



**2019 NSCAS Science Pilot Technical Report**  
May 2019



## Table of Contents

|   |     |
|---|-----|
| Section 1: Introduction .....                                       | 1   |
| 1.1. Pilot Overview .....   | 1   |
| 1.2. Background and Rationale .....                                 | 1   |
| 1.3. Development Timeline .....                                     | 2   |
| Section 2: Test Design and Development .....                        | 3   |
| 2.1. Pilot Test Design .....  | 3   |
| 2.2. Task Development .....   | 3   |
| 2.3. Prompt Types .....   | 4   |
| Section 3: Pilot Administration .....                               | 5   |
| 3.1. Cognitive Labs .....   | 5   |
| 3.2. Administration Training .....                                  | 5   |
| 3.3. Practice Tasks .....   | 5   |
| 3.4. Accommodations and Accessibility Features .....                | 6   |
| 3.5. Test Security .....  | 6   |
| Section 4: Test Results and Analyses .....                          | 7   |
| 4.1. Scoring and Reporting .....                                    | 7   |
| 4.2. Student Sample .....   | 8   |
| 4.3. Test Duration .....  | 9   |
| 4.4. Overall Test Scores .....                                      | 10  |
| 4.5. Classical Analysis .....                                       | 10  |
| 4.6. Distractor Analysis .....                                      | 13  |
| 4.7. Student Feedback Results .....                                 | 16  |
| 4.7.1. Feedback Question #1: Task Interestingness .....             | 18  |
| 4.7.2. Feedback Questions #2 and #3: Task and Item Difficulty ..... | 19  |
| 4.8. Cognitive Lab Results .....                                    | 21  |
| 4.9. Considerations for Future Development .....                    | 23  |
| Section 5: Fairness, Reliability, and Validity .....                | 24  |
| 5.1. Differential Item Functioning (DIF) .....                      | 24  |
| 5.1.1. DIF Methods .....  | 24  |
| 5.1.2. DIF Results .....  | 25  |
| 5.2. Reliability .....  | 26  |
| 5.3. Concurrent Validity .....                                      | 27  |
| References .....  | 28  |
| Appendix A: Task Development Evaluation Results .....               | 29  |
| Appendix B: Task Review Evaluation Results .....                    | 31  |
| Appendix C: Pilot Test Administration Manual .....                  | 33  |
| Appendix D: Cognitive Lab Parent Notification Letter .....          | 45  |
| Appendix E: Cognitive Lab Protocol .....                            | 50  |
| Appendix F: Administration Training Presentation .....              | 65  |
| Appendix G: Cognitive Lab Training Presentation .....               | 70  |
| Appendix H: Cognitive Lab Tasks and Items .....                     | 77  |
| Appendix I: Cognitive Lab Responses .....                           | 102 |
| Appendix J: Cognitive Lab Coding Results .....                      | 151 |
| Appendix K: Prompt-Level DIF Results .....                          | 154 |

## List of Tables

|   |    |
|---|----|
| Table 1.1. Development Timeline.....  | 2  |
| Table 2.1. Pilot Test Design .....  | 3  |
| Table 2.2. Prompt Types Included in the Pilot .....   | 4  |
| Table 4.1. Scoring Rules by Prompt Type .....   | 7  |
| Table 4.2. Score Range by Test Form .....   | 8  |
| Table 4.3. Number of Students in the Pilot Sample .....   | 8  |
| Table 4.4. Pilot Sample Demographics .....  | 8  |
| Table 4.5. Test Duration Descriptive Statistics .....   | 9  |
| Table 4.6. Raw Score Descriptive Statistics .....   | 10 |
| Table 4.7. Prompt-Level <i>P</i> -Values and Point-Biserial Correlations—Grade 5.....   | 11 |
| Table 4.8. Prompt-Level <i>P</i> -Values and Point-Biserial Correlations—Grade 8.....   | 12 |
| Table 4.9. Percentage of Students Selecting Each Answer Option—Grade 5, Form A.....   | 13 |
| Table 4.10. Percentage of Students Selecting Each Answer Option—Grade 5, Form B.....  | 14 |
| Table 4.11. Percentage of Students Selecting Each Answer Option—Grade 8, Form A.....  | 14 |
| Table 4.12. Percentage of Students Selecting Each Answer Option—Grade 8, Form B.....  | 14 |
| Table 4.13. Point-Biserial Correlations for Each Answer Option—Grade 5, Form A .....  | 15 |
| Table 4.14. Point-Biserial Correlations for Each Answer Option—Grade 5, Form B .....  | 15 |
| Table 4.15. Point-Biserial Correlations for Each Answer Option—Grade 8, Form A .....  | 16 |
| Table 4.16. Point-Biserial Correlations for Each Answer Option—Grade 8, Form B .....  | 16 |
| Table 4.17. Rating Percentage Distribution—Student Feedback Question #1 .....   | 18 |
| Table 4.18. Rating Percentage Distribution—Student Feedback Question #2 .....   | 19 |
| Table 4.19. Rating Percentage Distribution—Student Feedback Question #3 .....   | 20 |
| Table 4.20. Cognitive Lab Questions .....   | 21 |
| Table 5.1. Focal and Reference Groups for the DIF Demographic Variables .....   | 24 |
| Table 5.2. DIF Categories .....   | 25 |
| Table 5.3. DIF Summary Results .....  | 26 |
| Table 5.4. Reliability Results: Cronbach’s Alpha and Predicted Reliability ( <i>r</i> ) based on the<br>Spearman-Brown Formula..... | 27 |
| Table 5.5. Concurrent Validity between Pilot Test Scores and 2019 NSCAS Science Scores ...  | 27 |

## List of Figures

|  |    |
|--|----|
| Figure 2.1. Sample Task Model .....                        | 4  |
| Figure 4.1. Student Feedback Questions .....               | 17 |
| Figure 4.2. Ratings for Student Feedback Question #1 ..... | 18 |
| Figure 4.3. Ratings for Student Feedback Question #2 ..... | 19 |

## List of Abbreviations

Below is a list of abbreviations that appear in this technical report.

|              |   |
|--------------|---|
| CCC .....    | Crosscutting Concept                                    |
| CM .....     | choice multiple   |
| D&D .....    | drag-and-drop   |
| DCI.....     | Disciplinary Core Idea                                  |
| DIF .....    | differential item functioning                           |
| ELL.....     | English language learner                                |
| HT .....     | hot text  |
| MC .....     | multiple-choice   |
| NCCRS-S..... | Nebraska College and Career Ready Standards for Science |
| NDE .....    | Nebraska Department of Education                        |
| NSCAS.....   | Nebraska Student-Centered Assessment System             |
| SD .....     | standard deviation                                      |
| SEP.....     | Science and Engineering Practice                        |
| SES.....     | social economic status                                  |
| SSID.....    | state student ID  |
| TE .....     | text entry  |

## Section 1: Introduction

The purpose of this technical report is to summarize the development, administration, and results of the Spring 2019 pilot of the newly developed performance tasks for the Nebraska Student-Centered Assessment System (NSCAS) Summative Science assessments in Grades 5 and 8. The pilot assessment was designed by the Nebraska Department of Education (NDE) with support from its vendor NWEA® to measure three-dimensional science learning, incorporating elements of Disciplinary Core Ideas (DCIs), Science and Engineering Practices (SEPs), and Crosscutting Concepts (CCCs) from the Nebraska College and Career Ready Standards for Science (NCCRS-S; NDE, 2017). The purpose of the pilot was to gain feedback from Nebraska students on the newly developed tasks for use on the new science assessment that will be aligned to the NCCRS-S.

### 1.1. Pilot Overview

The Spring 2019 pilot tests were comprised of tasks and associated prompts written and reviewed by Nebraska educators in collaboration with NDE. The pilot was administered from March 4–15, 2019, in Grades 5 and 8 to glean meaningful information directly from students to inform NDE and NWEA on test development for Summer 2019. It was also intended to expose the new science tasks and associated prompts to Nebraska districts to prepare them for future administrations and to give them the opportunity to provide feedback on the prompt types. A small number of districts also participated in cognitive labs to document individual students' feedback about the tasks and prompts. The pilot was not intended to test students' knowledge or skills, and students' scores were not reported, although students' raw responses were provided to the NDE. Student responses will be analyzed by NDE and NWEA psychometricians and content specialists, and their feedback will be applied toward the development of the new NSCAS Science assessments.

### 1.2. Background and Rationale

The new NSCAS Science assessments in Grades 5 and 8 are intended to encompass new content standards and technologies. The development of these assessments began with the adoption of the NCCRS-S by the Nebraska State Board of Education on September 8, 2017. Informed by growing research in cognitive science and developmental psychology, these standards were guided by *A Framework for K–12 Science Education* (National Research Council, 2012) and are designed to reflect more recent research and thinking in science education. More specifically, they are more rigorous to better prepare students with the science knowledge, skills, and habits of mind to be ready for college, career, and civic responsibilities. This includes use of phenomena and three-dimensional thinking.

While Nebraska's current science assessment only uses standalone multiple-choice items, the new assessment design is based around performance tasks, which are phenomena-based scenarios with multiple prompts and commands leading the student into more complex thinking and a focus on doing science rather than knowing discrete science facts. The new assessment will use a variety of technology-enhanced prompt types to allow students to more fully show their thinking. The large change from the current assessment to the new NSCAS Science assessment resulted in this pilot project to allow NDE and NWEA to better design appropriate tasks and prompts.

### 1.3. Development Timeline

Table 1.1 presents the development timeline for the new NSCAS Science assessment. Following the Spring 2019 pilot, the next step is a full-scale field test in Spring 2020 to provide more accurate psychometric information to adjust the design and item bank as needed before the operational launch in Spring 2021. The tasks that survive the Spring 2020 field test will become operational in Spring 2021, so the data obtained from the Spring 2019 pilot and cognitive labs is important to inform how those tasks are developed in Summer 2019.

**Table 1.1. Development Timeline**

| Event                                  | Date(s)               |
|--|-----------------------|
| Adoption of NCCRS-S                    | September 8, 2017     |
| Visioning meeting led by Achieve       | November 2017         |
| Task force                             | January 2018          |
| Science Assessment Conference          | April 10–11, 2018     |
| Test Development Workshop              | July 17–20, 2018      |
| Task Review Committee                  | September 11–12, 2018 |
| Administration training                | February 25, 2019     |
| Cognitive lab training                 | February 25, 2019     |
| Pilot test window                      | March 4–15, 2019      |
| Cognitive labs                         | March 4–15, 2019      |
| Science ALD Workshop                   | May 1–2, 2019         |
| Science Phenomenon Development Meeting | June 17–21, 2019      |
| Science Assessment Development Meeting | July 8–12, 2019       |
| Full-scale field test                  | Spring 2020           |
| Operational launch                     | Spring 2021           |

## Section 2: Test Design and Development

### 2.1. Pilot Test Design

Each grade had two test forms, and each form consisted of two tasks that had between five and eight prompts, as shown in Table 2.1. Some prompts were multi-part (i.e., required more than one response). The new NSCAS Science assessment was designed around three-dimensional science learning, which includes tasks that require students to understand and use the three elements of DCIs, SEPs, and CCCs. The tasks and their associated prompts were linked to a common phenomenon that guided students through a complex cognitive process to engage with and use science. The prompts were questions or exercises to elicit responses that show the students' understanding of the DCIs, SEPs, and CCCs. The pilot assessment was intentionally short to reduce the time students spent away from the classroom and to limit the pilot development effort to ensure the quality and effectiveness of the tasks and prompts presented to the pilot sample.

**Table 2.1. Pilot Test Design**

| Grade | Form   | Task                 | #Prompts |
|-------|--------|----------------------|----------|
| 5     | Form A | T1: Lake Bubbles     | 5        |
|       |        | T2: Plant Problems   | 5        |
|       | Form B | T1: Something Smells | 5        |
|       |        | T2: TV Trouble       | 8        |
| 8     | Form A | T1: Fossil Discovery | 6        |
|       |        | T2: Sculpture Setup  | 6        |
|       | Form B | T1: Pirate Bay       | 6        |
|       |        | T2: Sculpture Setup  | 6        |

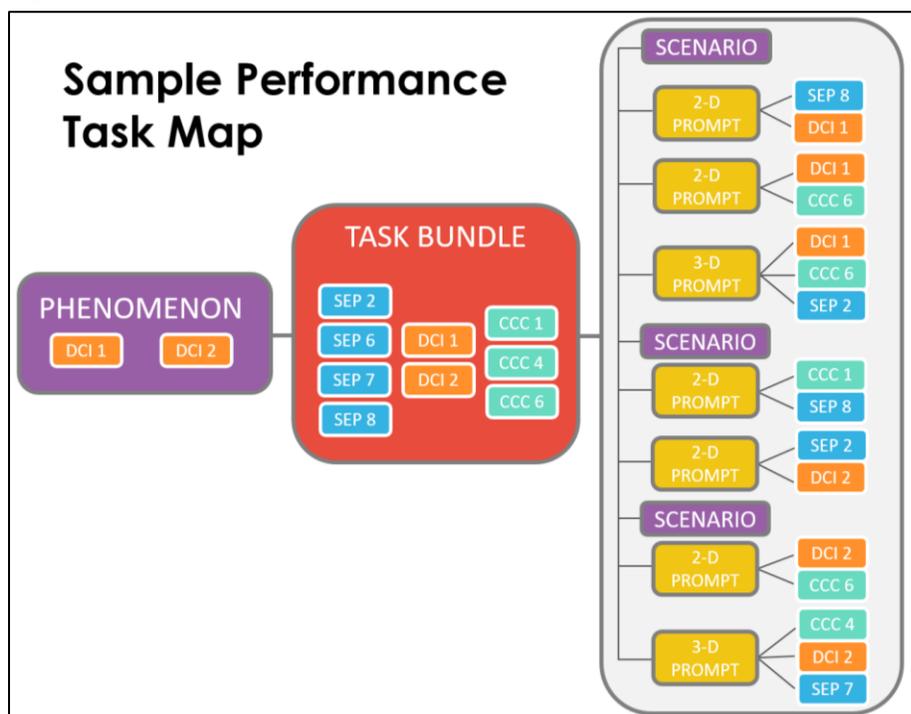
### 2.2. Task Development

A group of Nebraska teachers were recruited by NDE and brought together from July 17–20, 2018, to develop tasks and prompts for the 2019 pilot test. A total of 24 teachers participated. The writers developed five tasks per grade (10 tasks total) with between five and eight prompts per task. As shown in Table 2.1, eight of those tasks were included on the test forms (two tasks per form, four per grade). The other two tasks were provided as practice tasks per the NDE prior to the pilot administration and were therefore not included on the pilot forms.

The writers were guided in the vision of the new NSCAS Science assessment and began the development process by identifying a phenomenon that met NDE's criteria (e.g., it is observable, accessible, engaging, and explainable using grade-level appropriate science core ideas). Writers then thought about the steps needed for students to make sense of the phenomenon and identified SEPs and CCCs students would use in the sense-making process. At the end of the meeting, participants completed an evaluation survey. Appendix A presents the task development evaluation questions and results.

A task was built by introducing the phenomenon in a scenario that was bimodal (e.g., it had text and graphics) followed by prompts that were minimally two-dimensional. When additional information was needed, it was presented with another mini-scenario. Each task had at least one three-dimensional prompt. Some tasks culminated in an open-ended, evidence-based writing prompt. For the pilot, each test form contained one of these open-ended prompts. Figure 2.1 describes a sample task model and layout.

Figure 2.1. Sample Task Model



The newly developed tasks and prompts were further refined by a task review committee that met from September 11–12, 2018. The committee consisted of NDE staff, NWEA staff, and a group of 15 educators recruited by NDE who were not involved in writing the tasks. The tasks and prompts were reviewed for content and bias concerns. NDE determined which tasks would be placed on each form. NWEA then built the forms in the Qualtrics platform for the pilot. At the end of the review, participants completed an evaluation survey. Appendix B presents the task review evaluation questions and results.

### 2.3. Prompt Types

Table 2.2 describes the prompt types included on the Spring 2019 pilot assessments.

Table 2.2. Prompt Types Included in the Pilot

| Prompt Type                   | Description  |
|-------------------------------|--|
| Multiple-Choice               | Students select one response from multiple options.  |
| Choice Multiple (Multiselect) | Students select two or more responses from multiple options.   |
| Text Entry                    | Students respond to either fill-in-the-blank type items or open-ended prompts. Only 1 open-ended prompt was included in each form. |
| Drag-and-Drop                 | Students drag and drop items into groups. Within each group, students can rank items by dragging and dropping them into place.     |
| Hot Text                      | Students are presented with an image or text that has predefined regions to choose from.   |

## **Section 3: Pilot Administration**

The 2019 NSCAS Science pilot administration took place from March 4–15, 2019. Appendix C presents the pilot administration manual that includes all procedures and a proctor script. The tests were administered through Qualtrics, an online survey tool. The Qualtrics platform allowed the pilot to be administered to a larger number of students compared to a paper-pencil test, thereby improving the quality of the data. Participation was optional and at the discretion of each district. Each test form consisted of two tasks with associated prompts, along with a set of student feedback questions after each task. Students could rate their interest and the difficulty of the task and how well they understood each prompt. The pilot assessment was untimed and designed to provide students with as much time as needed to complete, although it was estimated that students would take no more than 45 minutes to complete each form. The test was to be taken in one sitting. Approximately 10,000 students participated in the pilot.

### **3.1. Cognitive Labs**

As part of the pilot administration, a small number of districts selected by the NDE participated in cognitive labs to collect more detailed information about the science tasks. The cognitive lab forms were also administered online via Qualtrics. Grade 5 students who participated in the cognitive lab received Form B, and Grade 8 students took Form A. These separate cognitive lab forms had specific questions to allow proctors to record deeper insight into student responses in a one-on-one setting. Appendix D presents the notification letter sent to parents and guardians that explained the purpose of the cognitive labs and asked for their consent to have their student participate. Appendix E presents the cognitive lab protocol that guided educators through the process.

Nine students participated in the Grade 5 cognitive lab, and six participated in Grade 8. An educator interacted with each student as they worked through the two science tasks and associated prompts. Students said what they were thinking out loud when reading the tasks and responding to the prompts and gave their perspective, including how easy or difficult the tasks and prompts were and why. Proctors asked the students to think aloud as they responded, guided students through the process, and documented their responses. The cognitive lab took no longer than 1.5 hours for most students.

### **3.2. Administration Training**

Two virtual trainings were conducted on February 25, 2019. One was meant for all participating districts, and a second training was specifically designed for districts participating in the cognitive lab. NDE sent an email to all districts with the link to join the virtual training. NDE then sent a separate email to those selected for the cognitive lab to inform them of the training and the link to join the training session. Appendix F presents the pilot administration training presentation, and Appendix G presents the cognitive lab training presentation. Meeting materials included a copy of the presentation and the test administration manual, as well as a recording of the training provided as a follow-up material.

