

−10.5 −6 −2 0 2 4.5 8

UCLA Cog Lab Study Participant Code_____

Item #13 (Math)	Start time:	End time:

Researcher notes each step to solve problem below:.

When student is done, say: *Tell me in words what you did to solve the problem.*

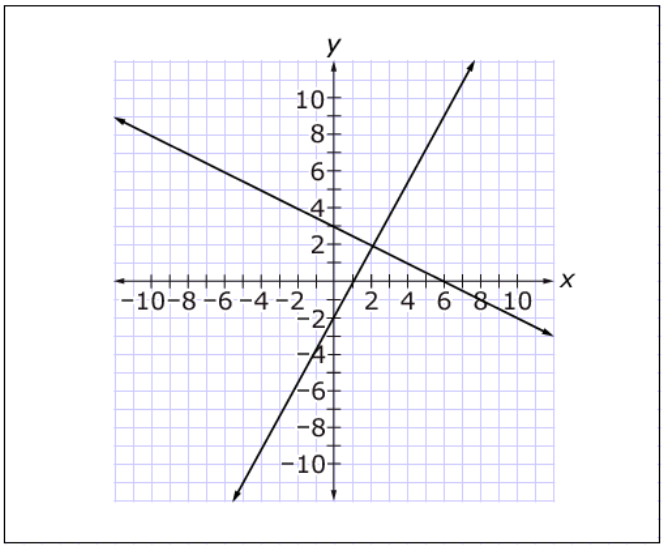
Instruction reading	☐ read aloud ☐ read silently ☐ did not appear to read
Problem reading	☐ read aloud ☐ read silently ☐ did not appear to read
Answer reading	☐ read aloud ☐ read silently ☐ did not appear to read
Did student re-read instructions or problem before answering?	☐ no ☐ yes Describe:
Prompts during silence:	*What made you do that in this problem?* *Tell me a little more about that step.* *What was going on through your mind just now?* *What were you thinking about just now?* *You can also tell me if you need more time to think.*
Did you understand what the question was asking you to do?	☐ no ☐ yes ☐ some/a little
Answer attempt	☐ incorrect ☐ correct ☐ no attempt
Was this problem easy, medium, or hard?	☐ easy ☐ medium ☐ hard
If medium or hard *why did you think it was medium or hard*?	☐ vocabulary ☐ didn't know what it was asking ☐ other (describe):
***IF SKIPPED** Why did you not answer the problem?*	☐ vocabulary ☐ didn't know what it was asking ☐ didn't know how to solve it ☐ other (describe):

Item	Claim	Domain	Target	DOK	CCSS-MC	CCSS-MP
#16	1	A-REI	J	1	A-REI.D.12	N/A

Click on the region of the graph that contains the solution set of the system of linear inequalities.

$y \leq -\frac{1}{2}x + 3$

$y \geq 2x - 2$



UCLA Cog Lab Study Participant Code_____

Item #16 (Math)	Start time:	End time:

Researcher notes each step to solve problem below:.

When student is done, say: *Tell me in words what you did to solve the problem.*

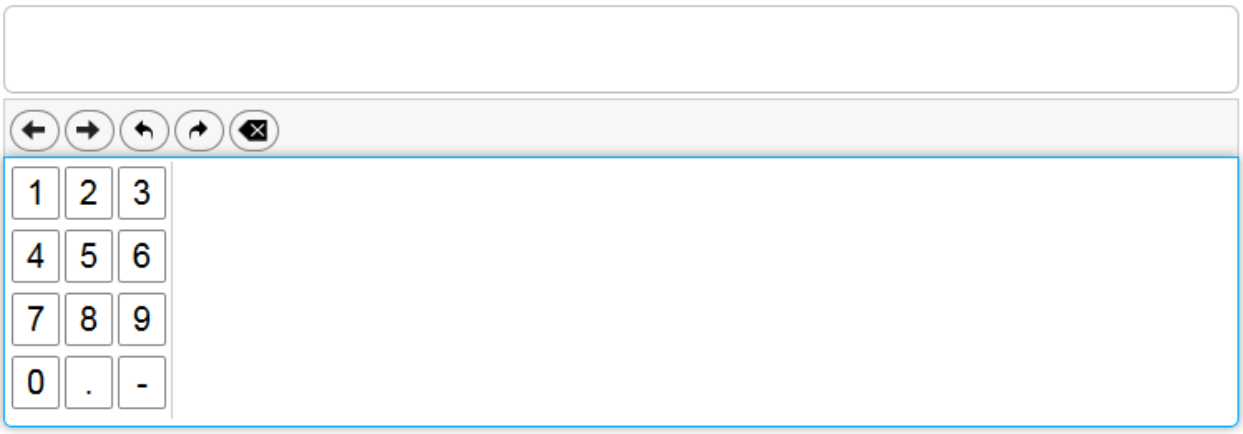
Instruction reading	☐ read aloud ☐ read silently ☐ did not appear to read
Problem reading	☐ read aloud ☐ read silently ☐ did not appear to read
Answer reading	☐ read aloud ☐ read silently ☐ did not appear to read
Did student re-read instructions or problem before answering?	☐ no ☐ yes Describe:
Prompts during silence:	*What made you do that in this problem?* T*ell me a little more about that step.* *What was going on through your mind just now?* *What were you thinking about just now?* *You can also tell me if you need more time to think.*
Did you understand what the question was asking you to do?	☐ no ☐ yes ☐ some/a little
Answer attempt	☐ incorrect ☐ correct ☐ no attempt
Was this problem easy, medium, or hard?	☐ easy ☐ medium ☐ hard
If medium or hard *why did you think it was medium or hard*?	☐ vocabulary ☐ didn't know what it was asking ☐ other (describe):
***IF SKIPPED** Why did you not answer the problem?*	☐ vocabulary ☐ didn't know what it was asking ☐ didn't know how to solve it ☐ other (describe):

Item	Claim	Domain	Target	DOK	CCSS-MC	CCSS-MP
#23	4	A-REI	E	3	A-REI.B.3	6

Nina has some money saved for a vacation she has planned.