### **3.3. Practice Tasks**

Two of the originally developed tasks were available for students and districts to review online prior to administering the pilot assessments. It was recommended that all students complete the appropriate practice task prior to participating in the pilot.

- Grade 5 Changing Coral: <http://bit.ly/prepilot5>
- Grade 8 Blue Skin: <http://bit.ly/prepilot8>

### **3.4. Accommodations and Accessibility Features**

The Qualtrics platform did not have any standard embedded accommodations or accessibility supports. The following guidelines were provided to educators during the administration training:

- Students may use approved non-embedded resources, such as multiplication charts or noise buffers, as specified by NDE policy.
- Zoom/magnifier – students should use the native device zoom feature to magnify the content on the page.
- Text-to-speech – students should use native screen readers or read aloud in accordance with the NDE accessibility manual.

A complete list of non-embedded universal tools, linguistic supports, and accommodations is provided in the NSCAS Summative & Alternate Accessibility Manual<sup>1</sup>.

### **3.5. Test Security**

The pilot administration manual was provided as a resource and included details on security. In a centralized testing process, it is critical that equity of opportunity, standardization of procedures, and fairness to students is maintained. Although the pilot assessment was not considered secure and tasks were released to the public, it was recommended that educators follow standard practices in test security as they would during a standardized assessment. All teachers need to be familiar with appropriate testing ethics and security practices related to testing. Professionalism, common sense, and practical procedures will provide the right framework for testing ethics. The NSCAS security manual<sup>2</sup> is intended to outline clear practices for appropriate security.

---

<sup>1</sup> The NSCAS Summative & Alternate Accessibility Manual is located online at <https://www.education.ne.gov/assessment/nscas-system/#nscas-accessibility>.

<sup>2</sup> The NSCAS Summative Assessment Security Manual is located online at <https://www.education.ne.gov/assessment/nscas-system/#nscas-security>.

## Section 4: Test Results and Analyses

### 4.1. Scoring and Reporting

All prompts were dichotomously scored, although the pilot scores were not reported. Student responses to the closed-ended prompts were scored by programs specifically developed for the pilot test, and student responses to the open-ended prompts were not included in the scoring (e.g., Question 5 on the Grade 5 Form A). Table 4.1 presents the scoring rules for each prompt type. Each prompt was worth 1 point.

**Table 4.1. Scoring Rules by Prompt Type**

| Prompt Type                   | Scoring Rules   |
|-------------------------------|---|
| Multiple-Choice               | <ul style="list-style-type: none"> <li>• 1 point if the response matches the key</li> <li>• 0 points any other response</li> </ul>  |
| Choice Multiple (Multiselect) | <ul style="list-style-type: none"> <li>• 1 point if the response matches the key exactly</li> <li>• 0 points any other response (e.g., if the student misses any correct option or adds any incorrect option)</li> </ul>  |
| Text Entry                    | Score just the numeric values: <ul style="list-style-type: none"> <li>• 1 point if the numeric value matches the key</li> <li>• 0 points any other response</li> </ul>  |
| Drag-and-Drop                 | Each D&D prompt is worth 1 point even though there are more than 1 boxes: <ul style="list-style-type: none"> <li>• 1 point if the item that the student dragged into each box is correct and no additional incorrect item(s) is dragged into each box</li> <li>• 0 points any other response</li> </ul> |
| Hot Text                      | <ul style="list-style-type: none"> <li>• 1 point if the response matches the key exactly</li> <li>• 0 points any other response (e.g., if the student did not highlight any correct sentence or highlighted any incorrect sentence)</li> </ul>  |

For choice multiple prompts, students had to select all the correct choices and could not select any incorrect choices to get 1 score point. For the drag-and-drop prompts, students needed to drag and drop the correct item into each box and could not drop any incorrect item into a box to get 1 score point. If an item had multiple parts, each part was worth 1 point and treated as a separate prompt for scoring purposes. The following prompts had multiple parts:<sup>3</sup>

- Grade 5 Form A: T2\_Q1, T2\_Q2
- Grade 5 Form B: T1\_Q2
- Grade 8 Form A: T1\_Q1, T1\_Q4, T1\_Q6, T2\_Q2, T2\_Q4, T2\_Q5
- Grade 8 Form B: T1\_Q5, T1\_Q6, T2\_Q2, T2\_Q4, T2\_Q5

Table 4.2 presents the possible score ranges by grade and form. The number of points possible depended on the number of prompts (or prompt parts) on each form. Since each prompt was worth only 1 point, the total number of points possible depended on how many close-ended prompts were on a test form. For example, Form A in Grade 5 had 11 close-ended prompts total across both tasks (including one prompt with two parts), so the maximum number of points possible for that form is 11. The possible score range varies by test form because the number of close-ended prompts on each form differed.

<sup>3</sup> Prompt codes are organized as follows: T1 = Task 1, T2 = Task 2, Q1 = Question 1, etc., which indicate the prompt number in the order they were presented in the task.

**Table 4.2. Score Range by Test Form**

| Grade | Form   | Score Range |
|-------|--------|-------------|
| 5     | Form A | 0–11        |
|       | Form B | 0–14        |
| 8     | Form A | 0–18        |
|       | Form B | 0–17        |

**4.2. Student Sample**

Table 4.3 presents the total number of students who took either the Grade 5 or Grade 8 form during the pilot. Table 4.4 presents the sample sizes and percentage of students who participated by various demographic variables including grade, gender, ethnicity, economic disadvantage status, disability status, English language learner (ELL) status, instructional program type, and homelessness status. All students who completed the test within the testing window were included. Around 1% of students did not finish the test and were excluded from the analyses. Overall, the student sample had similar demographic characteristics to Nebraska’s general population (NDE, 2018, pp. 103–107).

Student demographic information was retrieved by matching the state student IDs (SSIDs) to the information stored in the NWEA database. A small proportion of students who entered the wrong SSIDs were excluded from the demographic information in Table 4.4 because no demographic information could be retrieved for those students. However, their responses were included in all other psychometric analyses for the pilot.

**Table 4.3. Number of Students in the Pilot Sample**

| Grade | Pilot Sample |       |
|-------|--------------|-------|
|       | N            | %     |
| 5     | 4,852        | 48.4  |
| 8     | 5,172        | 51.6  |
| Total | 10,024       | 100.0 |

**Table 4.4. Pilot Sample Demographics**

| Demographic Variable |  | Grade 5 Form |       | Grade 8 Form |        | Total  |       |
|----------------------|--|--------------|-------|--------------|--------|--------|-------|
|                      |  | N            | %     | N            | %      | N      | %     |
| Gender               | Female                                 | 2,351        | 48.5  | 2,531        | 48.9   | 4,882  | 48.7  |
|                      | Male                                   | 2,501        | 51.5  | 2,641        | 51.1   | 5,142  | 51.3  |
|                      | Total                                  | 4,852        | 100.0 | 5,172        | 100.0  | 10,024 | 100.0 |
| Ethnicity            | American Indian or Alaskan Native      | 68           | 1.4   | 82           | 1.6    | 150    | 1.5   |
|                      | Asian                                  | 144          | 3.0   | 111          | 2.1    | 255    | 2.5   |
|                      | Black or African American              | 181          | 3.7   | 193          | 3.7    | 374    | 3.7   |
|                      | Hispanic                               | 899          | 18.5  | 941          | 18.2   | 1,840  | 18.4  |
|                      | Native Hawaiian/Other Pacific Islander | 11           | 0.2   | 11           | 0.2    | 22     | 0.2   |
|                      | White                                  | 3,380        | 69.7  | 3,674        | 71.0   | 7,054  | 70.4  |
|                      | Two or More Races                      | 169          | 3.5   | 160          | 3.1    | 329    | 3.3   |
| Total                | 4,852                                  | 100.0        | 5,172 | 100.0        | 10,024 | 100.0  |       |

| Demographic Variable       |                        | Grade 5 Form |       | Grade 8 Form |       | Total  |       |
|----------------------------|------------------------|--------------|-------|--------------|-------|--------|-------|
|                            |                        | N            | %     | N            | %     | N      | %     |
| Economic Disadvantage      | No                     | 2,827        | 58.3  | 3,318        | 64.2  | 6,145  | 61.3  |
|                            | Yes                    | 2,025        | 41.7  | 1,854        | 35.8  | 3,879  | 38.7  |
|                            | Total                  | 4,852        | 100.0 | 5,172        | 100.0 | 10,024 | 100.0 |
| Disability                 | No                     | 4,143        | 85.4  | 4,632        | 89.6  | 8,775  | 87.5  |
|                            | Yes                    | 709          | 14.6  | 540          | 10.4  | 1,249  | 12.5  |
|                            | Total                  | 4,852        | 100.0 | 5,172        | 100.0 | 10,024 | 100.0 |
| ELL                        | Current ELL            | 262          | 5.4   | 120          | 2.3   | 382    | 3.8   |
|                            | Former ELL             | 417          | 8.6   | 475          | 9.2   | 892    | 8.9   |
|                            | Non-ELL                | 4,173        | 86.0  | 4,577        | 88.5  | 8,750  | 87.3  |
|                            | Total                  | 4,852        | 100.0 | 5,172        | 100.0 | 10,024 | 100.0 |
| Instructional Program Type | ELL                    | 10           | 0.2   | 14           | 0.3   | 24     | 0.2   |
|                            | Free and Reduced Lunch | 44           | 0.9   | 104          | 2.0   | 148    | 1.5   |
|                            | N/A                    | 4,658        | 96.0  | 4,771        | 92.2  | 9,429  | 94.1  |
|                            | Other                  | 5            | 0.1   | 25           | 0.5   | 30     | 0.3   |
|                            | Section 504            | 1            | 0.0   | 1            | 0.0   | 2      | 0.0   |
|                            | Special Education      | 52           | 1.1   | 70           | 1.4   | 122    | 1.2   |
|                            | Talented and Gifted    | 72           | 1.5   | 187          | 3.6   | 259    | 2.6   |
|                            | Title 1                | 10           | 0.2   | –            | –     | 10     | 0.1   |
|                            | Total                  | 4,852        | 100.0 | 5,172        | 100.0 | 10,024 | 100.0 |
| Homeless                   | No                     | 4,829        | 99.5  | 5,153        | 99.6  | 9,982  | 99.6  |
|                            | Yes                    | 23           | 0.5   | 19           | 0.4   | 42     | 0.4   |
|                            | Total                  | 4,852        | 100.0 | 5,172        | 100.0 | 10,024 | 100.0 |

### 4.3. Test Duration

Table 4.5 presents the descriptive statistics of overall test duration in minutes, including the n-count, mean, standard deviation (SD), and the median, minimum, and maximum values. The pilot test was untimed, but it took an average of 25–30 minutes for students to finish the test. The timing data are skewed by outliers, or test durations larger than 3 SDs from the mean score (e.g., the maximum test duration of each form is very large). The outliers indicate that a very small proportion (0.1%) of the students completed the test at a later time, which resulted in these extreme large maximums. These outliers influenced the average time that students used to finish the test. The median, however, is less influenced by the outliers. It indicates that students finished the test around 21–25 minutes per form and is more similar across forms.

**Table 4.5. Test Duration Descriptive Statistics**

| Grade | Form   | N     | Test Duration in Minutes |       |        |      |        |
|-------|--------|-------|--------------------------|-------|--------|------|--------|
|       |        |       | Mean                     | SD    | Median | Min. | Max.*  |
| 5     | Form A | 2,795 | 30.7                     | 150.3 | 25.6   | 1.0  | 7934.5 |
|       | Form B | 2,521 | 24.8                     | 37.1  | 21.7   | 0.7  | 1265.6 |
| 8     | Form A | 3,090 | 31.2                     | 189.1 | 22.3   | 0.9  | 7173.0 |
|       | Form B | 2,779 | 28.0                     | 119.9 | 21.8   | 0.6  | 4430.4 |

\*Reasons for the large test durations could vary. For example, it is possible that students did not complete the test in one sitting.

#### 4.4. Overall Test Scores

Table 4.6 presents raw score descriptive statistics by test form. The average test scores are well below the maximum possible score points of each form, indicating that the test forms were hard for students. For example, the maximum possible score of Grade 5 Form A is 11. However, students' average score is only 3.2. The highest scores students received are lower than the maximum possible scores for three of the four forms, which suggests no one answered all the items correctly on these forms. The average score also varies by form, suggesting that the difficulty of the two forms at each grade level are not equivalent. Form A is more difficult than Form B for Grade 5 but easier than Form B for Grade 8.

**Table 4.6. Raw Score Descriptive Statistics**

| Grade | Form   | N     | Total Score |     |        |      |      | Max. Possible Score |
|-------|--------|-------|-------------|-----|--------|------|------|---------------------|
|       |        |       | Mean        | SD  | Median | Min. | Max. |                     |
| 5     | Form A | 2,783 | 3.2         | 1.6 | 3      | 0    | 10   | 11                  |
|       | Form B | 2,517 | 6.9         | 2.3 | 7      | 0    | 14   | 14                  |
| 8     | Form A | 3,081 | 6.4         | 3.2 | 6      | 0    | 16   | 18                  |
|       | Form B | 2,774 | 5.4         | 2.7 | 5      | 0    | 16   | 17                  |

#### 4.5. Classical Analysis

All prompts, or items, were analyzed for item difficulty and item discrimination based on classical test theory:

- Item difficulty is presented as the  $p$ -value that shows the proportion of students who answer an item correctly. The  $p$ -value is bound by 0.0 and 1.0 and is derived by dividing the number of students who got the item correct by the total number of students who answered it.
- Item discrimination refers to the ability of an item to differentiate students who understand the concept being measured from those who do not (i.e., low-ability vs. high-ability students). Students who do well on a test are expected to select the right answer to any given item, and students who do poorly are expected to select the wrong answer. This means that for a highly discriminating prompt, students who get the item correct will have a higher average test score than students who get the item incorrect.

Item discrimination is assessed by the correlation between how well students did on an item and how well they did on the entire test (i.e., their overall test score). This correlation is known as the point-biserial correlation that ranges between -1.0 and +1.0. An item with a high positive point-biserial correlation discriminates well between low-performing and high-performing students, whereas a negative point-biserial correlation indicates that lower-performing students did better on that item than higher-performing students. The point-biserial correlation coefficient ( $r_{pbi}$ ) for each item ( $i$ ) is derived as follows:

$$r_{pbi} = \frac{m_1 - m_x}{s_x} \sqrt{p_i / q_i}$$

where  $m_1$  is the mean value on the continuous variable  $x$  (such as total test score) for students who answer the item correctly;  $m_x$  is the mean value on the continuous variable  $x$  for the entire group;  $s_x$  is the standard deviation;  $p_i$  is the proportion of students answering item  $i$  correctly; and  $q_i$  is the proportion of students answering item  $i$  incorrectly.

Table 4.7 and Table 4.8 present the mean  $p$ -values and point-biserial correlations for each prompt. The results suggest that the prompts were generally hard for students. A low  $p$ -value of 0.2 or less (i.e., 20% or less of students responded correctly) indicates that the prompt is very hard and should be reviewed to determine if it is too difficult or mis-keyed. Prompts with a  $p$ -value lower than 0.2 are highlighted in the tables. Four of the six Grade 5 prompts highlighted are choice multiple, and two are drag-and-drop. Three of the nine Grade 8 prompts highlighted are choice multiple, two are drag-and-drop, and four are short text entry. These prompt types are relatively hard for students because they need to make sure all the correct choices are selected or put into the correct location for their answer to be considered correct.

Similarly, a point-biserial correlation lower than 0.2 indicates that the prompt does not discriminate between low- and high-performing students very well and needs to be reviewed. Prompts with a point-biserial lower than 0.2 are also highlighted in the tables. Three prompts at Grade 5 and one for Grade 8 have a point-biserial correlation lower than 0.2. Three out of these four prompts are very difficult based on their low  $p$ -values. A difficult prompt would have little variance in student responses since most students respond incorrectly. The resulting point-biserial correlation is typically low since both groups have the same score. Therefore, since few students answered these prompts correctly, they are not useful for differentiating students.

**Table 4.7. Prompt-Level  $P$ -Values and Point-Biserial Correlations—Grade 5**

| Form                   | Task               | Prompt ID            | Prompt Type* | N     | $P$ -value   | Point-biserial |       |
|------------------------|--------------------|----------------------|--------------|-------|--------------|----------------|-------|
| Form A                 | T1: Lake Bubbles   | T1_Q1                | MC           | 2,770 | 0.773        | 0.459          |       |
|                        |                    | T1_Q2                | D&D          | 2,751 | 0.051        | 0.317          |       |
|                        |                    | T1_Q3                | MC           | 2,758 | 0.551        | 0.435          |       |
|                        |                    | T1_Q4_Group          | D&D          | 2,679 | 0.047        | 0.345          |       |
|                        | T2: Plant Problems | T2_Q1_PartA          | MC           | 2,674 | 0.454        | 0.489          |       |
|                        |                    | T2_Q1_PartB          | CM           | 2,661 | 0.102        | 0.263          |       |
|                        |                    | T2_Q2_PartA          | MC           | 2,662 | 0.682        | 0.503          |       |
|                        |                    | T2_Q2_PartB          | MC           | 2,643 | 0.379        | 0.475          |       |
|                        |                    | T2_Q3                | HT           | 1,270 | 0.472        | 0.488          |       |
|                        |                    | T2_Q4                | CM           | 2,666 | 0.027        | 0.133          |       |
|                        |                    | T2_Q5                | CM           | 2,659 | 0.001        | 0.001          |       |
|                        | Form B             | T1: Something Smells | T1_Q1_Group  | D&D   | 2,495        | 0.601          | 0.474 |
|                        |                    |                      | T1_Q2_PartA  | MC    | 2,479        | 0.549          | 0.500 |
| T1_Q2_PartB            |                    |                      | MC           | 2,437 | 0.888        | 0.374          |       |
| T1_Q2_PartC            |                    |                      | MC           | 2,476 | 0.476        | 0.518          |       |
| T1_Q3                  |                    |                      | MC           | 2,463 | 0.728        | 0.423          |       |
| T1_Q4                  |                    |                      | MC           | 2,475 | 0.375        | 0.394          |       |
| T2: TV Trouble         |                    | T2_Q1                | MC           | 2,502 | 0.875        | 0.322          |       |
|                        |                    | T2_Q2                | MC           | 2,497 | 0.263        | 0.308          |       |
|                        |                    | T2_Q3                | MC           | 2,494 | 0.814        | 0.291          |       |
|                        |                    | T2_Q4                | MC           | 2,494 | 0.401        | 0.282          |       |
|                        |                    | T2_Q5                | MC           | 2,483 | 0.301        | 0.361          |       |
|                        |                    | T2_Q6                | MC           | 2,457 | 0.306        | 0.170          |       |
|                        |                    | T2_Q7                | MC           | 2,448 | 0.285        | 0.275          |       |
|                        |                    | T2_Q8                | CM           | 2,492 | 0.137        | 0.303          |       |
| <b>Grade 5 Average</b> |                    |                      |              |       | <b>0.422</b> | <b>0.356</b>   |       |

\*CM = choice multiple. D&D = drag-and-drop. HT = hot text. MC = multiple-choice.