- The vacation will cost a total of $1600.
- She will put $150 every week into her account to help pay for the vacation.
- She will have enough money for the vacation in 8 weeks.

If Nina was able to save $200 a week instead of $150 a week, how many fewer weeks would it take her to save enough money for the vacation? Enter the result in the response box.



UCLA Cog Lab Study Participant Code_____

Item #23 (Math)	Start time:	End time:

Researcher notes each step to solve problem below:

When student is done, say: *Tell me in words what you did to solve the problem.*

Instruction reading	☐ read aloud ☐ read silently ☐ did not appear to read
Problem reading	☐ read aloud ☐ read silently ☐ did not appear to read
Answer reading	☐ read aloud ☐ read silently ☐ did not appear to read
Did student re-read instructions or problem before answering?	☐ no ☐ yes Describe:
Prompts during silence:	*What made you do that in this problem?* T*ell me a little more about that step.* *What was going on through your mind just now?* *What were you thinking about just now?* *You can also tell me if you need more time to think.*
Did you understand what the question was asking you to do?	☐ no ☐ yes ☐ some/a little
Answer attempt	☐ incorrect ☐ correct ☐ no attempt
Was this problem easy, medium, or hard?	☐ easy ☐ medium ☐ hard
If medium or hard *why did you think it was medium or hard*?	☐ vocabulary ☐ didn't know what it was asking ☐ other (describe):
***IF SKIPPED** Why did you not answer the problem?*	☐ vocabulary ☐ didn't know what it was asking ☐ didn't know how to solve it ☐ other (describe):

Item	Claim	Domain	Target	DOK	CCSS-MC	CCSS-MP
#29	2	G-SRT	D	2	G-SRT.C.8	3

Emma is standing 10 feet away from the base of a tree and tries to measure the angle of elevation to the top. She is unable to get an accurate measurement, but determines that the angle of elevation is between 55 degrees and 75 degrees.

Decide whether each value given in the table is a reasonable estimate for the tree height. Select Reasonable or Not Reasonable for each height.

	Reasonable	Not Reasonable
4.2 feet	☐	☐
14.7 feet	☐	☐
24.4 feet	☐	☐
33.9 feet	☐	☐
39.1 feet	☐	☐
58.7 feet	☐	☐

Item #29 (Math)	Start time:	End time:

Researcher notes each step to solve problem below:

When student is done, say: *Tell me in words what you did to solve the problem.*

Instruction reading	☐ read aloud ☐ read silently ☐ did not appear to read
Problem reading	☐ read aloud ☐ read silently ☐ did not appear to read
Answer reading	☐ read aloud ☐ read silently ☐ did not appear to read
Did student re-read instructions or problem before answering?	☐ no ☐ yes Describe:
Prompts during silence:	*What made you do that in this problem?* T*ell me a little more about that step.* *What was going on through your mind just now?* *What were you thinking about just now?* *You can also tell me if you need more time to think.*
Did you understand what the question was asking you to do?	☐ no ☐ yes ☐ some/a little
Answer attempt	☐ incorrect ☐ correct ☐ no attempt
Was this problem easy, medium, or hard?	☐ easy ☐ medium ☐ hard
If medium or hard *why did you think it was medium or hard*?	☐ vocabulary ☐ didn't know what it was asking ☐ other (describe):
***IF SKIPPED** Why did you not answer the problem?*	☐ vocabulary ☐ didn't know what it was asking ☐ didn't know how to solve it ☐ other (describe):

ELA DIRECTIONS/SCRIPT

Now you will work on some English language arts problems.

Like the math problems, the reason we are doing this study is to test out the problems to find out if directions or wording is confusing and to see how different students solve each problem. As you solve the problems, you should say out loud what you are doing while you do it. However, if you need to read the passage or questions to yourself silently, it's okay. And if you need to think quietly about the problem before you answer, that's okay too. Just let us know if you need some time to think quietly.

For each problem, we want you to try your best, but don't worry if you don't think you know the right answer.

Item	Grade	Claim	Target	DOK	Standard(s)
#28	11	4	3	2	WLiteracy.8

A student is writing a research report about the Iditarod Trail and annual dogsled race in northern Alaska. Read the paragraphs from her report and the directions that follow.

Alaska's Iditarod National Historic Trail has a long and rich history of travel and trade routes, including during the Gold Rush of the nineteenth century. Located close to the Arctic Circle, the one-thousand-mile trail system is known for its rugged terrain and harsh winters. But it is primarily known for the internationally famous Iditarod Trail Sled Dog Race, commonly referred to as simply the Iditarod. The Iditarod was founded in the 1970s by a few Alaskans determined to commemorate the state's rich history of dogsledding and to honor a famous event: the halt of a serious diphtheria epidemic in 1925, when a crucial medical serum was delivered by dogsled to the Alaskan citizens in need.

The Iditarod is an epic challenge for man and man's best friend: it's a grueling 1150-mile trek from Anchorage to Nome, Alaska. Although held in early spring, the conditions are extremely harsh. Each team consists of a human captain (also called a musher) leading a team of more than a dozen sled dogs—mostly Siberian Huskies, Alaskan Malamutes, or Eskimo Dogs. Together, they brave frigid conditions, including frozen tundra and icy forests. "The Last Great Race on Earth," as it is also called, starts during the first week in March with teams generally reaching Nome in 10-15 days.

Select **two** credible sources that would **most likely** give the student more information for her paragraphs.

- [ ] "Snow Huskies"

  Commercial Film

  This 1998 PG-13 film is a heartwarming tale about a team of ragtag Alaskan Malamutes who save their owner's ranch by winning the nation's premier dogsled race.

- [ ] www.northernparks.com

  This website has information about all of Alaska's state parks for tourists planning a visit to the country's largest state.

- [ ] www.racetheiditarod.com

  Our business will provide you with all the information you will need to race in the Iditarod. Let our team of experts help you register your dogs, find local veterinarians, and secure housing and supplies.

- [ ] *Balto: Our Hero* by C.R. Benson

  This book, first published in 1957 by children's literature author C.R. Benson, tells the true dramatic tale of Balto, the lead sled dog in a race against time to deliver serum during the 1925 diphtheria epidemic in Nome, Alaska.

- [ ] "The Iditarod Trail Dog Sled Race"

  United States Geographical Society

  www.usgeog.com/iditarod

  In celebration of its 100-year history of documenting our planet's uniqueness, our staff ventured to the 49th state to cover the Iditarod. This article chronicles the history of one of the most grueling competitions in the world, nicknamed "The Last Great Race."

- [ ] "MUSH! Alaska's Most Challenging Race"

  Documentary Film

  This 2007 six-hour miniseries documentary by award-winning filmmaker, Bernard Kerns, details the history and route of the famous Iditarod.

Item #28 (ELA)	Start time:	End time:

Passage reading	☐ student read aloud ☐ student read silently ☐ did not appear to read
Task/item reading	☐ student read aloud ☐ student read silently ☐ did not appear to read
Re-reading of passage	☐ no ☐ yes
Prompts during silence:	*What was going on through your mind just now?* *What were you thinking about just now?* *You can also tell me if you need more time to think.*
Did you understand what the question was asking you to do?	☐ no ☐ yes ☐ some/a little
Answer attempt	☐ incorrect ☐ correct ☐ no attempt
Probe as necessary:	*Tell me in your own words what you did to answer this question.* *How did you make that decision?* or *How did you decide on that answer?*
Was this problem easy, medium, or hard?	☐ easy ☐ medium ☐ hard
If medium or hard *why did you think it was medium or hard*?	☐ vocabulary ☐ didn't know what it was asking ☐ other_____
IF SKIPPED*Why did you not answer the problem?*	☐ vocabulary ☐ didn't know what it was asking ☐ other_____

Item	Grade	Claim	Target	DOK	Standard(s)
#30	11	4	4	2	WLiteracy.8

A student is writing an argumentative report about the causes of sleepwalking. She found possible sources for her report. Read the sources and the directions that follow.

Source 1: "What is Sleepwalking?" by Mat Valerio

Sleepwalking is a sleep disorder that causes a person to get up and walk while still asleep, usually during the deepest stages of sleep. While sleepwalking, a person does not often respond when someone asks a question or touches him or her. Though sometimes a sleepwalker will verbally respond, the response will make no sense. A sleepwalking episode can include the person just walking quietly around a room or demonstrating very agitated behavior and trying to "escape" the room.

Source 2: "Is Sleepwalking Genetic?" by Chris Shue

Sleepwalking occurs most often during middle childhood through adolescence. Some adults also sleepwalk, but it is much more common among children. Genetics may play a big role in determining who will be a sleepwalker as the tendency runs in families. Environmental and medical conditions also may contribute to sleepwalking episodes. Sleepwalking was described in some of the earliest recorded medical literature, even before Hippocrates, the "father of medicine," lived.

Source 3: "Am I a Sleepwalker?" by Angelique Kandar

Research has shown that a variety of factors contribute to sleepwalking episodes. Sleepwalkers are usually operating under a severe lack of sleep or have an irregular or hectic sleep schedule. Often, they are under great amounts of stress or anxiety. Some medical conditions, such as abnormal heart rhythms, nighttime seizures, and **sleep apnea**, have been known to cause sleepwalking. Sometimes a person who has taken a certain medication experiences a sleepwalking episode.