**Table 4.8. Prompt-Level *P*-Values and Point-Biserial Correlations—Grade 8**

| Form   | Task                   | Prompt ID           | Prompt Type* | N     | <i>P</i> -value | Point-biserial |
|--------|------------------------|---------------------|--------------|-------|-----------------|----------------|
| Form A | T1: Fossil Discovery   | T1_Q1_PartA         | MC           | 3,062 | 0.790           | 0.375          |
|        |                        | T1_Q1_PartB         | CM           | 3,054 | 0.046           | 0.122          |
|        |                        | T1_Q2               | CM           | 3,059 | 0.165           | 0.329          |
|        |                        | T1_Q3               | MC           | 3,063 | 0.469           | 0.475          |
|        |                        | T1_Q4_PartA         | MC           | 3,058 | 0.746           | 0.418          |
|        |                        | T1_Q4_PartB         | MC           | 3,055 | 0.328           | 0.343          |
|        |                        | T1_Q5               | MC           | 3,059 | 0.756           | 0.466          |
|        |                        | T1_Q6_PartA         | MC           | 3,055 | 0.429           | 0.503          |
|        |                        | T1_Q6_PartB         | MC           | 3,053 | 0.279           | 0.319          |
|        | T2: Sculpture Setup    | T2_Q1_Group         | D&D          | 3,002 | 0.083           | 0.265          |
|        |                        | T2_Q2_PartA         | MC           | 3,042 | 0.370           | 0.443          |
|        |                        | T2_Q2_PartB         | TE           | 2,794 | 0.283           | 0.588          |
|        |                        | T2_Q3               | MC           | 3,035 | 0.598           | 0.465          |
|        |                        | T2_Q4_PartA         | MC           | 3,042 | 0.279           | 0.335          |
|        |                        | T2_Q4_PartB         | MC           | 3,041 | 0.417           | 0.344          |
|        |                        | T2_Q5_PartA         | TE           | 2,799 | 0.123           | 0.563          |
|        |                        | T2_Q5_PartB         | TE           | 2,803 | 0.263           | 0.511          |
|        |                        | T2_Q5_PartC         | TE           | 2,804 | 0.131           | 0.541          |
| Form B | T1: Pirate Bay         | T1_Q1_Group         | D&D          | 2,747 | 0.361           | 0.322          |
|        |                        | T1_Q2               | MC           | 2,754 | 0.537           | 0.247          |
|        |                        | T1_Q3               | CM           | 2,756 | 0.185           | 0.308          |
|        |                        | T1_Q4               | MC           | 2,759 | 0.539           | 0.334          |
|        |                        | T1_Q5_PartA         | MC           | 2,748 | 0.567           | 0.416          |
|        |                        | T1_Q5_PartB         | MC           | 2,755 | 0.245           | 0.216          |
|        |                        | T1_Q6_PartA         | MC           | 2,728 | 0.331           | 0.248          |
|        |                        | T1_Q6_PartB         | MC           | 2,752 | 0.222           | 0.263          |
|        |                        | T2: Sculpture Setup | T2_Q1_Group  | D&D   | 2,724           | 0.132          |
|        | T2_Q2_PartA            |                     | MC           | 2,745 | 0.359           | 0.461          |
|        | T2_Q2_PartB            |                     | TE           | 2,490 | 0.277           | 0.516          |
|        | T2_Q3                  |                     | MC           | 2,727 | 0.586           | 0.465          |
|        | T2_Q4_PartA            |                     | MC           | 2,736 | 0.284           | 0.294          |
|        | T2_Q4_PartB            |                     | MC           | 2,744 | 0.412           | 0.314          |
|        | T2_Q5_PartA            |                     | TE           | 2,514 | 0.104           | 0.544          |
|        | T2_Q5_PartB            |                     | TE           | 2,519 | 0.256           | 0.457          |
|        | T2_Q5_PartC            |                     | TE           | 2,514 | 0.130           | 0.503          |
|        | <b>Grade 8 Average</b> |                     |              |       |                 | <b>0.345</b>   |

\*CM = choice multiple. D&D = drag-and-drop. HT = hot text. MC = multiple-choice. TE = text entry.

#### 4.6. Distractor Analysis

A distractor is an incorrect or inferior answer option in a multiple-choice item. Distractor analysis is an extension of item analysis and is often conducted similarly to item difficulty and item discrimination. The quality of distractors is an important component of an item’s overall quality. Distractors should be clearly incorrect yet plausible and attractive to lower-ability students. Distractor analyses were conducted for all the multiple-choice prompts in the pilot to evaluate the quality of the distractors. Specifically, the following distractor analyses were conducted:

1. Calculate the percentage of students that selected each answer option. Use the results to determine if each option is attractive and selected by a certain proportion of students.
2. Calculate the point-biserial correlation for each option. While the correct answer should have a positive point-biserial correlation with the total score, the distractors should exhibit negative point-biserial correlations (i.e., lower-ability students would likely choose the distractors while higher-ability students would not).

Table 4.9 – Table 4.12 present the percentage of students that selected each answer option of a multiple-choice prompt by grade and form. Bolded numbers indicate the percentage of students that selected the correct answer, or the key. Highlighted numbers indicate when the percentage of students who selected a distractor is higher than the percentage of students who selected the key. These prompts should be reviewed to determine if the distractors are too attractive or the prompt has double keys.

**Table 4.9. Percentage of Students Selecting Each Answer Option—Grade 5, Form A**

| Prompt ID   | N     | %Students Selecting Each Answer Option |             |             |             |             | Other* | NR* |
|-------------|-------|--|-------------|-------------|-------------|-------------|--------|-----|
|             |       | Option 1                               | Option 2    | Option 3    | Option 4    | Option 5    |        |     |
| T1_Q1       | 2,795 | 10.2                                   | 7.3         | <b>76.6</b> | 5.0         | –           | –      | 0.9 |
| T1_Q2_PartA | 2,795 | 18.7                                   | 7.5         | 7.2         | 7.3         | <b>56.7</b> | 0.5    | 2.0 |
| T1_Q2_PartB | 2,795 | 10.9                                   | <b>30.4</b> | 22.5        | 27.5        | 6.4         | 0.3    | 2.0 |
| T1_Q2_PartC | 2,795 | <b>36.9</b>                            | 8.8         | 14.5        | 20.4        | 17.0        | 0.3    | 2.1 |
| T1_Q2_PartD | 2,795 | 6.4                                    | <b>40.2</b> | 20.5        | <b>26.9</b> | 3.5         | 0.3    | 2.1 |
| T1_Q2_PartE | 2,795 | 25.5                                   | 10.9        | <b>32.8</b> | 15.1        | 13.2        | 0.5    | 2.0 |
| T1_Q3       | 2,795 | 2.9                                    | <b>54.4</b> | 22.0        | 13.1        | 6.2         | –      | 1.3 |
| T2_Q1_PartA | 2,795 | 5.4                                    | 14.9        | 43.4        | 32.0        | –           | –      | 4.3 |
| T2_Q2_PartA | 2,795 | 30.3                                   | <b>64.9</b> | –           | –           | –           | –      | 4.8 |
| T2_Q2_PartB | 2,795 | 20.1                                   | 13.6        | <b>35.8</b> | 25.0        | –           | –      | 5.4 |

\*Other = students selected something other than 1–5. NR = no response.

**Table 4.10. Percentage of Students Selecting Each Answer Option—Grade 5, Form B**

| Prompt ID   | N     | %Students Selecting Each Answer Option |             |             |             |             | NR* |
|-------------|-------|--|-------------|-------------|-------------|-------------|-----|
|             |       | Option 1                               | Option 2    | Option 3    | Option 4    | Option 5    |     |
| T1_Q2_PartA | 2,521 | 23.8                                   | <b>54.0</b> | 12.5        | 8.0         | –           | 1.7 |
| T1_Q2_PartB | 2,521 | <b>85.8</b>                            | 5.6         | 5.3         | –           | –           | 3.3 |
| T1_Q2_PartC | 2,521 | 19.9                                   | <b>46.8</b> | 17.5        | 14.0        | –           | 1.8 |
| T1_Q3       | 2,521 | 12.4                                   | 4.8         | <b>71.2</b> | 9.3         | –           | 2.3 |
| T1_Q4       | 2,521 | 12.5                                   | <b>36.8</b> | 11.7        | <b>37.2</b> | –           | 1.8 |
| T2_Q1       | 2,521 | <b>86.8</b>                            | 4.4         | 0.8         | 7.2         | –           | 0.8 |
| T2_Q2       | 2,521 | 18.4                                   | <b>31.6</b> | <b>26.0</b> | 23.0        | –           | 1.0 |
| T2_Q3       | 2,521 | <b>51.1</b>                            | 3.7         | 9.7         | 5.0         | <b>29.5</b> | 1.1 |
| T2_Q4       | 2,521 | 15.9                                   | 30.8        | <b>39.7</b> | 12.6        | –           | 1.1 |
| T2_Q5       | 2,521 | 19.4                                   | 25.9        | 23.6        | <b>29.7</b> | –           | 1.5 |
| T2_Q6       | 2,521 | <b>36.5</b>                            | <b>29.8</b> | 15.0        | 16.2        | –           | 2.5 |
| T2_Q7       | 2,521 | 14.0                                   | <b>27.7</b> | <b>38.0</b> | 17.5        | –           | 2.9 |

\*NR = no response.

**Table 4.11. Percentage of Students Selecting Each Answer Option—Grade 8, Form A**

| Prompt ID   | N     | %Students Selecting Each Answer Option |             |             |             |             | NR* |
|-------------|-------|--|-------------|-------------|-------------|-------------|-----|
|             |       | Option 1                               | Option 2    | Option 3    | Option 4    | Option 5    |     |
| T1_Q1_PartA | 3,090 | <b>78.3</b>                            | 20.8        | –           | –           | –           | 0.9 |
| T1_Q3       | 3,090 | 9.7                                    | 18.9        | 7.4         | 16.6        | <b>46.5</b> | 0.9 |
| T1_Q4_PartA | 3,090 | 13.6                                   | 7.7         | <b>73.8</b> | 3.9         | –           | 1.0 |
| T1_Q4_PartB | 3,090 | <b>55.1</b>                            | 5.7         | 5.7         | <b>32.4</b> | –           | 1.1 |
| T1_Q5       | 3,090 | 6.6                                    | 6.6         | <b>74.9</b> | 10.8        | –           | 1.0 |
| T1_Q6_PartA | 3,090 | 14                                     | 32.6        | <b>42.4</b> | 9.8         | –           | 1.1 |
| T1_Q6_PartB | 3,090 | <b>27.6</b>                            | <b>48.9</b> | 15.2        | 7.1         | –           | 1.2 |
| T2_Q2_PartA | 3,090 | <b>40.1</b>                            | 15.3        | 6.6         | <b>36.5</b> | –           | 1.6 |
| T2_Q3       | 3,090 | 13.9                                   | <b>58.7</b> | 14.1        | 11.5        | –           | 1.8 |
| T2_Q4_PartA | 3,090 | 8.2                                    | 10.4        | <b>52.4</b> | <b>27.5</b> | –           | 1.6 |
| T2_Q4_PartB | 3,090 | 23.6                                   | <b>41.1</b> | 20.5        | 13.2        | –           | 1.6 |

\*NR = no response.

**Table 4.12. Percentage of Students Selecting Each Answer Option—Grade 8, Form B**

| Prompt ID   | N     | %Students Selecting Each Answer Option |             |             |             | NR* |
|-------------|-------|--|-------------|-------------|-------------|-----|
|             |       | Option 1                               | Option 2    | Option 3    | Option 4    |     |
| T1_Q2       | 2,779 | 6.8                                    | <b>53.3</b> | 36.1        | 3.0         | 0.9 |
| T1_Q4       | 2,779 | <b>53.5</b>                            | 9.6         | 9.8         | 26.4        | 0.7 |
| T1_Q5_PartA | 2,779 | <b>56.1</b>                            | 9.9         | 32.9        | –           | 1.1 |
| T1_Q5_PartB | 2,779 | 6.7                                    | <b>33.8</b> | <b>34.4</b> | <b>24.3</b> | 0.9 |
| T1_Q6_PartA | 2,779 | <b>32.5</b>                            | 11.0        | <b>54.6</b> | –           | 1.8 |
| T1_Q6_PartB | 2,779 | <b>34.0</b>                            | 15.2        | <b>22.0</b> | <b>27.9</b> | 1.0 |
| T2_Q2_PartA | 2,779 | <b>40.4</b>                            | 15.7        | 7.2         | <b>35.5</b> | 1.2 |
| T2_Q3       | 2,779 | 13.1                                   | <b>57.5</b> | 15.9        | 11.6        | 1.9 |
| T2_Q4_PartA | 2,779 | 7.7                                    | 11.3        | <b>51.5</b> | <b>27.9</b> | 1.5 |
| T2_Q4_PartB | 2,779 | 22.9                                   | <b>40.7</b> | 21.5        | 13.6        | 1.3 |

\*NR = no response.

Table 4.13 –Table 4.16 present the point-biserial correlation for each answer option of a multiple-choice prompt by grade and test form. If the key and distractors work well, the key should have a positive point-biserial correlation with the overall test score, and the distractors should exhibit negative point-biserial correlations. As shown by the bolded numbers that indicate the answer keys, all keys have positive point-biserial correlations, and most distractors have negative point-biserial correlations. However, several distractors have positive point-biserial correlations, which indicates that students who selected these distractors tend to have higher total scores. These distractors are highlighted in the tables and should be reviewed or modified.

**Table 4.13. Point-Biserial Correlations for Each Answer Option—Grade 5, Form A**

| Prompt ID   | N     | Point-Biserial Correlation |              |              |              |              |
|-------------|-------|----------------------------|--------------|--------------|--------------|--------------|
|             |       | Option 1                   | Option 2     | Option 3     | Option 4     | Option 5     |
| T1_Q1       | 2,770 | -0.289                     | -0.266       | <b>0.459</b> | -0.160       | –            |
| T1_Q2_PartA | 2,724 | -0.076                     | -0.054       | -0.023       | -0.050       | <b>0.132</b> |
| T1_Q2_PartB | 2,731 | <b>0.028</b>               | <b>0.192</b> | -0.056       | -0.144       | -0.034       |
| T1_Q2_PartC | 2,729 | <b>0.020</b>               | -0.026       | -0.022       | <b>0.073</b> | -0.059       |
| T1_Q2_PartD | 2,728 | -0.029                     | -0.098       | <b>0.050</b> | <b>0.119</b> | -0.083       |
| T1_Q2_PartE | 2,727 | <b>0.036</b>               | -0.043       | <b>0.027</b> | -0.004       | -0.033       |
| T1_Q3       | 2,758 | -0.139                     | <b>0.435</b> | -0.267       | -0.148       | -0.130       |
| T2_Q1_PartA | 2,674 | -0.147                     | -0.274       | <b>0.489</b> | -0.233       | –            |
| T2_Q2_PartA | 2,662 | -0.503                     | <b>0.503</b> | –            | –            | –            |
| T2_Q2_PartB | 2,643 | -0.295                     | -0.223       | <b>0.475</b> | -0.072       | –            |

**Table 4.14. Point-Biserial Correlations for Each Answer Option—Grade 5, Form B**

| Prompt ID   | N     | Point-Biserial Correlation |              |              |              |              |
|-------------|-------|----------------------------|--------------|--------------|--------------|--------------|
|             |       | Option 1                   | Option 2     | Option 3     | Option 4     | Option 5     |
| T1_Q2_PartA | 2,479 | -0.338                     | <b>0.500</b> | -0.198       | -0.138       | –            |
| T1_Q2_PartB | 2,437 | <b>0.374</b>               | -0.259       | -0.253       | –            | –            |
| T1_Q2_PartC | 2,476 | -0.244                     | <b>0.518</b> | -0.265       | -0.170       | –            |
| T1_Q3       | 2,463 | -0.227                     | -0.230       | <b>0.423</b> | -0.213       | –            |
| T1_Q4       | 2,475 | -0.252                     | <b>0.394</b> | -0.161       | -0.113       | –            |
| T2_Q1       | 2,502 | <b>0.322</b>               | -0.212       | -0.089       | -0.211       | –            |
| T2_Q2       | 2,497 | -0.034                     | -0.175       | <b>0.308</b> | -0.096       | –            |
| T2_Q3       | 2,494 | <b>0.210</b>               | -0.082       | -0.202       | -0.171       | <b>0.017</b> |
| T2_Q4       | 2,494 | -0.149                     | -0.073       | <b>0.282</b> | -0.148       | –            |
| T2_Q5       | 2,483 | -0.154                     | -0.124       | -0.116       | <b>0.361</b> | –            |
| T2_Q6       | 2,457 | -0.027                     | <b>0.170</b> | -0.059       | -0.118       | –            |
| T2_Q7       | 2,448 | -0.133                     | <b>0.275</b> | -0.082       | -0.097       | –            |

**Table 4.15. Point-Biserial Correlations for Each Answer Option—Grade 8, Form A**

| Prompt ID   | N     | Point-Biserial Correlation |              |              |              |              |
|-------------|-------|----------------------------|--------------|--------------|--------------|--------------|
|             |       | Option 1                   | Option 2     | Option 3     | Option 4     | Option 5     |
| T1_Q1_PartA | 3,062 | <b>0.375</b>               | -0.375       | –            | –            | –            |
| T1_Q3       | 3,063 | -0.184                     | -0.326       | -0.086       | -0.086       | <b>0.475</b> |
| T1_Q4_PartA | 3,058 | -0.225                     | -0.264       | <b>0.418</b> | -0.175       | –            |
| T1_Q4_PartB | 3,055 | -0.131                     | -0.225       | -0.187       | <b>0.343</b> | –            |
| T1_Q5       | 3,059 | -0.271                     | -0.248       | <b>0.466</b> | -0.226       | –            |
| T1_Q6_PartA | 3,055 | -0.289                     | -0.307       | <b>0.503</b> | -0.012       | –            |
| T1_Q6_PartB | 3,053 | <b>0.319</b>               | <b>0.046</b> | -0.323       | -0.191       | –            |
| T2_Q2_PartA | 3,042 | -0.190                     | -0.184       | -0.216       | <b>0.443</b> | –            |
| T2_Q3       | 3,035 | -0.208                     | <b>0.465</b> | -0.293       | -0.165       | –            |
| T2_Q4_PartA | 3,042 | -0.143                     | -0.177       | -0.112       | <b>0.335</b> | –            |
| T2_Q4_PartB | 3,041 | -0.094                     | <b>0.344</b> | -0.222       | -0.115       | –            |

**Table 4.16. Point-Biserial Correlations for Each Answer Option—Grade 8, Form B**

| Prompt ID   | N     | Point-Biserial Correlation |              |              |              |          |
|-------------|-------|----------------------------|--------------|--------------|--------------|----------|
|             |       | Option 1                   | Option 2     | Option 3     | Option 4     | Option 5 |
| T1_Q2       | 2,754 | -0.162                     | <b>0.247</b> | -0.139       | -0.093       | –        |
| T1_Q4       | 2,759 | <b>0.334</b>               | -0.242       | -0.250       | -0.047       | –        |
| T1_Q5_PartA | 2,748 | <b>0.416</b>               | -0.258       | -0.273       | –            | –        |
| T1_Q5_PartB | 2,755 | -0.138                     | <b>0.072</b> | -0.194       | <b>0.216</b> | –        |
| T1_Q6_PartA | 2,728 | <b>0.248</b>               | -0.227       | -0.090       | –            | –        |
| T1_Q6_PartB | 2,752 | -0.127                     | -0.136       | <b>0.263</b> | 0.000        | –        |
| T2_Q2_PartA | 2,745 | -0.216                     | -0.156       | -0.223       | <b>0.461</b> | –        |
| T2_Q3       | 2,727 | -0.174                     | <b>0.465</b> | -0.293       | -0.192       | –        |
| T2_Q4_PartA | 2,736 | -0.108                     | -0.186       | -0.088       | <b>0.294</b> | –        |
| T2_Q4_PartB | 2,744 | -0.077                     | <b>0.314</b> | -0.220       | -0.090       | –        |

#### 4.7. Student Feedback Results

At the end of each task, the following student feedback questions were administered to collect information about students' interest and perception of the task and prompts. Figure 4.1 presents an example of what these questions looked like at the end of a task. Students answered based on a rating scale of 1 to 5, with 1 being not interesting or not difficult and 5 being very interesting or very difficult. For the last Student Feedback question, students provided a rating for each prompt associated with the task.