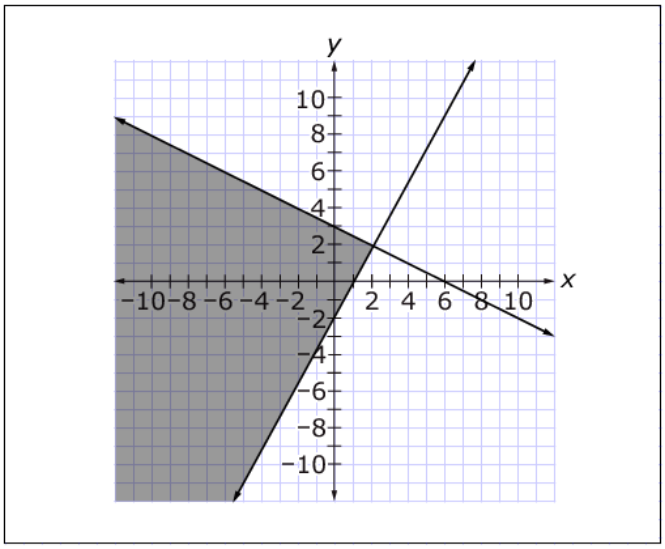
The student wrote down some claims to use in her report. Look at the claims on the table. Determine if the information in the sources supports each claim. Click on the boxes to show the claims that each source supports. A source may have more than one box selected.

	Source 1	Source 2	Source 3
Claim 1: Some outside influences make a person more likely to sleepwalk.	☐	☐	☐
Claim 2: If your mother was a sleepwalker, it is more likely that you will be, too.	☐	☐	☐
Claim 3: When people are sleepwalking they are not aware of their surroundings.	☐	☐	☐
Claim 4: Going to bed at the same time every night can help some children not to sleepwalk.	☐	☐	☐

Item #30 (ELA)	Start time:	End time:

Passage reading	☐ student read aloud ☐ student read silently ☐ did not appear to read
Task/item reading	☐ student read aloud ☐ student read silently ☐ did not appear to read
Re-reading of passage	☐ no ☐ yes
Prompts during silence:	*What was going on through your mind just now?* *What were you thinking about just now?* *You can also tell me if you need more time to think.*
Did you understand what the question was asking you to do?	☐ no ☐ yes ☐ some/a little
Answer attempt	☐ incorrect ☐ correct ☐ no attempt
Probe as necessary:	*Tell me in your own words what you did to answer this question.* *How did you make that decision?* or *How did you decide on that answer?*
Was this problem easy, medium, or hard?	☐ easy ☐ medium ☐ hard
If medium or hard *why did you think it was medium or hard*?	☐ vocabulary ☐ didn't know what it was asking ☐ other_____
IF SKIPPED*Why did you not answer the problem?*	☐ vocabulary ☐ didn't know what it was asking ☐ other_____

Appendix E:
Answer Keys and Rubrics

Item	Key/Exemplar	Rubric
Math 13	**Part A** Step 1: $2(x + 9) = 4(x + 7) + 2$ Step 2: $2x + 18 = 4x + 28 + 2$ **Step 3: $2x + 18 = 4x + 26$** Step 4: $-8 = 2x$ Step 5: $-4 = x$ **Part B** −10.5 **−6** −2 0 2 4.5 8	(2 points) The student selects Step 3 in Part B and indicates -6 as the correct response in Part B. (1 point) The student gets Part A or Part B correct but not both.
Math 16		(1 point) The student correctly selects the region containing the solution set.
Math 23	2	(1 point) Student enters the correct number of weeks.

Item	Key/Exemplar	Rubric
Math 29		(1 point) The student correctly identifies the reasonable estimates.
ELA 28	E, F	(1 point) The student selects the correct two options.
ELA 30	Claim 1: Sources 2 and 3 Claim 2: Source 2 Claim 3: Source 1 Claim 4: Source 3	(1 point) The student selects the correct five options.
ELA 22 (bonus)	Part A: B Part B: A	(1 point) The student selects the correct option for Part A and selects the correct option for Part B.

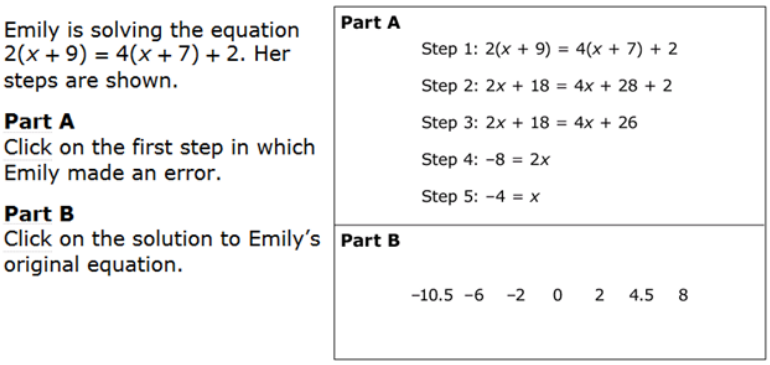
Appendix F:
Correct and Incorrect Items by Student Participant

	Math				ELA		
ID	Item 13	Item 16	Item 23	Item 29	Item 28	Item 30	Total Correct
B103	Correct	Correct	Correct	X	Correct	X	4
A103	Correct	Correct	X	Correct	Correct	X	4
B106	Correct	Correct	Correct	X	X	X	3
B108	Correct	Correct	Correct	X	X	X	3
B105	Correct	X	X	X	Correct	X	2
B201	Correct	X	Correct	X	X	X	2
C102	Correct	—	X	X	X	Correct	2
B101	X	X	X	X	Correct	Correct	2
A104	Correct	Correct	X	X	X	N/A	2
B102	X	Correct	Correct	—	X	N/A	2
B104	Correct	—	X	Correct	X	N/A	2
C101	Correct	X	X	X	X	X	1
B107	X	Correct	X	X	X	X	1
C103	X	—	X	X	Correct	X	1
A102	X	X	X	X	X	X	0
A106	X	X	X	X	X	X	0
A105	X	X	X	—	X	X	0

Note. — indicates participant skipped the item. X indicates incorrect. N/A indicates participant did not reach the item in the allotted time.