1. How interesting was the task you just completed?
2. How difficult was the task as a whole?
3. Please rate the difficulty to understand how to respond to each question.

**Figure 4.1. Student Feedback Questions**

**Student Feedback**

**Feedback Question #1**

How interesting was the task you just completed?

1 star is not interesting and 5 stars is very interesting.

☆☆☆☆☆

**Feedback Question #2**

How difficult was the task as a whole?

1 star is not difficult and 5 stars is very difficult.

☆☆☆☆☆

**Feedback Question #3**

Please rate the difficulty **to understand how to respond to each question.**

1 star is not difficult and 5 stars is very difficult.

Question 1 ☆☆☆☆☆

Question 2 ☆☆☆☆☆

Question 3 ☆☆☆☆☆

Question 4 ☆☆☆☆☆

Question 5 ☆☆☆☆☆

#### 4.7.1. Feedback Question #1: Task Interestingness

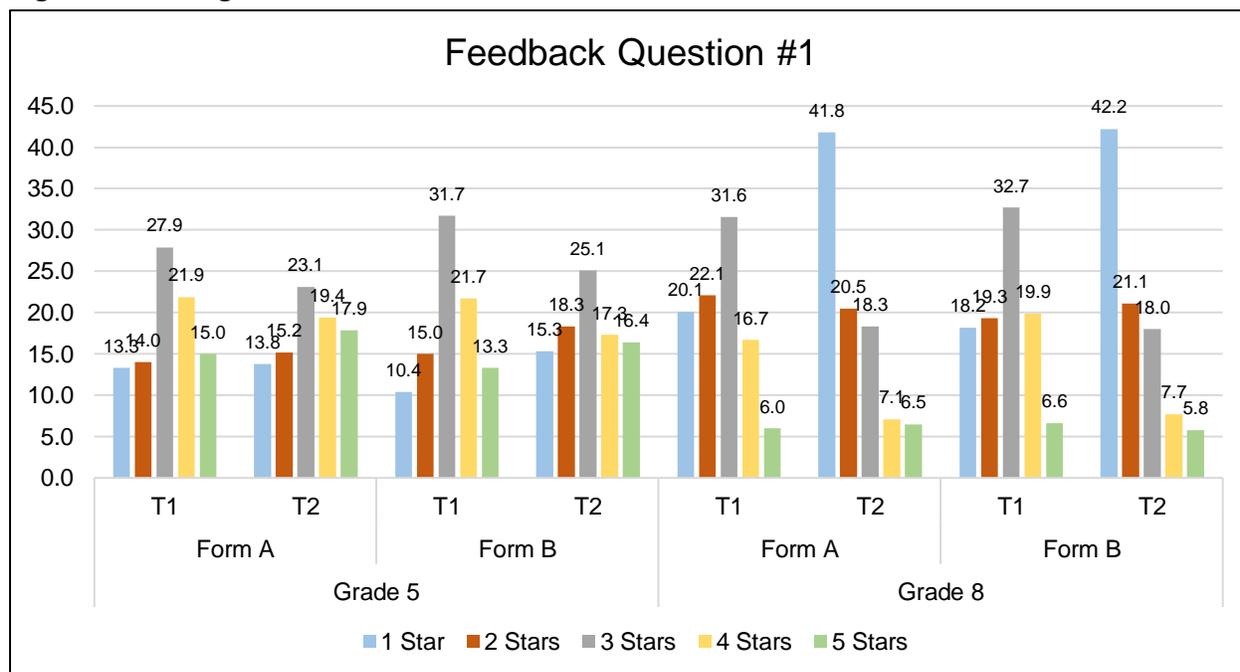
Table 4.17 presents the rating results (i.e., the percentage of students that gave 1 to 5 stars) for Feedback Question #1, along with the average rating across all students. Figure 4.2 presents a graphical representation of the results. As shown in the table and chart, the least interesting task was Task 2: Sculpture Setup on both the Grade 8 forms. The most interesting tasks (i.e., the tasks with the biggest percentages of 5-star ratings) are all from the Grade 5 forms.

**Table 4.17. Rating Percentage Distribution—Student Feedback Question #1**

| Grade | Form   | Task                 | N     | 1 Star | 2 Stars | 3 Stars | 4 & 5 Stars Combined | NR*  | Average |
|-------|--------|----------------------|-------|--------|---------|---------|----------------------|------|---------|
| 5     | Form A | T1: Lake Bubbles     | 2,795 | 13.3   | 14.0    | 27.9    | 36.9                 | 7.8  | 3.1     |
|       |        | T2: Plant Problems   | 2,795 | 13.8   | 15.2    | 23.1    | 37.3                 | 10.5 | 3.1     |
|       | Form B | T1: Something Smells | 2,521 | 10.4   | 15.0    | 31.7    | 35.0                 | 7.9  | 3.1     |
|       |        | T2: TV Trouble       | 2,521 | 15.3   | 18.3    | 25.1    | 33.7                 | 7.5  | 3.0     |
| 8     | Form A | T1: Fossil Discovery | 3,090 | 20.1   | 22.1    | 31.6    | 22.7                 | 3.5  | 2.7     |
|       |        | T2: Sculpture Setup  | 3,090 | 41.8   | 20.5    | 18.3    | 13.6                 | 5.8  | 2.1     |
|       | Form B | T1: Pirate Bay       | 2,779 | 18.2   | 19.3    | 32.7    | 26.5                 | 3.1  | 2.8     |
|       |        | T2: Sculpture Setup  | 2,779 | 42.2   | 21.1    | 18.0    | 13.5                 | 5.3  | 2.1     |

\*NR = no response.

**Figure 4.2. Ratings for Student Feedback Question #1**



#### 4.7.2. Feedback Questions #2 and #3: Task and Item Difficulty

Table 4.18 presents the rating results (i.e., the percentage of students that gave 1 to 5 stars) for Feedback Question #2, along with the average rating across all students. Figure 4.3 presents a graphical representation of the results. As shown in the table and chart, the least difficult task was Task 1: Fossil Discovery on the Grade 8 Form A. The most difficult task was Task 2: Sculpture Setup on both the Grade 8 forms. Students also thought these were the least interesting tasks according to the Feedback Question #1 results.

**Table 4.18. Rating Percentage Distribution—Student Feedback Question #2**

| Grade | Form   | Task                 | N     | 1 Star | 2 Stars | 3 Stars | 4 Stars | 5 Stars | NR*  | Average |
|-------|--------|----------------------|-------|--------|---------|---------|---------|---------|------|---------|
| 5     | Form A | T1: Lake Bubbles     | 2,795 | 6.3    | 15.9    | 32.4    | 25.3    | 11.8    | 8.3  | 3.2     |
|       |        | T2: Plant Problems   | 2,795 | 8.3    | 18.4    | 31.4    | 21.0    | 10.6    | 10.4 | 3.1     |
|       | Form B | T1: Something Smells | 2,521 | 8.1    | 19.8    | 33.2    | 22.5    | 8.8     | 7.7  | 3.0     |
|       |        | T2: TV Trouble       | 2,521 | 7.8    | 15.0    | 30.2    | 26.7    | 12.8    | 7.6  | 3.2     |
| 8     | Form A | T1: Fossil Discovery | 3,090 | 8.1    | 23.4    | 36.6    | 22.8    | 5.7     | 3.3  | 2.9     |
|       |        | T2: Sculpture Setup  | 3,090 | 5.5    | 8.9     | 20.9    | 30.9    | 28.4    | 5.4  | 3.7     |
|       | Form B | T1: Pirate Bay       | 2,779 | 7.1    | 21.5    | 39.7    | 21.0    | 7.6     | 3.0  | 3.0     |
|       |        | T2: Sculpture Setup  | 2,779 | 6.4    | 8.7     | 21.0    | 30.2    | 28.1    | 5.6  | 3.7     |

\*NR = no response.

**Figure 4.3. Ratings for Student Feedback Question #2**

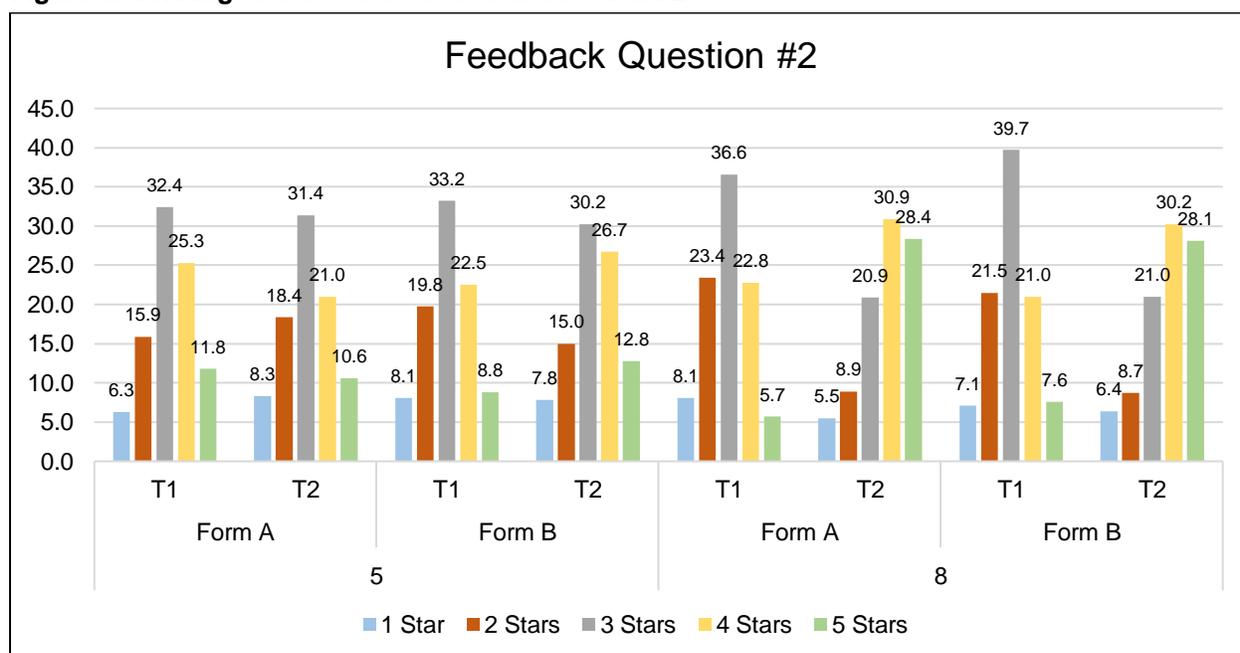


Table 4.19 presents the results for the prompt-level difficulty responses for Student Feedback Question #3. Students' perceptions of the prompt difficulty are related to the empirical difficulty indicated by the *p*-values. The correlation between the average difficulty rating in Table 4.19 and the *p*-values presented in Table 4.7 and Table 4.8 is  $-.488$ , which indicates that the higher the difficulty rating is, the lower the *p*-value. Thus, students' perceptions of the prompts' difficulty provided from the Student Feedback questions were moderately related to empirical item difficulty.

**Table 4.19. Rating Percentage Distribution—Student Feedback Question #3**

| Grade               | Form                | Task                 | Prompt # | N     | 1 Star | 2 Stars | 3 Stars | 4 Stars | 5 Stars | N/A | Average |
|---------------------|---------------------|----------------------|----------|-------|--------|---------|---------|---------|---------|-----|---------|
| 5                   | Form A              | T1: Lake Bubbles     | 1        | 2,795 | 32.1   | 22.0    | 19.8    | 9.1     | 9.4     | 7.8 | 2.4     |
|                     |                     |                      | 2        | 2,795 | 14.9   | 26.2    | 26.7    | 14.4    | 10.3    | 7.5 | 2.8     |
|                     |                     |                      | 3        | 2,795 | 11.1   | 22.0    | 30.5    | 17.6    | 10.7    | 8.0 | 2.9     |
|                     |                     |                      | 4        | 2,795 | 11.1   | 17.4    | 26.3    | 23.3    | 14.8    | 7.2 | 3.1     |
|                     |                     |                      | 5        | 2,795 | 9.8    | 10.2    | 15.4    | 21.1    | 35.6    | 8.0 | 3.7     |
|                     |                     | T2: Plant Problems   | 1        | 2,795 | 27.3   | 23.8    | 20.6    | 9.6     | 9.8     | 8.8 | 2.5     |
|                     |                     |                      | 2        | 2,795 | 18.8   | 24.9    | 25.0    | 13.7    | 8.5     | 9.1 | 2.7     |
|                     |                     |                      | 3        | 2,795 | 18.0   | 21.6    | 25.5    | 15.1    | 10.5    | 9.3 | 2.8     |
|                     |                     |                      | 4        | 2,795 | 12.7   | 17.8    | 26.3    | 21.3    | 12.6    | 9.3 | 3.0     |
|                     |                     |                      | 5        | 2,795 | 18.9   | 19.6    | 21.0    | 15.4    | 15.6    | 9.5 | 2.9     |
|                     | Form B              | T1: Something Smells | 1        | 2,521 | 40.3   | 20.4    | 15.1    | 8.6     | 9.6     | 6.0 | 2.2     |
|                     |                     |                      | 2        | 2,521 | 10.1   | 19.8    | 29.2    | 21.9    | 13.2    | 5.8 | 3.1     |
|                     |                     |                      | 3        | 2,521 | 19.3   | 25.6    | 26.4    | 13.8    | 9.1     | 5.8 | 2.7     |
|                     |                     |                      | 4        | 2,521 | 14.1   | 20.0    | 24.4    | 21.5    | 13.6    | 6.3 | 3.0     |
|                     |                     |                      | 5        | 2,521 | 20.3   | 17.4    | 19.2    | 18.0    | 18.4    | 6.7 | 3.0     |
|                     |                     | T2: TV Trouble       | 1        | 2,521 | 32.6   | 22.6    | 20.4    | 9.5     | 8.8     | 6.1 | 2.4     |
|                     |                     |                      | 2        | 2,521 | 13.7   | 22.5    | 31.6    | 17.3    | 8.8     | 6.1 | 2.8     |
|                     |                     |                      | 3        | 2,521 | 25.3   | 23.7    | 23.2    | 12.8    | 8.9     | 6.1 | 2.5     |
| 8                   | Form A              | T1: Fossil Discovery | 4        | 2,521 | 15.7   | 23.9    | 31.7    | 18.0    | 7.6     | 3.1 | 2.8     |
|                     |                     |                      | 5        | 3,090 | 26.1   | 22.3    | 25.7    | 14.6    | 8.1     | 3.1 | 2.5     |
| 6                   |                     |                      | 3,090    | 15.3  | 22.5   | 31.0    | 17.9    | 9.8     | 3.5     | 2.8 |         |
| T2: Sculpture Setup |                     |                      | 1        | 3,090 | 9.9    | 14.9    | 24.7    | 21.8    | 24.2    | 4.5 | 3.4     |
|                     |                     |                      | 2        | 3,090 | 8.4    | 18.0    | 28.7    | 22.2    | 18.0    | 4.7 | 3.2     |
|                     |                     |                      | 3        | 3,090 | 9.4    | 16.6    | 28.0    | 22.2    | 19.1    | 4.7 | 3.3     |
|                     |                     | 4                    | 3,090    | 8.3   | 14.3   | 27.6    | 24.7    | 20.4    | 4.7     | 3.4 |         |
|                     |                     | 5                    | 3,090    | 8.9   | 11.5   | 22.0    | 23.1    | 30.0    | 4.4     | 3.6 |         |
|                     |                     | 6                    | 3,090    | 9.0   | 11.0   | 19.3    | 23.0    | 32.6    | 5.1     | 3.6 |         |
| Form B              |                     | T1: Pirate Bay*      | 1        | 2,779 | 15.7   | 21.3    | 28.4    | 19.7    | 12.4    | 2.5 | 2.9     |
|                     |                     |                      | 2        | 2,779 | 17.0   | 27.5    | 30.8    | 14.8    | 7.1     | 2.9 | 2.7     |
|                     |                     |                      | 3        | 2,779 | 16.7   | 25.4    | 31.5    | 15.1    | 8.4     | 2.8 | 2.7     |
|                     | 4                   |                      | 2,779    | 12.8  | 23.7   | 30.4    | 20.4    | 9.6     | 3.1     | 2.9 |         |
|                     | 5                   |                      | 2,779    | 13.0  | 21.0   | 31.3    | 19.6    | 12.0    | 3.0     | 3.0 |         |
|                     | T2: Sculpture Setup | 1                    | 2,779    | 9.5   | 15.1   | 26.3    | 20.7    | 24.3    | 4.1     | 3.4 |         |
|                     |                     | 2                    | 2,779    | 8.1   | 15.9   | 31.0    | 22.3    | 18.4    | 4.3     | 3.3 |         |
|                     |                     | 3                    | 2,779    | 8.2   | 15.2   | 30.9    | 22.0    | 19.1    | 4.6     | 3.3 |         |
|                     |                     | 4                    | 2,779    | 7.5   | 15.4   | 26.7    | 24.5    | 21.5    | 4.5     | 3.4 |         |
|                     |                     | 5                    | 2,779    | 7.9   | 11.9   | 22.2    | 24.0    | 29.4    | 4.5     | 3.6 |         |
|                     |                     | 6                    | 2,779    | 8.9   | 10.5   | 20.2    | 22.6    | 32.5    | 5.4     | 3.6 |         |

\*The survey question for Grade 8, Form B, Task 1 did not include a question for Prompt 6.

#### 4.8. Cognitive Lab Results

The cognitive lab forms included the questions in Table 4.20. Proctors documented the students' responses as they progressed through the forms. Appendix H presents the tasks and prompts for Grade 5 Form B and Grade 8 Form A used in the cognitive labs, and Appendix I presents the responses to the questions in Table 4.20. Some questions were content-specific, while others garnered more general responses that were grouped together by similarities and analyzed based on the assignment of codes. Appendix J presents these coding results.

**Table 4.20. Cognitive Lab Questions**

| <b>Cognitive Lab Questions</b>   |
|--|
| 1. After they have finished ask the student:<br>A. Did you understand the text?<br>B. What are you wondering after reading the passage?<br>C. What science knowledge might you need to understand smells?<br>Record their answers and any observations about the student's interaction with the scenario.                                      |
| 2. Ask the student to describe what they are thinking as they read and answer the question. If the student is not talking, you may ask "What are you thinking?" or "Why did you answer that way?" or "What would help you answer this question?".  |
| 3. The student's comments provided evidence that the item elicited thinking about:<br>1. [varies by item]<br>2. [varies by item]<br>3. Other   |
| 4. If "other" what thinking did the item elicit?   |
| 5. While reading and interacting with the question, the student appeared to...<br>1. rush to answer the question without careful consideration of the question, the text, and how to respond.<br>2. be engaged with the question, taking time to consider and respond.<br>3. be able to understand how to generate a response to the question. |
| 6. Ask the student: Was it easy or hard to know how to respond to the question?<br>1. Easy<br>2. Medium<br>3. Hard   |
| 7. Ask the student: What would help you answer the question?   |
| 8. Ask the student: Was it easy or hard to use the given information to answer the question?<br>1. Easy<br>2. Medium<br>3. Hard  |
| 9. Ask the student: Why was it easy or hard?   |
| 10. Record any additional notes, comments, or observations made about the student's interaction with this question.  |
| 11. Ask the student: Tell me what you learned about _____. Record their response.  |
| 12. Ask the student: What are the science tests like in your class? Record their response.   |
| 13. Ask the student: How does that compare with what you just did? Record their response.  |
| 14. Ask the student: Why did you rate your interest in the task with 1, 2, 3, 4, or 5 stars?   |
| 15. Ask the student: Why did you rate the difficulty of the task with 1, 2, 3, 4, or 5 stars?  |
| 16. Ask the student: Why did you rate the difficulty of the questions with 1, 2, 3, 4, or 5 stars?   |

After reviewing the cognitive lab responses, there are some general themes that should be considered carefully by NDE and NWEA before item development in Summer 2019:

- The students had difficulty with the graphic gap match prompts. There was confusion about where to place responses. The force diagram in Grade 8 (Task 2 in Forms A and B) was especially confusing. This was a concern prior to the administration since the Qualtrics platform cannot do graphical gap match prompts. However, the NWEA engine can use graphic gap match prompts and does so in a way that is more intuitive for students.
- Students were concerned that they had not worked with the content in class yet, or they had worked on the content the previous year.
- In Grade 5, students felt that the pilot assessments were generally more difficult than their classroom tests. The diagrams helped, but word choice made the assessment more difficult. That is likely a construct-irrelevant concern that will need to be addressed in development. Construct-irrelevance is when factors unrelated to the concept being tested affect a student's ability to respond correctly. For example, an ELL student may not know what a nickel or dollar is, so a math problem asking students how many nickels are in a dollar has a construct-irrelevant feature that may negatively impact their score, even though it is a simple math problem.
- In Grade 5, there seemed to be a considerable difference in the background knowledge students brought to the pilot. One comment stated, "it wasn't boring but I know a lot of stuff about that so it wasn't as interesting," while another said, "taught me some things I didn't know about how the sun is lower in winter, higher in summer; explains why the seasons are the way they are." This dichotomy may make it difficult to find truly novel or familiar phenomena.
- In Grade 8, the Sculpture Setup task was considered more difficult than the Fossil Discovery task. Some students felt that being able to see all the information throughout the task was very beneficial. A few considered it to be very easy because of that. When they could not go back, they considered the prompts to be more difficult.

Overall, most students indicated that they understood each task and that they knew how to respond to the prompts. Students also appeared to be engaged with the prompts, taking time to consider and respond. While students often noted that the graphics were helpful, they also often indicated that the graphics and prompts could provide clearer direction and more information throughout the cognitive lab. Being able to refer back to previous prompts and graphics helped students respond, and a few students commented that the progression of the prompts throughout the task was a nice feature of this assessment. Students found many of the pilot tasks to be interesting and noted that they learned something from them. Students commented that their classroom tests were easier than the pilot, especially in Grade 5, often because they know what to expect on them and the format of the prompts are easier to respond to. However, they also commented that the pilot provided more graphics and information, which they liked.

#### 4.9. Considerations for Future Development

The following considerations should be reviewed prior to the beginning of the development process in Summer 2019:

1. *Art*: A greater emphasis should be placed on making the art accessible for visually impaired students. Whether the impairment is color blindness or loss of visual acuity, art must still allow students to grasp the concept being displayed. However, almost all students found the graphics to be useful. Thus, continued use of graphics is important, just more care should be made in creating them.
- *Hot Text*: This prompt type did not seem to work as well as intended for selecting evidence text out of paragraphs. While the formatting of the hot text prompts in Qualtrics are different than in the NWEA test engine, concerns raised by NDE during the task review meeting are understandable. NWEA suggests not using this prompt type until these concerns are resolved.
- *Multiple choice vs. choice multiple prompts*: Students mentioned that the different shapes of the selection boxes told them that they would be able to choose more than one answer. That difference is not present in the NWEA engine, so directions or command lines should be used to ensure that students know what their response is expected to look like.
- *Word Choice*: The Grade 5 students felt that the word choice and phrasing was difficult. Using lower grade-level words for non-content material and simpler and more precise sentences should resolve this concern.

## Section 5: Fairness, Reliability, and Validity

This chapter presents results from the following analyses to provide evidence of fairness, reliability, and validity:

- Differential item functioning (DIF) analysis
- Internal consistency (Cronbach's alpha)
- Concurrent validity (Pearson correlation coefficient)

### 5.1. Differential Item Functioning (DIF)

DIF analysis was conducted to examine the fairness of the pilot test. DIF is a statistical procedure that flags items for potential bias. The fundamental measurement assumption of DIF is that the probability of a correct response to a test item is a function of the item's difficulty and the student's ability. This function is expected to remain invariant to other person characteristics unrelated to ability such as gender and ethnicity. Therefore, if two students with the same ability respond to the same item, they are assumed to have an equal probability of answering the item correctly.

To test this assumption, responses to items by students sharing an aspect of a person characteristic (e.g., gender) are compared to responses to the same items by other students who share a different aspect of the same characteristic (e.g., males vs. females). The group representing students in a specific demographic group is referred to as the *focal* group. The group comprised of students from outside this group is referred to as the *reference* group. When the fundamental measurement assumption does not hold (i.e., students with the same ability in different groups of interest have different probabilities of correctly answering an item), the item is said to be functioning differently for the two groups.

The presence of DIF in an item suggests that the item is functioning unexpectedly regarding the groups included in the comparison. However, the cause of the unexpected functioning is not revealed in a DIF analysis. It may be that item content is inadvertently providing an advantage or disadvantage to members of one of the two groups. Content experts who have special knowledge of the groups involved are often in a good position to identify a cause of this type. DIF may also result from differential instruction closely associated with group membership.

#### 5.1.1. DIF Methods

There are a total of 60 pilot prompts (items) across grades and forms. The Mantel-Haenszel (MH) (1959) procedure was used to detect DIF for these prompts (Holland & Thayer, 1988). The MH method has been widely used in educational measurement due to its easy implementation in testing programs. The procedure compares the ratio of the probabilities of two groups of students (i.e., focal and reference groups) answering an item correctly across all score levels. Table 5.1 presents the four demographic variables analyzed for DIF, along with the focal and reference groups for each variable.

**Table 5.1. Focal and Reference Groups for the DIF Demographic Variables**

| Demographic Variable         | Focal Group                | Reference Group                |
|------------------------------|----------------------------|--------------------------------|
| ELL                          | Former ELLs                | Non-ELLs                       |
| Gender                       | Female                     | Male                           |
| Social Economic Status (SES) | Economically disadvantaged | Not economically disadvantaged |
| Ethnicity                    | Hispanic                   | White                          |

A large enough sample size of students from each demographic group involved in the comparison is required to ensure that each comparison has adequate power to detect DIF. DIF was not conducted if the sample size for either the reference group or the focal group was less than 200 or if the total sample size (reference + focal group) was less than 500. For ELL, the sample size of current ELL students was too small to be included in the DIF analyses, so DIF was conducted by comparing former and non-ELL students. For ethnicity, Hispanic was the only minority ethnicity group with a big enough sample size to be included in the DIF analyses.

The results are categorized based on the Educational Testing Service (ETS)'s method of classifying DIF (Zwick, 2012; Zieky, 1993), as shown in Table 5.2. This method allows items exhibiting negligible DIF (Category A) to be differentiated from those exhibiting moderate DIF (Category B) and strong DIF (Category C). Categories B and C have a further breakdown as “+” (DIF is in favor of the focal group) or “-” (DIF is in favor of the reference group).

**Table 5.2. DIF Categories**

| DIF Category | Level of DIF | Definition  |
|--------------|--------------|---|
| A            | Negligible   | <ul style="list-style-type: none"> <li>Absolute value of the Mantel-Haenszel delta difference (MH D-DIF) is not significantly different from 0 or is less than one.</li> </ul>  |
| B            | Moderate     | <ul style="list-style-type: none"> <li>Absolute value of the MH D-DIF is significantly different from 0 but not from one, and is at least 1; or</li> <li>Absolute value of the MH D-DIF is significantly different from 1, but less than 1.5.</li> <li>Positive values are classified as “B+” and negative values as “B-”.</li> </ul> |
| C            | Strong       | <ul style="list-style-type: none"> <li>Absolute value of the MH D-DIF is significantly different from 1, and is at least 1.5; and</li> <li>Absolute value of the MH D-DIF is larger than 1.96 times the standard error of MH D-DIF.</li> <li>Positive values are classified as “C+” and negative values are “C-”.</li> </ul>          |

### 5.1.2. DIF Results

Table 5.3 presents the summary results of the DIF analyses, including the number of prompts exhibiting DIF by ELL status, gender, social economic status (SES), and ethnicity. Appendix K presents the prompt-level DIF results, including the sample sizes. Prompts with Category C DIF are highlighted. In general, most prompts are classified in Category A. Only a small percentage of prompts are classified as Category C (0–3.3%) or Category B (1.7–8.3%). Only 1–2 prompts exhibit Category C DIF for the ELL, SES, and ethnicity variables, and none exhibit Category C DIF for gender. Among the five prompts that showed strong DIF, four of them favor the reference group. There are 1–5 prompts that exhibit Category B DIF for each demographic variable.

Some hypotheses from content experts about the causes of DIF include the following:

- Low SES students may not have sufficient previous experience with online forms.
- There are difficult words for ELL students in the stem of some prompts.
- Some prompts have a series of facts, images or graphs, and complex structure for students with limited language or computer experience to understand.

**Table 5.3. DIF Summary Results**

| DIF Category | ELL       |              | Gender    |              | SES       |              | Ethnicity |              |
|--------------|-----------|--------------|-----------|--------------|-----------|--------------|-----------|--------------|
|              | #Items    | %            | #Items    | %            | #Items    | %            | #Items    | %            |
| A            | 53        | 88.3         | 58        | 96.7         | 58        | 96.7         | 53        | 88.3         |
| B+           | 3         | 5.0          | –         | –            | –         | –            | 3         | 5.0          |
| B-           | 2         | 3.3          | 2         | 3.3          | 1         | 1.7          | 2         | 3.3          |
| C+           | 1         | 1.7          | –         | –            | –         | –            | –         | –            |
| C-           | 1         | 1.7          | –         | –            | 1         | 1.7          | 2         | 3.3          |
| <b>Total</b> | <b>60</b> | <b>100.0</b> | <b>60</b> | <b>100.0</b> | <b>60</b> | <b>100.0</b> | <b>60</b> | <b>100.0</b> |

## 5.2. Reliability

Reliability refers to the consistency of scores obtained from the assessment. It reflects the absence of random measurement error. When the measurement error is large, reliability is small, and vice versa. Increasing reliability by minimizing error is an important goal for any test. Different sources of measurement error affect scores. The effect of each particular source of error has a corresponding reliability coefficient that describes the influence of that source on scores. One source of measurement error is the items selected for a test. Internal consistency will be high if measurement error due to items is low. Tests are considered of sound reliability when the reliability coefficients range from 0.80 and above. Increasing the test length can increase the reliability to a large extent.

The internal consistency of the pilot test was analyzed by Cronbach’s alpha that indicates how well prompts that reflect the same construct yield similar results. Table 5.4 presents the reliability results. As shown in the table, Cronbach’s alpha varies across forms. Grade 8 Form A, which has the most prompts at 18, has the highest Cronbach’s alpha. The reliability of the other three forms are low. Potential causes of the low reliability include assumptions of Cronbach’s alpha, students’ low motivation for the pilot test, and the relatively short test length compared to a regular state assessment. The estimate of reliability from Cronbach’s alpha is likely an underestimate of the reliability of the pilot test. The use of Cronbach’s alpha to estimate reliability assumes that items are intended to be of equal difficulty, but the pilot prompts were expected to vary in difficulty due to cognitive complexity and other construct-relevant sources.

Based on the test duration data, some students went through the test quickly and may have been rapid guessing (i.e., the minimum test duration of each form is close to 1 minute or less). The medium test duration time was only around 20–25 minutes, which is much shorter than a regular state assessment. Based on the Spring 2018 test duration results for the current NSCAS Science assessment, a regular Nebraska state science assessment takes around 70 minutes or less for students to complete (NDE, 2018, p. 109), which would allow roughly six tasks.

A longer test length generally increases the reliability of an assessment. Therefore, the Spearman-Brown formula was used to determine whether the reliability would increase if the pilot test was longer (i.e., closer in length to the current summative version). Using the Spearman-Brown formula, the pilot test length was increased by three times (e.g., six tasks instead of two and around three times the number of prompts). The results (i.e., the predicted reliabilities) are presented in Table 5.4. The reliabilities of the current NSCAS Science assessments are at 0.8 or above (NDE, 2018, pp. 131–132). Based on the Spearman-Brown results, when the new science assessment is three times longer and when students have higher motivation during the administration, it is likely that the test will meet the reliability target of 0.8.

**Table 5.4. Reliability Results: Cronbach’s Alpha and Predicted Reliability (*r*) based on the Spearman-Brown Formula**

| Grade | Form   | Cronbach’s Alpha | Predicted <i>r</i> based on Spearman-Brown (3x longer test length) |
|-------|--------|------------------|--|
| 5     | Form A | 0.413            | 0.679  |
|       | Form B | 0.511            | 0.758  |
| 8     | Form A | 0.714            | 0.882  |
|       | Form B | 0.569            | 0.798  |

### 5.3. Concurrent Validity

Validity is defined as “the degree to which evidence and theory support the interpretations of test scores for proposed uses. Validity is, therefore, the most fundamental consideration in developing tests and evaluating tests” (AERA, APA, & NCME, 2014, p. 11). It is not a quantifiable property but an ongoing process, beginning at initial conceptualization of the construct, continuing throughout the entire testing process, and extending into the interpretation and use of test scores.

Validity evidence for the pilot is presented as concurrent validity to determine how well scores from the pilot test correspond to the scores obtained from another established and validated test designed to assess the same domain area. Specifically, scores from the 2019 NSCAS Science assessments were used. Concurrent validity requires that both tests are administered to the same students within a short amount of time. The pilot test was administered during the test window from March 4–15, 2019, and the 2019 NSCAS Science test scores were extracted from students who took the test between March 18 and April 5, 2019.

Concurrent validity is expressed in the form of a Pearson correlation coefficient between students’ scores on two tests with a value between 0.0 to 1.0. Table 5.5 presents the correlations between the pilot test scores and students’ total scale scores from the 2019 NSCAS Science assessment. It also presents the correlations by domain score. The correlations between the pilot and the summative test scores are moderate, indicating that students’ pilot test scores and their summative scores are related to some extent but are not highly related as the pilot test was designed to measure slightly different constructs.

**Table 5.5. Concurrent Validity between Pilot Test Scores and 2019 NSCAS Science Scores**

| Grade | Form   | N     | Overall Correlation | Correlation by Domain |          |       |                 |
|-------|--------|-------|---------------------|-----------------------|----------|-------|-----------------|
|       |        |       |                     | Inquiry               | Physical | Life  | Earth and Space |
| 5     | Form A | 736   | 0.486               | 0.373                 | 0.478    | 0.423 | 0.383           |
|       | Form B | 512   | 0.460               | 0.402                 | 0.305    | 0.377 | 0.420           |
| 8     | Form A | 1,050 | 0.636               | 0.562                 | 0.525    | 0.546 | 0.544           |
|       | Form B | 1,001 | 0.596               | 0.507                 | 0.506    | 0.507 | 0.518           |

## References

- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer, and H. I. Brown (Eds.), *Test validity* (pp. 129–145). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719–748.
- National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. Committee on a Conceptual Framework for New K–12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education Washington, DC: The National Academies Press.
- Nebraska Department of Education (NDE). (2017). *Nebraska’s college and career ready standards for science*. Retrieved from [https://cdn.education.ne.gov/wp-content/uploads/2017/10/Nebraska\\_Science\\_Standards\\_Final\\_10\\_23.pdf](https://cdn.education.ne.gov/wp-content/uploads/2017/10/Nebraska_Science_Standards_Final_10_23.pdf).
- Nebraska Department of Education (NDE). (2018). *Spring 2018 NSCAS summative ELA, mathematics, and science technical report*. Retrieved from <https://www.education.ne.gov/assessment/technical-reports/>.
- Zwick, R. (2012). *A review of ETS differential item functioning assessment procedures: Flagging rules, minimum sample size requirements, and criterion refinement* (ETS RR-12-08). Princeton, NJ: ETS.
- Zieky, M. (1993). DIF statistics in test development. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 337–347). Hillsdale, NJ: Erlbaum.