Appendix G:
Results by Item

Math Item 13



Item 13 of the math section was composed of two parts. In Part A, students were asked to identify the step where an error occurred in an incorrectly solved equation. This was a multiple-choice item with five answer options. In Part B students were asked to solve the equation. Most students (93.3%) reported that they understood Item 13. One student (not included in the previous count) reported that he understood Part A, and only somewhat understood Part B. The mean perceived difficulty on a scale from 1 to 3 (1 = *easy*, 3 = *hard*) for this item was 1.2 (*n* = 13), indicating that students tended to perceive this item as easy.

All students in our sample attempted this problem (*N* = 17). Of those, 10 students (58.8%) solved both Part A and Part B correctly (two points, based on Smarter Balanced scoring rubric) and seven students (41.2%) solved just one of the parts correctly (one point, based on Smarter Balanced scoring rubric). For most students (five out of the eight who answered incorrectly), actually solving the problem (Part B) appeared to be more difficult than identifying the step where an error occurred (Part A). This could be because students needed to engage in mathematical computation in Part B of the item, which involves the features of mathematical reasoning, number facility, and the importance of being exact or accurate.

Students who completed Part B of the item correctly demonstrated mathematical reasoning by selecting the correct mathematical method to complete the computation. They further demonstrated number facility by implementing the mathematical methods correctly. Lastly, they demonstrated the importance of being exact or accurate by identifying the correct error step in Part A and performing the mathematical computation accurately in Part B. The majority of students who did not complete the item successfully did not perform the calculation in Part B correctly (e.g., some students subtracted instead of added) indicating a lack of background knowledge or misconception about the mathematical computation needed.

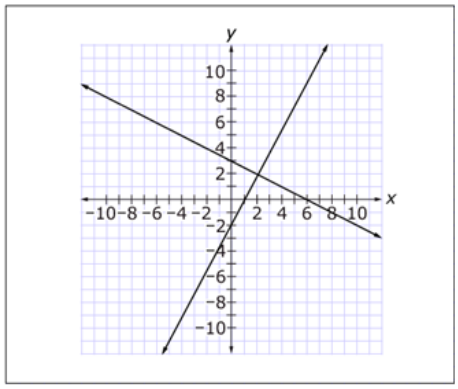
For example, one student (C101) initially made a careless computational error; she added instead of subtracted and when she did not see her answer among the choices, she looked back at her work to find the mistake, and re-did the problem to get it correct. Another student (B101) correctly recognized the error in Part A, but then went on to solve Part B with that error (using 26 instead of 30) and failed to solve Part B correctly. Some of the students also selected the incorrect step as the error step in Part A, similarly indicating a lack of knowledge related to correct computation.

Math Item 16

Click on the region of the graph that contains the solution set of the system of linear inequalities.

$y \leq -\frac{1}{2}x + 3$

$y \geq 2x - 2$



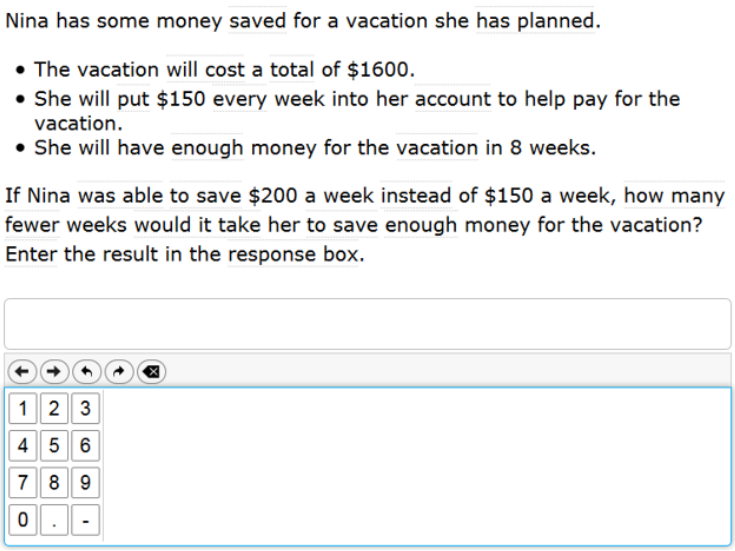
Item 16 was a hotspot type item where students were asked to select the region of the graph that contained the solution set for the system of linear inequalities provided. Only one region was correct (left bottom); however, students were able to select more than one region. Most students (78.5%) reported they understood Item 16. The mean perceived difficulty for this item was 1.8 (*N* = 17), indicating that on average students perceived the item to have medium difficulty.

Fourteen students attempted to complete this item. Of those 14 students, 50% solved the item correctly. Of the items rated, this item contained the fewest features. However, one feature that made this item unique was interacting with computers. Computer interaction afforded a novel way for students to provide their answer, which confused some students about the objective of the problem as well as how to provide their answer (i.e., some students selected more than one region). Additional reasons why students may not have solved this item correctly may include lack of knowledge or ability related to the following features: analyzing data or information, mathematical reasoning, and processing information.