**Appendix A: Task Development Evaluation Results**

A total of 20 people responded.

**Part 1: Feedback on the Meeting**

Read each of the statements below and mark your level of agreement. Please bubble only one of the five options for each statement

| Survey Question  | Strongly Disagree |   | Disagree |      | Neutral |      | Agree |      | Strongly Agree |       |
|--|-------------------|---|----------|------|---------|------|-------|------|----------------|-------|
|  | N                 | % | N        | %    | N       | %    | N     | %    | N              | %     |
| <b>Training</b>  |                   |   |          |      |         |      |       |      |                |       |
| 1. The opening session presentations helped me understand the goals of this meeting.           | –                 | – | –        | –    | 2       | 10.0 | 10    | 50.0 | 8              | 40.0  |
| 2. The training helped me understand the process for writing tasks.                            | –                 | – | –        | –    | 3       | 15.0 | 13    | 65.0 | 4              | 20.0  |
| 3. My facilitator clarified the meeting tasks for me while we worked.                          | –                 | – | –        | –    | 1       | 5.0  | 4     | 20.0 | 15             | 75.0  |
| 4. The training materials were useful.   | –                 | – | –        | –    | 1       | 5.0  | 8     | 40.0 | 11             | 55.0  |
| <b>Process</b>   |                   |   |          |      |         |      |       |      |                |       |
| 5. My group considered Nebraska's state standards when we wrote tasks.                         | –                 | – | –        | –    | –       | –    | –     | –    | 20             | 100.0 |
| 6. My group considered how content is taught across schools and districts.                     | –                 | – | –        | –    | –       | –    | 5     | 25.0 | 15             | 75.0  |
| 7. My group considered cognitive complexity when writing tasks.                                | –                 | – | –        | –    | 4       | 20.0 | 8     | 40.0 | 8              | 40.0  |
| 8. Having the opportunity to make revisions while we worked improved the quality of our items. | –                 | – | –        | –    | –       | –    | 4     | 20.0 | 16             | 80.0  |
| 9. Participating in the session increased my understanding of Nebraska's standards.            | –                 | – | –        | –    | –       | –    | 2     | 10.0 | 18             | 90.0  |
| <b>Overall</b>   |                   |   |          |      |         |      |       |      |                |       |
| 10. Overall, my group valued my opinions.  | –                 | – | –        | –    | 1       | 5.0  | 5     | 25.0 | 14             | 70.0  |
| 11. Communication prior to the meeting prepared me for the work.                               | –                 | – | 2        | 10.0 | 2       | 10.0 | 10    | 50.0 | 6              | 30.0  |
| 12. The facilitator was prepared each day.   | –                 | – | –        | –    | –       | –    | 2     | 10.0 | 18             | 90.0  |
| 13. The facilitator made sure all opinions were heard.   | –                 | – | –        | –    | –       | –    | 3     | 15.0 | 17             | 85.0  |
| 14. The meeting rooms were comfortable and clean.  | –                 | – | 1        | 5.9  | 1       | 5.9  | 5     | 29.4 | 10             | 58.8  |
| 15. The meeting rooms had appropriate materials to complete the work.                          | –                 | – | –        | –    | –       | –    | 4     | 20.0 | 16             | 80.0  |
| 16. The food service met my needs.   | –                 | – | –        | –    | 1       | 5.0  | 4     | 20.0 | 15             | 75.0  |

**Part 2: Information About You**

| Survey Question                               | N  | %     |
|---|----|-------|
| 17. On which content area panel did you work? |    |       |
| English Language Arts                         | –  | –     |
| Mathematics                                   | –  | –     |
| Science                                       | 20 | 100.0 |
| 18. On which grade panel did you serve?       |    |       |
| Grade 3                                       | –  | –     |
| Grade 4                                       | –  | –     |
| Grade 5                                       | 16 | 80.0  |
| Grade 6                                       | 1  | 5.0   |
| Grade 7                                       | 1  | 5.0   |
| Grade 8                                       | 14 | 70.0  |
| 19. Have you served on a committee before?    |    |       |
| Yes   | 11 | 57.9  |
| No  | 8  | 42.1  |
| 20. What grade do you teach?                  |    |       |
| 3rd grade                                     | 2  | 10.0  |
| 4th grade                                     | 1  | 5.0   |
| 5th grade                                     | 5  | 25.0  |
| 6th grade                                     | 3  | 15.0  |
| 7th grade                                     | 5  | 25.0  |
| 8th grade                                     | 6  | 30.0  |
| Other grades                                  | 1  | 5.0   |
| N/A   | 3  | 15.0  |
| 21. What content area do you teach?           |    |       |
| English Language Arts                         | 4  | 20.0  |
| Mathematics                                   | 4  | 20.0  |
| Science                                       | 19 | 95.0  |
| N/A   | 1  | 5.0   |

**Part 3: Your Turn to Elaborate**

22. Share comments or additional feedback that will help us understand your perspective about the meeting. Thank you!

- Great experience!
- Dairy free was not always honored (butter = dairy!).
- I learned a lot. This is a valuable first step in preparation for the new standards.
- The experience this week was painfully useful and valuable. The struggle through the process created invaluable insight into the method of the process and the need to be thorough.
- This week was an amazing learning experience for me! I would love to return for more work/learning! Sara is amazing! Food service did not meet my needs on pasta day.
- I feel as though we were experiencing the assessment process while creating. It would've been helpful to review the process a bit more before diving into our work.
- The slide shared on Friday that represented the "structure" of our work was helpful. I recommend showing it sooner for future work. We also should have had the task format options from NWEA sooner. Thanks!
- The food was awesome!

**Appendix B: Task Review Evaluation Results**

A total of 6 people responded.

**Part 1: Feedback on the Meeting**

Read each of the statements below and mark your level of agreement. Please bubble only one of the five options for each statement

| Survey Question   | Strongly Disagree |      | Disagree |      | Neutral |      | Agree |      | Strongly Agree |      |
|---|-------------------|------|----------|------|---------|------|-------|------|----------------|------|
|   | N                 | %    | N        | %    | N       | %    | N     | %    | N              | %    |
| <b>Training</b>   |                   |      |          |      |         |      |       |      |                |      |
| 23. The opening session presentations helped me understand the goals of this meeting.           | -                 | -    | -        | -    | -       | -    | 3     | 50.0 | 3              | 50.0 |
| 24. The training helped me understand the process for reviewing items.                          | -                 | -    | -        | -    | -       | -    | 3     | 50.0 | 3              | 50.0 |
| 25. My facilitator clarified the meeting tasks for me while we worked.                          | -                 | -    | -        | -    | -       | -    | 2     | 33.3 | 4              | 66.7 |
| 26. The training materials were useful.   | -                 | -    | -        | -    | -       | -    | 4     | 66.7 | 2              | 33.3 |
| <b>Process</b>  |                   |      |          |      |         |      |       |      |                |      |
| 27. My group considered Nebraska's state standards when we discussed items.                     | -                 | -    | -        | -    | -       | -    | 1     | 16.7 | 5              | 83.3 |
| 28. My group considered how content is taught across schools and districts.                     | -                 | -    | 1        | 16.7 | -       | -    | 3     | 50.0 | 2              | 33.3 |
| 29. My group considered cognitive complexity when reviewing items.                              | -                 | -    | -        | -    | 1       | 16.7 | 4     | 66.7 | 1              | 16.7 |
| 30. Having the opportunity to make revisions while we worked improved the quality of our items. | -                 | -    | -        | -    | -       | -    | 4     | 66.7 | 2              | 33.3 |
| 31. Participating in the review increased my understanding of Nebraska's standards.             | -                 | -    | -        | -    | -       | -    | 2     | 33.3 | 4              | 66.7 |
| <b>Overall</b>  |                   |      |          |      |         |      |       |      |                |      |
| 32. Overall, my group valued my opinions.   | -                 | -    | -        | -    | -       | -    | 3     | 50.0 | 3              | 50.0 |
| 33. Communication prior to the meeting prepared me for the work.                                | -                 | -    | -        | -    | 5       | 83.3 | -     | -    | 1              | 16.7 |
| 34. The facilitator was prepared each day.  | -                 | -    | -        | -    | -       | -    | 1     | 16.7 | 5              | 83.3 |
| 35. The facilitator made sure all opinions were heard.  | -                 | -    | -        | -    | -       | -    | 2     | 33.3 | 4              | 66.7 |
| 36. The meeting rooms were comfortable and clean.   | -                 | -    | -        | -    | -       | -    | 2     | 33.3 | 4              | 66.7 |
| 37. The meeting rooms had appropriate materials to complete the work.                           | -                 | -    | -        | -    | -       | -    | 1     | 16.7 | 5              | 83.3 |
| 38. The food service met my needs.  | 1                 | 16.7 | -        | -    | -       | -    | 1     | 16.7 | 4              | 66.7 |

**Part 2: Information About You**

| Survey Question                               | N | %     |
|---|---|-------|
| 39. On which content area panel did you work? |   |       |
| ELA   | – | –     |
| Mathematics                                   | – | –     |
| Science                                       | 6 | 100.0 |
| 40. On which grade panel did you serve?       |   |       |
| Gr 3 item                                     | – | –     |
| Gr 4 item                                     | – | –     |
| Gr 5 item                                     | 2 | 33.3  |
| Gr 6 item                                     | – | –     |
| Gr 7 item                                     | – | –     |
| Gr 8 item                                     | 4 | 66.7  |
| Bias  | 1 | 16.7  |
| 41. Have you served on a committee before?    |   |       |
| Yes   | 4 | 66.7  |
| No  | 2 | 33.3  |
| 42. What grade do you teach?                  |   |       |
| 3rd grade                                     | – | –     |
| 4th grade                                     | – | –     |
| 5th grade                                     | 1 | 16.7  |
| 6th grade                                     | – | –     |
| 7th grade                                     | 1 | 16.7  |
| 8th grade                                     | 2 | 33.3  |
| Other   | 4 | 66.7  |
| 43. What content area do you teach?           |   |       |
| ELA   | 1 | 16.7  |
| Mathematics                                   | 1 | 16.7  |
| Science                                       | 6 | 100.0 |

**Part 3: Your Turn to Elaborate**

44. Share comments or additional feedback that will help us understand your perspective about the meeting. Thank you!

- Three stars next to question #9: Participants in the review increased my understanding of Nebraska's standards. Two stars next to question #16: The food service met my needs.
- Eggs left out likely caused food poisoning.
- Dynamic presentation helped me to better understand how to write, how to evaluate, and how to discuss with others. NCCRSS
- Note on question #19: I worked on a committee with STARS. This information is not being disseminated very well to classroom teachers. Maybe this needs work through ESU level.

**Appendix C: Pilot Test Administration Manual**



**TEST ADMINISTRATION MANUAL**

Spring 2019 Science Pilot

2018–2019

### Contributors

Nebraska Student-Centered Assessment System General Summative assessments are administered by the Nebraska Department of Education (NDE):

301 Centennial Mall South  
Lincoln, Nebraska 68509

(402) 471-2495

The assessment contractor is NWEA. NWEA can be reached by calling Customer Service at (855) 225-9926 or by emailing [NWEANebraska@nwea.org](mailto:NWEANebraska@nwea.org).

Copyright © 2019 by the Nebraska Department of Education. No part of this publication may be reproduced, copied or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the Nebraska Department of Education.

## Table of Contents

---

### Contents

|  |    |
|--|----|
| Part 1—General Overview .....                  | 4  |
| Test Administration Key Dates .....            | 4  |
| Student Participation in the Test .....        | 4  |
| Scheduling the Test .....                      | 5  |
| NSCAS Security .....                           | 5  |
| Part 2—Administering the Science Pilot .....   | 5  |
| Browser Requirements .....                     | 5  |
| Preparing for the Science Pilot .....          | 6  |
| Administering the Science Tests .....          | 6  |
| General Test Settings .....                    | 6  |
| Proctor Responsibilities .....                 | 7  |
| Part 3—Test Administration Instructions .....  | 8  |
| Additional Questions .....                     | 11 |
| Appendix—The Do's and Don'ts of Security ..... | 12 |

## Part 1—General Overview

In order to gain feedback from Nebraska districts on newly developed science tasks, NWEA and NDE worked collaboratively to develop science tasks in grades 5 and 8 to pilot with Nebraska districts. This pilot is intended to glean more meaningful information directly from students to inform NWEA and NDE on assessment development for summer 2019. It is also intended to expose these new science questions to Nebraska districts to prepare them for future administrations and to allow them the opportunity to provide feedback on the new question types. NDE and NWEA recognize the potential impact this pilot toward meaningfully measuring scientific literacy.

This pilot is not intended to test student skills or knowledge, and students' scores will not be reported. This test may not be representative of what will appear on a future assessment. Student responses will be reviewed by NDE.

The Science Pilot tests are comprised of questions written and/or reviewed by Nebraska educators in collaboration with NDE. These tests will be administered through Qualtrics, an online survey tool.

This *Test Administration Manual* has been designed to help districts and schools administer the pilot tests accurately and efficiently. Please take the time to read this manual to become familiar with the administration of the assessments.

### Test Administration Key Dates

| Date              | Activity                              |
|-------------------|---------------------------------------|
| February 25, 2019 | Attend Science Pilot virtual training |
| March 4–15, 2019  | Pilot test window                     |

### Student Participation in the Test

The purpose of the Science Pilot assessment is to test new assessments designed to measure three dimensional science learning. Participation is optional and at the discretion of each district.

There are a set of student feedback questions after each task. Students can rate their interest and the difficulty of the task as well as how well they understood how to respond to each question of the task.

In addition to gathering student feedback on the tasks, select educators will also be asked to conduct a Cognitive Lab exercise to understand individual students' thought processes when answering the science questions. In each cognitive lab, an educator will interact with a student during the assessment, ask specific questions, and report the student's responses. Separate training and directions will be provided to those selected educators.

## Scheduling the Test

The Science Pilot assessment is untimed and designed to provide students with as much time as needed to complete; however, the estimated test-taking time is no more than 45 minutes to complete the two tasks of each test form. The test should be taken in one sitting.

This pilot will take place March 4 through March 15, 2019.

## NSCAS Security

In a centralized testing process, it is critical that equity of opportunity, standardization of procedures, and fairness to students is maintained. Although this Science Pilot is not considered secure and tasks will be released to the public, it is recommended that educators follow standard practices in test security as they would during a standardized assessment.

It is important to note that all teachers need to be familiar with appropriate testing ethics and security practices related to testing. Professionalism, common sense, and practical procedures will provide the right framework for testing ethics. The *NSCAS Security Manual* is intended to outline clear practices for appropriate security.

## Part 2—Administering the Science Pilot

### Browser Requirements

Qualtrics operates on supported versions of the following browsers:

- Apple Safari®
- Google Chrome™
- Microsoft Edge
- Microsoft Internet Explorer
- Mozilla Firefox

Versions of these browsers no longer supported by their maker may not function as intended and may expose the user to inherent security vulnerabilities. Qualtrics recommends that users always use the most up-to-date version of their browser. For more information on the Qualtrics platform and technical support, please visit:

<https://www.qualtrics.com/support/survey-platform/getting-started/help-and-feedback/#LoginBrowserCompatibility>

## Preparing for the Science Pilot

Posted to Qualtrics is one released task for each grade. Educators and students can use the released tasks for students to experience what these new, innovative questions types are like and what the pilot will look like.

Grade 5 – Changing Coral: <http://bit.ly/prepilot5>

Grade 8 – Blue Skin: <http://bit.ly/prepilot8>

Please take note of the question types that are used in this pilot:

- Multiple Choice – allows the student to choose one or multiple options from a list of possible answers
- Text Entry – allows the student to respond to open-ended questions
- Pick, Group, and Rank (drag and drop) – allows students to drag and drop items into groups. Within each group, respondents can rank items by dragging and dropping them into place.
- Hot Spot – the student is presented with an image or text that has predefined regions to select from.

## Administering the Science Tests

Although not a standardized assessment, Proctors should follow the same procedures when administering the Science Pilot. Those directions can be found in [Part 3—Test Administration Instructions](#).

There will be two unique test forms per grade. Each test form will contain two tasks. Students will only be expected to complete one of the test forms. The URL to these test forms will be provided by NDE. These test forms should be assigned to students at random. It is recommended that for any group of students in either grade 5 or 8 that will participate in the pilot, half should be assigned one test form in the respective grade and the other half should be assigned the other test form.

## General Test Settings

As noted, participation in the Science Pilot is optional. There are no standard embedded accommodations/accessibility supports provided through the Qualtrics platform, but educators should assist any student who needs accommodations as described in the [NSCAS General Summative & Alternate Accessibility Manual](#).

Students may use approved non-embedded resources, such as multiplication charts or noise

buffers, as specified by NDE policy. A complete list of non-embedded universal tools, linguistic supports, and accommodations is included in the [NSCAS General Summative & Alternate Accessibility Manual](#).

Take note of the following important available test settings that should be used in accordance with the [NSCAS General Summative & Alternate Accessibility Manual](#):

- Zoom/magnifier – students should use the native device zoom feature to magnify the content on the page.
- Text-to-Speech – students should use native screen readers or read aloud in accordance with the NDE accessibility manual.
- Calculator – students may use a four-function calculator to answer questions in grade 8. Students in grade 8 must be provided access to a calculator. Students may not share calculators during the assessment. The calculator function of a smartphone may not be used during a test.

## Proctor Responsibilities

Some of the major tasks Proctors are responsible for include:

- Arranging the testing room
- Ensuring that all students have scratch paper
- Ensuring that all students have a basic, four-function calculator
- Restricting electronic devices of any type (smart phones, cell phones, PDAs, wristwatches with electronic displays, calculators, iPods®, MP3 players, etc.). These devices should be placed into secure storage, or into student backpacks, and kept elsewhere throughout the test.
- Prohibiting talking or sharing of responses with other students
- Adhering to all security requirements

Prepare students for testing by informing them of the scheduled tests in advance. Explain to the students why they are being tested and how the results will be used. Students can sense the importance the Proctor places on the tests, and their performance may be affected accordingly. Students should realize that doing their best is important.

Proctors will need the 10-digit State Student ID for students participating in this pilot. This should be provided to each individual student at the time of testing, preferably on individual cards for each student that can be collected at the end of the test and securely destroyed.

Good organization of test materials and well-executed procedures will help the administration proceed smoothly. The following guidelines are recommended:

- Provide a testing location that has comfortable seating, sufficient workspace, and good lighting.
- Arrangements for rooms and seating should be announced in advance in order to eliminate confusion when testing begins. The room where students take the test should be as free from outside disturbance as possible.
- A “Testing in Progress” sign should be posted on the door to prevent interruptions.

### Part 3—Test Administration Instructions

Administration of the Science Pilot is an important professional responsibility. The usefulness of the test results depends on the accuracy of each student’s performance. Experience shows that student performance is highly dependent upon the student’s motivation and attitude toward the test, the preparedness of the Proctor, the physical arrangements for testing, and adherence to instructions. To ensure accurate and reliable results, the Proctor **must** become thoroughly familiar with the procedures described in this manual before administering the test.

When the test session is started, follow the script below.

In order to ensure accurate achievement results, it is essential that all Proctors follow the same procedures when administering the tests.

#### Instructions for Science Pilot

Read aloud **word for word** the material that is printed in **bold type** and preceded by the word “SAY.”

*The material that is italicized is information for you and should not be read to the students.*

Read the directions to the students exactly as they are written using a natural tone and manner. If you make a mistake in reading a direction, stop and say, “No, that is wrong. Listen again.” Then read the direction again. Be sure students understand the directions and how to respond. Be careful not to inadvertently give hints or clues that indicate an answer. Begin the test when all students are present.

**SAY: Today you will take the Science Pilot. Give this test your best effort. There will be no scores reported. Some questions will be easy; others will be more difficult. Be sure to read the entire question and all of the answer choices carefully. If you do not know the answer to a question, use your best test-taking skills to eliminate some answer choices. Then choose the best answer of the ones that remain. Remember there is no penalty for guessing.**

**You should answer each question before you move on to the next question. Once you move to the next page, you will not be able to go back.**

**You may use scratch paper on this test.**

*Read to grade 8:*

- You may use a calculator on this test.

The test is not timed, and you will be given a reasonable amount of time to finish.

If you need help once the test has started, raise your hand and the Proctor (teacher) will come to you. The Proctor (teacher) is not allowed to provide you with any additional information during the test. The Proctor (teacher) cannot help you with any words.

Are there any questions?

*Answer all questions. When all students are ready, continue.*

**SAY:** We will begin by accessing the test site.

*Provide students with the URL to the form provided by NDE. Assist students as needed.*

**SAY:** Enter your 10-Digit State Student ID in the text box.

*Assist students as needed.*

**SAY:** This test will include several different types of questions. Multiple choice questions will ask you to select an answer usually from among four choices. Multiple select questions will ask you to select multiple correct answers usually from among five or more choices. Some questions will ask you to construct an answer by following the directions given. For these enhanced questions, follow the instructions in the question. You may use information provided in any part of the task to answer the questions.

- On a page, you should answer each question before you move on to the next question. Once you move to the next page, you will not be able to go back.
- If you need to leave your computer, raise your hand and ask the Proctor. Students may take breaks at the discretion of the proctor. The test must be completed in one sitting.
- At the end of each task, there will be a list of questions where you can provide your feedback to the task. Please answer those to the best of your ability.
- This test is not timed. You will know you are finished when you see a screen that says, "Congratulations, you have finished the test."

When you come to the end of the session, please sit quietly or read until the

**Proctor provides additional instructions.**

**Are there any questions?**

*Answer all questions. When all students are ready, continue.*

**SAY: When you are ready, you may start the test.**

*Text-to-Speech is available using your device's native read aloud feature but should only be enabled for students with a documented need, such as an English Learner or a student on an IEP or 504 plan, to be consistent with the requirements for use in the NSCAS General Summative assessment.*

*While students are working, walk around the room to see that they are following directions. Although these questions are not considered secure, educators and students should follow the same guidance and policies for maintaining test security.*

## Additional Questions

If you have any additional questions not covered in this manual, please contact an NWEA Nebraska Customer Service Representative by phone at (855) 225-9926 or by email at [NWEANebraska@nwea.org](mailto:NWEANebraska@nwea.org).

Customer Service Representatives will make every attempt to answer your questions and escalate the issue when appropriate. Emailed support requests are handled within 24 hours. All contacts to the support team are tracked and documented. Representatives are available from 7:00 a.m. to 5:00 p.m. Central Time, Monday through Friday.

## Appendix—The Do’s and Don’ts of Security

| Do's   | Don't's  |
|--|--|
| <p>Do control access to all cell phones and personal electronic devices.</p>   | <p>Do not discuss, disseminate, or otherwise reveal the contents of the test to anyone.</p>  |
| <p>Do attend any district or school training for the administration of the test in order to be properly informed of the procedures to follow, including securing test materials.</p> | <p>Do not keep, copy, reproduce, or use any reading, mathematics, or science test, test item, any specific test content, or examine responses to an item or any section of a secured test in any manner inconsistent with the instructions provided by and through the Nebraska Department of Education.</p> |
| <p>Do move around the testing site to ensure students are adhering to the instructions given.</p>  | <p>Do not leave students unattended with testing materials.</p>  |
| <p>Do collect scratch paper and return it to the School Test Coordinator.</p>  | <p>Do not possess any secure test materials at any time other than during the actual administration of the test. Proctors should be given their secure materials the morning of the administration of the test, and materials must be counted and collected at the end of each day of testing.</p>           |
| <p>Do follow appropriate accommodation procedures as found in the NSCAS <i>General Summative and Alternate Accessibility Manual</i>.</p>   | <p>Do not allow students to leave the testing site with test materials for any reason.</p>   |
| <p>Do make students feel comfortable and relaxed.</p>  | <p>Do not allow students to look ahead to other content areas before being instructed to do so.</p>  |
| <p>Do escort all students and carry all secure testing materials to alternate sites for extended time, etc.</p>  | <p>Do not coach or provide feedback in any way, which includes answering any questions relating to the contents.</p>   |
| <p>Do have test booklets or test tickets/online set-up ready for students ahead of time.</p>   | <p>Do not alter, influence, or interfere with a test response in any way or instruct the student to do so. Students who move to alternate testing sites for an extended time should be escorted, and school personnel should carry all secure testing materials to the new testing location.</p>             |
| <p>Do remove from the wall all curriculum materials that relate to the tested content.</p>   | <p>Do not complete any unanswered item or provide actual answers to students.</p>  |
| <p>Do maintain standardized testing procedures.</p>  | <p>Do not place students in situations in which they can discuss test items or answers, such as during a break.</p>  |
| <p>Do adhere to ACT’s regulations for students with extended time.</p>   |  |
| <p>Do read the NSCAS <i>Security Manual</i> and all applicable test administration manuals before testing.</p>   |  |

## Appendix D: Cognitive Lab Parent Notification Letter

February 19, 2019

Dear Parent or Guardian:

The Nebraska Department of Education (NDE) will be engaging a small number of students to participate in cognitive labs to collect more detailed information about potential science tasks for a future state science assessment. Your student is among the children selected to participate in a cognitive lab taking place March 4–15, 2019.

In cognitive labs, students tell educators what they are thinking when answering test questions. Students also give their perspective on the items. An educator will interact with your student as he or she works through two science tasks and the associated test questions. Your student will be asked to think aloud as he or she responds. The educator will guide your student through the process and document his or her responses. The appendix included with this letter presents an example of the cognitive lab directions and questions that will be used by the educator. The cognitive lab will take no longer than 1.5 hours for most students.

This is a research study. The information collected will be used to help NDE write interesting and engaging test questions for students. The cognitive labs are not intended to test students' skills or knowledge, and student scores will not be reported. Your student's grades will *not* be affected. Student names will not be collected, and student responses will be completely confidential. Only combined group responses will be reviewed by NDE. Participation is voluntary. **Please sign this letter below with your consent and return to me by March 1, 2019.**

There is no need to study in preparation for the cognitive lab, but please encourage your student to do their best. NDE depends on student responses to collect information to inform future science assessment design and development.

We are excited that our school is participating in this research study and appreciate your support!

Sincerely,  
School Principal

\_\_\_\_\_  
Parent/Guardian Printed Name

\_\_\_\_\_  
Parent/Guardian Signature

**Appendix: Cognitive Lab Example**

State Student ID

---

---

Ask the student to read the introduction text only out loud.

After they have finished ask the student:

- A) Did you understand the text?
- B) What are you wondering after reading the passage?
- C) What science knowledge might you need to understand changing coral?

Record their answers and any observations about the student's interaction with the scenario.

---

---

---

---

---

Ask the student to read aloud and answer Question 1.

Ask the student to describe what they are thinking as they read and answer the question. If the student is not talking, you may ask "What are you thinking?" or "Why did you answer that way?" or "What would help you answer this question?".

---

---

---

---

---

The student's comments provided evidence that the item elicited thinking about:

- Human activities have an impact on Earth's ecosystems.
- Evidence that supports a particular claim.
- Other
- None of the above

---

If "other" what thinking did the item elicit?

---

---

---

---

---

While reading and interacting with the question, the student appeared to...

- rush to answer the question without careful consideration of the question, the stimuli, and how to respond.
- be engaged with the question, taking time to consider and respond.
- be able to understand how to generate a response to the question.

---

Ask the student:

Was it easy or hard to know how to respond to the question?

- Easy
  - Medium
  - Hard
-

What would help you answer the question?

---

---

---

---

Ask the student:

Was it easy or hard to use the given information to answer the question?

- Easy
- Medium
- Hard

Ask the student:

Why was it easy or hard?

---

---

---

---

Record any additional notes, comments, or observations made about the student's interaction with this question.

---

---

---

---

Ask the student:  
What are the science tests like in your class?

Record their response.

---

---

---

---

Ask the student:  
How does that compare with what you just did?

Record their response.

---

---

---

---

Record any additional notes, comments, or observations made about the student's interaction with this question.

---

---

---

---

## Appendix E: Cognitive Lab Protocol



### **NSCAS Science Cognitive Lab Protocol** February 2019



© 2019 Nebraska Department of Education. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

**Table of Contents**

1. Cognitive Lab Overview .....1

2. Cognitive Lab Protocol .....2

    2.1. Preparation .....2

    2.2. Logistics .....2

    2.3. Cognitive Lab Process .....2

        Step 1: General Introduction .....2

        Step 2: Think Aloud Practice .....3

        Step 3: Conduct the Cognitive Interview.....5

        Step 4: Concluding Questions .....5

        Step 5: Repeat Steps 3 and 4 for the Second Task.....6

    2.4. Closing.....6

Appendix A: Links to the Pilot Test and Data Collection Tool .....7

Appendix B: Data Collection Tool Example .....8

## 1. Cognitive Lab Overview

Thank you for assisting the Nebraska Department of Education by conducting cognitive labs with individual students for the Nebraska Student-Centered Assessment System (NSCAS) Science assessment from March 4–15, 2019. The purpose of conducting these cognitive labs is to collect detailed information about the process students take when answering test questions to inform future test development for the NSCAS Science assessment.

During each cognitive lab, you will interact with one student as he/she work through two science tasks and his/her associated test questions. The student will respond to the questions online, and their responses will be recorded in the system. You will need to encourage the student to think aloud while she/he completes each question. While the student thinks aloud, you will document your observations of her/his thought processes and other feedback from the student in the data collection tool.

Below is an overview of the cognitive lab process:

- Step 1. Introduce yourself to the student and explain the purpose of the cognitive lab.
- Step 2. Explain the expectations for thinking aloud when answering questions and practice with the student.
- Step 3. Begin the cognitive interview by letting the student read the scenario aloud and answer the first test question associated with the first science task. Encourage the student to think aloud by asking questions such as "How did you decide this was the answer? What evidence from the scenario did you use to decide the answer?" If the student is not talking, you may want to ask her/him to point to the evidence or evidences she/he used in the scenario to draw the conclusion. Document the student's responses in the data collection tool accessed through Qualtrics. Repeat these steps until the student finishes all the questions for the first task.
- Step 4. Each science task includes a few questions at the end to determine how well the student engaged with the task. The student will respond to those questions, and you will ask additional follow-up questions provided in the data collection tool and record his/her responses.
- Step 5. Repeat Steps 3 and 4 for the second science task.

Below is a list of things to remember when conducting the cognitive lab:

- DO NOT tell the student if his/her answer is right or wrong.
- DO NOT ask questions that lead to a student's response.
- DO NOT interrupt the student to ask why she/he is responding in a particular way. Only prompt if she/he are not thinking aloud.
- Encourage the student to be more descriptive if he/she is not talking much, but do not put words in his/her mouth.
- Sit next to the student but not in his/her personal space.

## 2. Cognitive Lab Protocol

### 2.1. Preparation

Before conducting the cognitive labs, please take the following preparation steps:

- Review the pilot test to familiarize yourself with the test questions. The pilot test can be accessed from the first link in Appendix A. You are welcome to answer the test questions, but your responses will not be scored.
- Review the cognitive lab data collection tool on Qualtrics. It can be accessed from the second link in Appendix A.
- Read this entire document before engaging with the student.
- Make sure both you and the student each have a laptop or tablet with a keyboard for use during the cognitive lab.

### 2.2. Logistics

Please adhere to the following logistics when conducting the cognitive labs:

- **Setting:** Each cognitive lab is conducted in a one-on-one setting rather than in a group. Make sure you interact with only one student during the cognitive lab.
- **Seating:** Follow the school's procedures to escort the student from her/his classroom to an empty testing room or a corner of a quiet room.
- **Materials:**
  - This cognitive lab protocol.
  - Laptop or tablet with a keyboard for both yourself and the student
  - Data collection tool (log in to your computer and have the data collection tool on Qualtrics ready to record your observations)
  - Pilot test form (let the student log in to her/his computer and access the form)

### 2.3. Cognitive Lab Process

During each cognitive lab, please follow Steps 1–5 described below.

#### Step 1: General Introduction

---

- a. Introduce yourself and the purpose of the cognitive lab. For example, you can say:

*Hello! My name is (YOUR NAME). I'm (YOUR ROLE). Thank you for helping me today. Today you will be participating in a special study. Its goal is to find better ways to design science tests for all students. We need your help creating a science test.*

- b. Inform the student that his/her responses will be kept confidential and will not be graded. For example, you can say:

*You will not be asked to write your name on any of the work you do, and no one in the school will see your answers. The information you provide will be kept confidential. Even though this is called a test, your responses will not be graded, and this test will not affect your grade in your class. We are really interested in figuring out how we can make the test questions better for students.*

- c. Answer any questions the student may have. For example, you can say:

*Do you have any questions before we begin?*

### **Step 2: Think Aloud Practice**

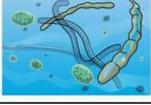
---

- a. Explain what it means to think aloud. For example, you can say:

*For this special study, I am interested in capturing what you notice when you are thinking as you respond to questions. The best way to find out what you are thinking is to have you read and talk out loud while you are working on the test. As you answer the questions, I would like you to say out loud things that you would normally think to yourself.*

- b. Practice thinking aloud with the student. For example, you can say:

To give you some practice with thinking out loud, I'd like you to think out loud while responding to a science question. I'll do one first. I'm going to answer this question.

| Organism    | Picture   | How it obtains energy |
|-------------|---|-----------------------|
| Dolphin     |    | eats fish             |
| Triggerfish |    | eats coral            |
| Coral       |    | eats algae            |
| Algae       |    | from the Sun          |
| Bacteria    |  | from dead organisms   |

Use the information given about the organisms in the ocean ecosystem to predict the organism that will be MOST affected when sediment in the ocean increases.

- A) dolphin
- B) triggerfish
- C) coral
- D) algae
- E) bacteria

*First, I carefully read the information and question, then I read through all the answer options. I think about sediments increasing in the ocean and how those could affect the organisms listed. I think sediments block sunlight and could cover organisms. I see that algae needs the Sun for energy, so I select that organism. I think about how the other organisms but I think the algae will be affected the most.*

*Now tell me how you would answer this same question.*

- c. Continue to practice with the student until she/he understands how to think aloud. Answer any questions the student may have.

### **Step 3: Conduct the Cognitive Interview**

---

- a. Let the student begin the test by reading aloud the scenario associated with the first science task. After she/he finishes, ask the student some follow-up questions and record her/he answers in the data collection tool, along with any observations you made about the student's interaction with the scenario. An example of the follow-up questions is provided in Appendix B.
- b. Ask the student to read aloud the first test question associated with the scenario.
- c. If the student does not understand the question, ask "What do you need to help you understand?"
- d. Let the student speak aloud all her/his thoughts while solving the question and responding online.
- e. If the student is not talking much, ask some follow-up questions to make the data collection as complete as possible. Examples of these questions are provided in the Appendix B.

Below are things to keep in mind during the data collection:

- Do not ask questions that lead to a student's response.
- Be as objective and non-biased as possible.
- Focus on the task at hand (the particular item or task). Do not ask the student a question in general terms.
- Ask the student to explain why the answer options are not right if he/she does not provide that information.
- Capture the frustration aspect of taking the test (and performing the think aloud exercise) if a student seems hung up with something.

- f. Record your observations of the student's responses and interaction with the test question in the data collection tool on Qualtrics. Examples of the questions in the data collection tool to record your observations are provided in Appendix B. Please record your observations in the corresponding fields.
- g. Repeat these steps until the student finishes all the questions for the first task.

### **Step 4: Concluding Questions**

---

- a. After the student finishes all the questions of a task, he/she will respond to a few follow-up **Student Feedback** questions online to determine if she/he thought the task was interesting and if he/she could understand the task.

- b. Based on the student's ratings in the **Student Feedback** section, ask follow-up questions such as "Why was the question interesting (or not interesting)?" and "Why was the question hard to understand?" Record the student's responses in the data collection tool. Examples of these questions are provided in Appendix B.

**Step 5: Repeat Steps 3 and 4 for the Second Task**

---

- a. Repeat Steps 3 and 4 for the second science task, recording all observations in the data collection tool.

**2.4. Closing**

Thank the student for her/his hard work. Dismiss the student according to school procedures. Answer any questions the student may have.

If a student does not finish in allotted time, the student should proceed through the remaining pages and select submit.