Students who completed the item successfully first analyzed the set of linear inequalities to determine how each aspect of the inequalities could be represented in the graph (i.e., analyzing data or information). Second, these students determined that they needed to utilize the rise, run, and slope of the inequalities as data points to verify the graph region (i.e., mathematical reasoning). Lastly, the students engaged in the verification process utilizing the rise, run, and, slope data points (i.e., processing information).

For example B106 graphed the two inequalities using the provided scratch paper and verified the two lines presented on the screen (processing information). However, he was unsure of what to do next, due to a lack of mathematics background knowledge and ultimately guessed (incorrectly). When asked about the item's difficulty, he responded with, "I feel like it would be easy if I could remember what to do." Students who possessed the math background knowledge (such as A104 and B108) were able to quickly and correctly solve the problem without the use of scratch paper (by analyzing the data or information and using mathematical reasoning in their heads).

Math Item 23



Item 23 required students to solve a word problem with an unknown variable, where students were not provided with all of the information needed to solve the problem. This was an open-ended item that required students to provide a constructed response. Most students (88.2%) reported they understood the item. The mean perceived difficulty for this item was 2.1 (*N* = 17), indicating that the students tended to perceive this item to be of medium difficulty.

Students solved the problem and used the numbers provided to enter the information. All students in our sample attempted this problem (*N* = 17). Of those, five students (29.4%) solved this problem correctly and 12 students solved this problem incorrectly (70.6%). The main reasons why students did not solve this problem correctly were related to the features of problem sensitivity, mathematical reasoning, deductive reasoning, and processing information. Related to problem sensitivity, a unique feature to this item, students did not solve the item correctly when they failed to recognize that something was wrong or missing in the problem. Similarly, these students did not engage in deductive reasoning to determine that the word problem included an unmentioned amount of savings, which prevented them from adding the amounts up correctly. These students also had difficulty with mathematical reasoning, in that they were not clear about what mathematical method to use to solve the problem.

Additionally, these students also neglected to engage in processing the information provided in the item by auditing and verifying the data and numbers.

On the other hand, students who completed the item correctly, appeared to have an "a-ha" moment where they first realized that something was wrong with the problem (i.e., problem sensitivity). After having audited and verified the data provided in the item (i.e., processing information), they realized (i.e., deductive reasoning) that Nina started off with $400 in her bank account. After the students had this realization, they set up the correct equation (i.e., mathematical reasoning) and plugged in the correct numbers to complete the item successfully. For instance, while reading and thinking about the problem silently, B108 said aloud, rhetorically, "Did she already have money in the bank?" and appeared momentarily confused. He then went on to solve the problem correctly, and during the retrospective report, he remarked, "I just kind of did the math and figured out she already has 400 dollars in the bank." Similarly, during the retrospective report, B106 remarked, "I was a little confused at first, because when you solve it, it equals 1200 not 1600, so she wouldn't have enough. Then it says that 'Nina has some money saved for a vacation.' Has SOME money? She only put in 1200, so she has to have 400 already."

Math Item 29

Emma is standing 10 feet away from the base of a tree and tries to measure the angle of elevation to the top. She is unable to get an accurate measurement, but determines that the angle of elevation is between 55 degrees and 75 degrees.

Decide whether each value given in the table is a reasonable estimate for the tree height. Select Reasonable or Not Reasonable for each height.

	Reasonable	Not Reasonable
4.2 feet	☐	☐
14.7 feet	☐	☐
24.4 feet	☐	☐
33.9 feet	☐	☐
39.1 feet	☐	☐
58.7 feet	☐	☐

Item 29 of the math section was a nontraditional, multiselect item. Students had to determine if six answers were reasonable or not reasonable for the height of a tree given data provided for the angle and length. All students (100.0%) reported they understood this item. The mean perceived difficulty for this item was 2.3 (*N* = 17), indicating medium difficulty.

Fifteen students in our sample attempted this problem. Of those, two students (13.3%) solved this problem correctly and 13 students solved this problem incorrectly (86.7%). All students who attempted this problem were able to solve at least a portion (correctly categorize some of the answers). Of those with partially correct answers, 10 of 12 correctly categorized at least four numbers as "reasonable" or "not reasonable" (83.3%). One of the two students who

completed this item correctly guessed the answer, but did not use mathematical reasoning. The student imagined trees she had seen before and said, “So when it’s asking for a reasonable estimate for the tree height I would probably just go with numbers in the middle because that makes sense” (A103).

Item 29 contained the highest number of features of the items selected. A majority of the students who did not solve this item correctly did not engage in mathematical reasoning (i.e., did not select a particular mathematical method to solve the problem), but instead would estimate heights for the tree based on reason, logic, and some mathematical knowledge (i.e., estimating the quantifiable characteristics of products, events, or information). Some of these students selected some correct “reasonable” heights for the tree based on this process (i.e., received a partially correct score based on our partial scoring criteria). These students would also create a partial visualization of the triangle described in the problem, but were unable to create a complete visualization.

The key features that appeared to be present in a student’s correct solution process were making decisions and solving problems, mathematical reasoning, visualization, flexibility of closure, and reading comprehension. To solve the problem correctly, students needed to first read the item prompt and associated narrative, and to clearly understand both aspects, which indicates the importance of reading and written comprehension for this particular item. Students then needed to analyze and evaluate the information and data provided in the narrative to determine the best approach to solving the item (i.e., making decisions and solving problems). As part of the decision-making process, students also needed to detect the particular triangle delineated in the narrative (i.e., flexibility of closure) and subsequently create a visualization of said triangle. After having represented the triangle and determined that the item required using mathematical procedures, students needed to utilize basic trigonometry (i.e., mathematical reasoning) and calculate either sine, cosine, or tangent (SOHCAHTOA).

Only three students, all from School B (B101, B103, B106), correctly reasoned that this problem requires basic trigonometry and the computing of sine, cosine, or tangent (mathematical reasoning), but ultimately none of the three were able to solve the problem correctly, perhaps due to incomplete knowledge or computational errors. B104, who did not mention SOHCAHTOA but got the item correct by guessing, explained, “I would probably just go with numbers in the middle because that makes sense,” noting that she had made a mental image of the tree in her head.

ELA Item 28

A student is writing a research report about the Iditarod Trail and annual dogsled race in northern Alaska. Read the paragraphs from her report and the directions that follow.

Alaska's Iditarod National Historic Trail has a long and rich history of travel and trade routes, including during the Gold Rush of the nineteenth century. Located close to the Arctic Circle, the one-thousand-mile trail system is known for its rugged terrain and harsh winters. But it is primarily known for the internationally famous Iditarod Trail Sled Dog Race, commonly referred to as simply the Iditarod. The Iditarod was founded in the 1970s by a few Alaskans determined to commemorate the state's rich history of dogsledding and to honor a famous event: the halt of a serious diphtheria epidemic in 1925, when a crucial medical serum was delivered by dogsled to the Alaskan citizens in need.

The Iditarod is an epic challenge for man and man's best friend: it's a grueling 1150-mile trek from Anchorage to Nome, Alaska. Although held in early spring, the conditions are extremely harsh. Each team consists of a human captain (also called a musher) leading a team of more than a dozen sled dogs—mostly Siberian Huskies, Alaskan Malamutes, or Eskimo Dogs. Together, they brave frigid conditions, including frozen tundra and icy forests. "The Last Great Race on Earth," as it is also called, starts during the first week in March with teams generally reaching Nome in 10-15 days.

Select **two** credible sources that would **most likely** give the student more information for her paragraphs.

- ☐ "Snow Huskies"
  Commercial Film
  This 1998 PG-13 film is a heartwarming tale about a team of ragtag Alaskan Malamutes who save their owner's ranch by winning the nation's premier dogsled race.
- ☐ www.northernparks.com
  This website has information about all of Alaska's state parks for tourists planning a visit to the country's largest state.
- ☐ www.racetheiditarod.com
  Our business will provide you with all the information you will need to race in the Iditarod. Let our team of experts help you register your dogs, find local veterinarians, and secure housing and supplies.
- ☐ *Balto: Our Hero* by C.R. Benson
  This book, first published in 1957 by children's literature author C.R. Benson, tells the true dramatic tale of Balto, the lead sled dog in a race against time to deliver serum during the 1925 diphtheria epidemic in Nome, Alaska.
- ☐ "The Iditarod Trail Dog Sled Race"
  United States Geographical Society
  www.usgeog.com/iditarod
  In celebration of its 100-year history of documenting our planet's uniqueness, our staff ventured to the 49th state to cover the Iditarod. This article chronicles the history of one of the most grueling competitions in the world, nicknamed "The Last Great Race."
- ☐ "MUSH! Alaska's Most Challenging Race"
  Documentary Film
  This 2007 six-hour miniseries documentary by award-winning filmmaker, Bernard Kems, details the history and route of the famous Iditarod.

Item 28 of the ELA section was a multiselect item. Students read a passage about an Alaskan dog race and were asked to select two credible sources that would most likely provide more supporting information for the passage, from six choices.

All students in our sample attempted this problem (*N* = 17). Of those, five students (29.4%) solved this problem correctly and 12 students solved this problem incorrectly (70.6%). The two correct answers were "The Iditarod Trail Dog Sled Race" and "Mush! Alaska's Most Challenging Race." Eleven of the 12 students who completed the item incorrectly were partially correct based on our partial scoring criteria. Most participants missed this item because they chose "www.racetheiditarod.com" instead of "Mush! Alaska's Most Challenging Race." Most participants reported that they understood the item (94.1%). The mean perceived difficulty for this item was 1.2 (*n* = 16), indicating that students tended to perceive this item to as easy.

With respect to particular career features, students who did not solve the item correctly did not appear to engage in thorough reading or written comprehension or fully understand what they read. Many of the students selected sources based on irrelevant or incorrect details from the narrative, or did not seem to fully understand what was meant by "credible sources" within this context. These students seemed to select sources somewhat haphazardly without analyzing the facts and details of the text (i.e., analyzing data or information), or without

identifying key information and comparing and contrasting salient details (i.e., identifying objects, actions, and events).

The students who solved the item correctly, on the other hand, appeared to fully comprehend what the question was asking (i.e., written comprehension) and what the paragraphs were conveying (i.e., reading comprehension). These students proceeded to analyze the facts and key details provided within the text and their similarities and differences (i.e., analyzing data or information, and identifying objects, actions, and events), and use these key facts and details to compare, categorize, and verify with the details provided about the source options. (i.e., making decisions and solving problems and processing information). As part of this process, students also needed to use logic and reasoning to identify the strengths and weaknesses of each source based on the key details provided about the source, and how they compared to the salient details and facts within the paragraphs (i.e., critical thinking, and judging the quality of things, services, or people), as well as obtain information from all the sources (i.e., getting information).

For instance, A105 noted generally that sources "about the race" are better than sources "not about the race." Students eliminated options through this process; for instance, C101 dismissed "www.northernparks.com" because "it was like touring; it wasn't really about the race" and B108 remarked "Snow Huskies is a film so it's like a Disney/Hollywood thing, which could be cool but no information, really." Meanwhile, students with potentially lower reading comprehension skills or deductive reasoning skills made answer choices by honing in on the word "Iditarod," such as B201, who chose "www.racetheiditarod.com" and "The Iditarod Trail Dog Sled Race" because "these websites used the name of the trail in it so figured it [sic] would be useful."

ELA Item 30

A student is writing an argumentative report about the causes of sleepwalking. She found possible sources for her report. Read the sources and the directions that follow.

Source 1: "What is Sleepwalking?" by Mat Valerio

Sleepwalking is a sleep disorder that causes a person to get up and walk while still asleep, usually during the deepest stages of sleep. While sleepwalking, a person does not often respond when someone asks a question or touches him or her. Though sometimes a sleepwalker will verbally respond, the response will make no sense. A sleepwalking episode can include the person just walking quietly around a room or demonstrating very agitated behavior and trying to "escape" the room.

Source 2: "Is Sleepwalking Genetic?" by Chris Shue

Sleepwalking occurs most often during middle childhood through adolescence. Some adults also sleepwalk, but it is much more common among children. Genetics may play a big role in determining who will be a sleepwalker as the tendency runs in families. Environmental and medical conditions also may contribute to sleepwalking episodes. Sleepwalking was described in some of the earliest recorded medical literature, even before Hippocrates, the "father of medicine," lived.

Source 3: "Am I a Sleepwalker?" by Angelique Kandar

Research has shown that a variety of factors contribute to sleepwalking episodes. Sleepwalkers are usually operating under a severe lack of sleep or have an irregular or hectic sleep schedule. Often, they are under great amounts of stress or anxiety. Some medical conditions, such as abnormal heart rhythms, nighttime seizures, and **sleep apnea**, have been known to cause sleepwalking. Sometimes a person who has taken a certain medication experiences a sleepwalking episode.

The student wrote down some claims to use in her report. Look at the claims on the table. Determine if the information in the sources supports each claim. Click on the boxes to show the claims that each source supports. A source may have more than one box selected.

	Source 1	Source 2	Source 3
Claim 1: Some outside influences make a person more likely to sleepwalk.	☐	☐	☐
Claim 2: If your mother was a sleepwalker, it is more likely that you will be, too.	☐	☐	☐
Claim 3: When people are sleepwalking they are not aware of their surroundings.	☐	☐	☐
Claim 4: Going to bed at the same time every night can help some children not to sleepwalk.	☐	☐	☐

Item 30 of the ELA section was a nontraditional, multiselect item. Students had to read three passages (sources) about sleepwalking and determine which sources supported each claim.

Fourteen students in our sample attempted this problem. Of those, two students (14.3%) solved this problem correctly and 12 students solved this problem incorrectly (85.7%). All students who attempted this problem were able to solve at least a portion (correctly categorize some of the answers). Of those with partially correct answers, most participants missed Claim 1 which had two correct answers. This could partly explain the low frequency of correct responses. All students who attempted it reported they understood the item (100.0%). The mean perceived difficulty for this item was 1.1 (n = 14), indicating that students tended to perceive this item to as easy.

The key processes demonstrated to solve Item 30 correctly included analyzing data or information, making decisions and solving problems, processing information, getting information, and written and reading comprehension. Similar to Item 28, students needed to comprehend what the question was asking (i.e., written comprehension), what the paragraphs were conveying (i.e., reading comprehension), and in particular key facts and data provided in the sources as well as the claims (i.e., analyzing data or information). It became apparent that a majority of the students missed the key detail within the item instructions that more than one

source could be selected for a claim. Limited written comprehension was perhaps a contributing error in this respect.

Moreover, students who completed the item correctly appeared to engage in a similar process as in Item 28 where they obtained relevant information from the sources (i.e., getting information), and utilized these key facts and details from the sources to compare, categorize, and verify with the details provided within the claims (i.e., making decisions and solving problems and processing information). Students also needed to recognize differences and similarities across the three sources (i.e., identifying objects, actions, events). As C101 noted, "They were all like one topic, but different questions." Ultimately, reading comprehension plays a large role in this English language arts item. A106, for example, demonstrated lower reading comprehension skills throughout the cognitive lab; for this particular item she homed in on the word "children" in Claim 4 and incorrectly chose Source 2, which also contains the word "children," suggesting that she relied on shallow reading skills rather than the more complex reading comprehension needed to solve this problem correctly.

UCLA

CRESST

NATIONAL CENTER FOR RESEARCH ON EVALUATION, STANDARDS, AND STUDENT TESTING

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)

Graduate School of Education & Information Studies
University of California, Los Angeles
300 Charles E. Young Drive North
GSE&IS Bldg., Box 951522
Los Angeles, CA 90095-1522

(310) 206-1532
www.cresst.org