Appendix A: Links to the Pilot Test and Data Collection Tool

**Appendix A: Links to the Pilot Test and Data Collection Tool**

Grade 5 Form B: [https://nde.qualtrics.com/jfe/form/SV\\_5szG93Jzpzp5cciF](https://nde.qualtrics.com/jfe/form/SV_5szG93Jzpzp5cciF)

Grade 8 Form A: [https://nde.qualtrics.com/jfe/form/SV\\_cFUC658Tvl1hIMJ](https://nde.qualtrics.com/jfe/form/SV_cFUC658Tvl1hIMJ)

Appendix B: Data Collection Tool Example

**Appendix B: Data Collection Tool Example**

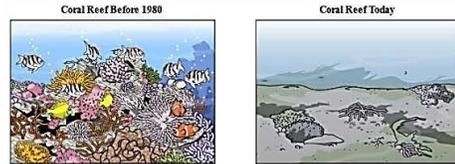
This appendix presents sample questions included in the data collection tool for a scenario and one question.

**Grade 5 or 8, Test Form A or B**

Q1: Student State ID: \_\_\_\_\_

Ask the student to read the scenario out loud.

Scientists have been investigating coral reef ecosystems over time. They have observed many changes. In some areas, corals reefs have disappeared. The scientists are wondering what is causing these changes.



**Information about Coral Reefs**

- Coral reefs are made from animals that live in the same place their entire lives.
- Corals need shallow, clear, warm water to survive.
- Most corals live with algae, which are tiny, plant-like organisms that need sunlight to grow.

Q2: After the student has finished ask the student:

- Did you understand the text?
- What are you wondering after reading the passage?
- What science knowledge might you need to understand changing coral?

Record their answers and any observations about the student's interaction with the scenario.

---

---

---

---

---

Appendix B: Data Collection Tool Example

Ask the student to read aloud and answer **Question 1**.

**Question 1**

Some scientists claim that there is a land-based cause for this ocean problem. Which evidence would support their claim?

- A) The ocean today is warmer than it was in 1980.
- B) The ocean today has more algae than it did in 1980.
- C) There is more oxygen in the ocean today than there was in 1980.
- D) There is more soil and sediment in the ocean today than there was in 1980.

Q3: Ask the student to describe what they are thinking as they read and answer the question. If the student is not talking, you may ask "What are you thinking?" or "Why did you answer that way?" or "What would help you answer this question?"

---



---



---

Q4: The student's comments provided evidence that the question elicited thinking about :

- Human activities have an impact of Earth's ecosystems.
- Evidence that supports a particular claim.
- Other

If "other" what thinking did the item elicit?

---



---

Q5: While reading and interacting with the question, the student appeared to ...

- rush to answer the question without careful consideration of the question, the stimuli, and how to respond.
- be engaged with the question, taking time to consider and respond.
- be able to understand how to generate a response to the question.

Q6: Ask the student, "Was the question easy or hard to know how to respond to the question?"

- Easy
- Medium
- Hard

Q7: Ask the student, "What would help you answer the question?"

---



---

Appendix B: Data Collection Tool Example

Q8: Ask the student, "Was it easy or hard to use the given information to answer the question?"

- Easy
- Medium
- Hard

Q9: Ask the student, "Why was it easy or hard?"

---

---

---

Q10: Record any additional notes, comments, or observations made about the student's interaction with this question.

---

---

---

*Note: Q3-10 is repeated for each of the 5-8 questions in the task. After the student completes these, there are 3 questions about the whole task followed by the Student Feedback section.*

Q: Ask the student, "Tell me what you learned about changing coral or why coral reefs are changing."

---

---

---

Q: Ask the student, "What are the science tests like in your class?"

---

---

---

Q: Ask the student, "How does that compare with what you just did?"

---

---

---

Q: Ask the student to read aloud and answer the first question in the **Student Feedback** section.

**Student Feedback**

**Feedback Question #1**

How interesting was the task you just completed?

1 star is not interesting and 5 stars is very interesting.



Appendix B: Data Collection Tool Example

Q: Why did you rate your interest in the task with 1, 2, 3, 4, or 5 stars?

---

---

---

Ask the student to read aloud and answer the second question in the **Student Feedback** section.

**Feedback Question #2**

How difficult was the task as a whole?

1 star is not difficult and 5 stars is very difficult.



Q: Why did you rate the difficulty of the whole task with 1, 2, 3, 4, or 5 stars?

---

---

---

Q: Ask the student to read aloud and answer the 3<sup>rd</sup> and final question in the **Student Feedback** section.

**Feedback Question #3**

Please rate the difficulty **to understand how to respond to each question.**

1 star is not difficult and 5 stars is very difficult.

- Question 1 ★★★★★
- Question 2 ★★★★★
- Question 3 ★★★★★
- Question 4 ★★★★★
- Question 5 ★★★★★

Q: Why did you rate the difficulty of the questions with 1, 2, 3, 4, or 5 stars?

---

---

---

Q: If needed, ask the student to click the go-to-next-page prompt after completing the task. Note that the student cannot review the task or revise responses after he/she clicks the go-to-next-page

Appendix B: Data Collection Tool Example

Select here to go to the next page.

Q: If the student did not enter his/her student state ID at the beginning of the task, the student will be sent back to the beginning of the task. Ask the student to enter her/his student state ID and scroll through the 1<sup>st</sup> task then select the blue prompt to move to the 2<sup>nd</sup> task and complete the test.

 Please limit the response to 10 characters.

State Student ID

Q: When the student you and see this message, the student test and your Cognitive Lab Data Collection Tool are complete and submitted.

We thank you for your time spent taking this survey.  
Your response has been recorded.

## Appendix F: Administration Training Presentation



### Spring 2019 Science Pilot Training

February 2019

#### Welcome

- Jeremy Heneger, Director of Assessment, NDE
- Sara Cooper, Science Education Specialist, NDE
- Rhonda True, Enhanced Assessment Grant Coordinator, NDE
- Roy Beven, Senior Content Specialist Science, NWEA

#### Agenda

- NSCAS Summative Science Overview
- Spring Pilot Overview
- Science Pilot Administration
- Science Pilot Preparation
- Resources
- Questions & Answers



#### Development Timeline



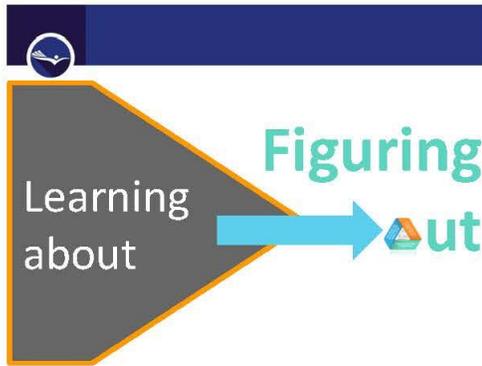
#### 3 Dimensional Standards



#### FRAMEWORK INSPIRED

To remain true to the intent of the Framework, indicators are 1 **EXAMPLE** of how to put the 3 dimensions together.

Students can and should use multiple CCCs and SEPs to engage in *figuring out* phenomena related to any DCI.



### 3 Prioritized Instructional Shifts



### Nebraska's Claims

**Overall Claim**  
Students can demonstrate the scientific literacy necessary to be civic-minded decision makers and demonstrate readiness for college, career, and lifelong learning through application of science and engineering practices and crosscutting concepts within and among the disciplines of science.

**Critical Consumers of Information**  
Students can gather, analyze, and communicate information from multiple sources to use as evidence to make sense of familiar and unfamiliar phenomena and problems.

**Interconnectedness of Science**  
Students can make connections between disciplinary core ideas within the physical science, life science, and Earth and Space sciences domains, across multiple science domains, and across multiple content areas (such as mathematics and English language arts) to make sense of familiar and unfamiliar phenomena and problems.

### Features to Maintain Consistency

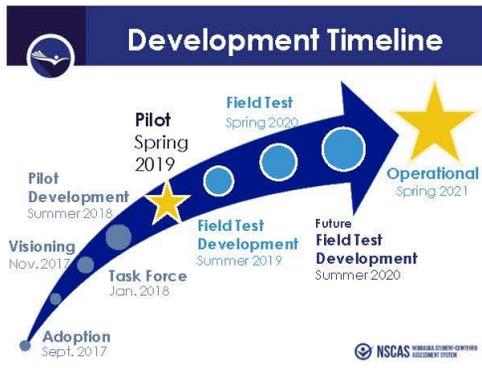
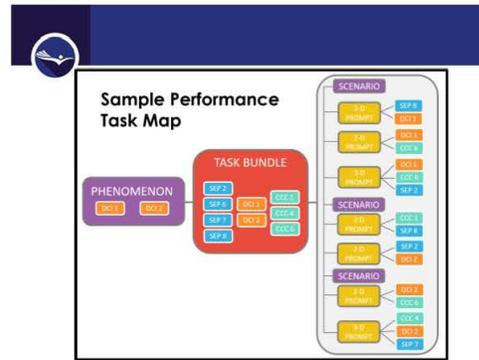
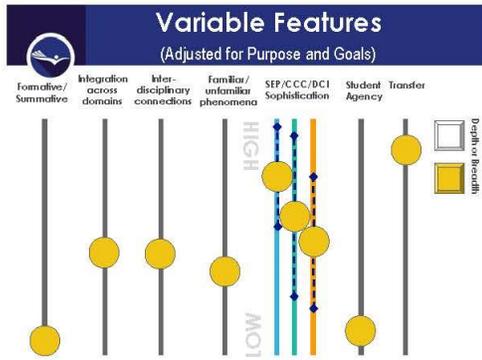
|                               |                             |                                 |                              |  |
|-------------------------------|-----------------------------|---------------------------------|------------------------------|--|
| Phenomena and Problem-focused | Engage diverse sense-making | Require reasoning with evidence | Grade appropriate 3D targets | Demonstrate science understanding by doing science |
|-------------------------------|-----------------------------|---------------------------------|------------------------------|--|

### Assessment System Components

\*Common Thread: Professional learning for educators\*

| Curriculum Embedded Tasks | Task Library | Monitoring Tasks | Statewide Summative |
|---------------------------|--------------|------------------|---------------------|
| (K-12)                    | (K-12)       | (3,4,6,7,9,10)   | (5,8,11)**          |

\*\*ACT for 3rd year cohort



### Spring 2019 Science Pilot Overview

**Purpose:**

- Gather feedback on the new science tasks with new types of questions
- Inform task development in the Summer of 2019 by evaluating the new style and question types
- Not** to predict performance on the Nebraska College and Career Ready Standards for Science. As a result, student scores will not be provided.

**Participants:**

- All districts were encouraged to participate. Participation is voluntary.
- All students in grade 5 and 8 are eligible to participate.

**Pilot Test Window:** March 4 – March 15, 2019

nwea

### Administration

**Administration Platform:**

- The science pilot assessment will be administered through Qualtrics, a web-based assessment system
  - Supported browsers include:
    - Apple Safari
    - Google Chrome
    - Microsoft Edge
    - Microsoft Internet Explorer
    - Mozilla Firefox
- For more information on the Qualtrics platform and technical support, please visit: <https://www.qualtrics.com/support/survey-platform/getting-started/help-and-feedback/#LoginBrowserCompatibility>

nwea

### Administration Information

**Forms:**

- There are **two** forms at each grade level.
  - Grade 5 form A
  - Grade 5 form B
  - Grade 8 form A
  - Grade 8 form B
- Links to the forms will be posted on the NDE Student Assessment website
- Students should be randomly assigned to a form.
  - Recommendation:** For any group of students in either grade 5 or 8, assign half of the students form A, and the other half form B

nwea

## Administration Information

**Forms:**

- The first question will require the student enter his/her 10-digit state ID. No additional student information will be required.
- Each form has two tasks, each with one or two scenarios.
- Each task includes from 5 to 8 questions.
- After each task, students will be asked to provide feedback on their interest, the difficulty of the questions, as well as how well they understood the task.

**Test Duration and Scheduling:**

- The pilot assessment is not timed. The majority of students should complete a form within 45 minutes.
- Students should complete the entire form in one test session.

nwea

## Administration - Accessibility

All students should be provided scratch paper and a calculator

There are no standard embedded accommodations/accessibility supports provided through the Qualtrics platform.

- Students may use approved non-embedded resources, such as multiplication charts or note buffers, as specified by NDE policy.
- Zoom/magnifier – students should use the native device zoom feature to magnify the content on the page.
- Text-to-speech – students should use native screen readers or read aloud in accordance with the NDE accessibility manual

A complete list of non-embedded universal tools, linguistic supports, and accommodations is included in the [NSCAS General Summative & Alternate Accessibility Manual](#).

nwea

## Preparation Tips

- Ensure device browsers meet the technical requirements
- Provide students with an opportunity to practice using the sample tasks. This provides students with an opportunity to gain familiarity with the Qualtrics interface and exposes students to the various types of questions on the pilot assessments
- Review the Science Pilot Test Administration Manual
- Determine procedures for student assignments of Form A or Form B
- Ensure students know or have access to their state ID to participate in the pilot

nwea

## Administration Resources

**Sample Tasks:**

- Two tasks are available for students and districts to review prior to the administering the pilot assessments.
  - Grade 5 Changing Coral: <http://bit.ly/prep1a15>
  - Grade 8 Blue Skin: <http://bit.ly/prep1a18>
- **Recommendation:** All students participating in the pilot should complete the appropriate practice task prior to participating in the pilot.

**Pilot Administration Manual:**

- Includes all procedures and proctor script
- Posted to the Assessment Portal

nwea

## Feedback

- **Embedded Student Feedback Questions**
  - After each task, students will be prompted to provide feedback
- **Cognitive Labs**
  - A subset of districts will be participating in cognitive labs during the administration. This will be used to gather additional student feedback.

## Help Desk

- Policy questions: Contact NDE
  - Phone: 402-471-2475
  - Email: [nde.stateassessment@nebraska.gov](mailto:nde.stateassessment@nebraska.gov)
- Science Pilot inquires or support: Contact NWEA
  - Phone: (855) 225-9926
  - Email: [INWE@Nebraska@nwea.org](mailto:INWE@Nebraska@nwea.org)
  - 7:00 a.m. – 5:00 p.m. Central Time (CT), Monday – Friday

nwea



## Questions & Answers

nwea

## Appendix G: Cognitive Lab Training Presentation




Spring 2019 Science  
Cognitive Lab Pilot Training  
February 2019

### Welcome

- Jeremy Heneger, Director of Assessment, NDE
- Sara Cooper, Science Education Specialist, NDE
- Rhonda True, Enhanced Assessment Grant Coordinator, NDE
- Roy Beven, Senior Content Specialist Science, NWEA

### Agenda

- Introduction
- Cognitive Lab Overview
- Preparation and Logistics
- Cognitive Lab Protocol
- Resources
- Questions & Answers 

### Development Timeline



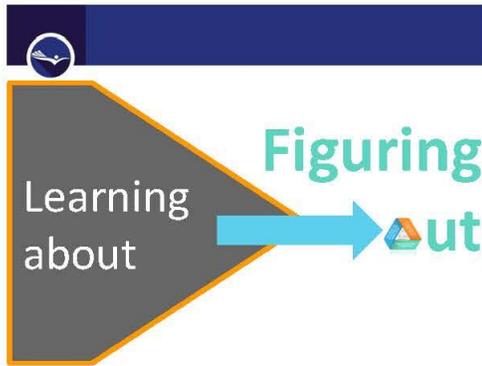
### 3 Dimensional Standards



### FRAMEWORK INSPIRED

To remain true to the intent of the Framework, indicators are 1 **EXAMPLE** of how to put the 3 dimensions together.

Students can and should use multiple CCCs and SEPs to engage in *figuring out* phenomena related to any DCI.



### 3 Prioritized Instructional Shifts

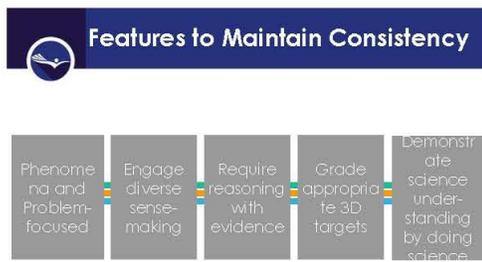


### Nebraska's Claims

**Overall Claim**  
Students can demonstrate the scientific literacy necessary to be civic-minded decision makers and demonstrate readiness for college, career, and lifelong learning through application of science and engineering practices and crosscutting concepts within and among the disciplines of science.

**Critical Consumers of Information**  
Students can gather, analyze, and communicate information from multiple sources to use as evidence to make sense of familiar and unfamiliar phenomena and problems.

**Interconnectedness of Science**  
Students can make connections between disciplinary core ideas within the physical science, life science, and Earth and Space sciences domains, across multiple science domains, and across multiple content areas (such as mathematics and English language arts) to make sense of familiar and unfamiliar phenomena and problems.

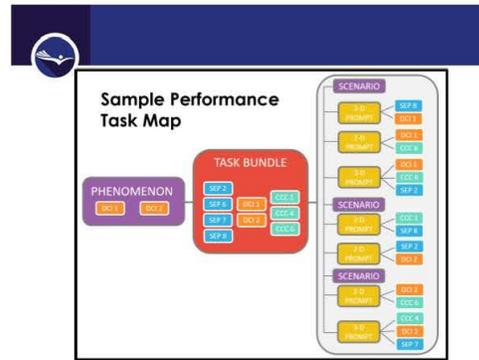
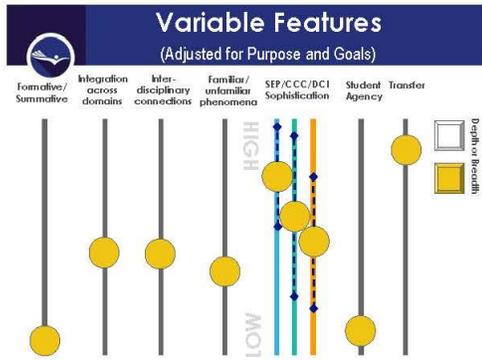


### Assessment System Components

**\*Common Thread: Professional learning for educators\***

| Curriculum Embedded Tasks | Task Library | Monitoring Tasks | Statewide Summative |
|---------------------------|--------------|------------------|---------------------|
| (K-12)                    | (K-12)       | (3,4,6,7,9,10)   | (5,8,11)**          |

\*\*ACT for 3rd year cohort



### Purpose of the Cognitive Labs

- To collect detailed information about the process students take when answering test questions on the NSCAS Science Pilot assessment from March 4–15, 2019, to inform future test development

nwea

### Cognitive Lab Overview

- You will interact with one student as they work through two science tasks and their associated test questions.
- The student will respond to the questions online while thinking aloud, being encouraged by you to do so.
- While the student thinks aloud, you will document your observations of their thought processes and other feedback from the student in the data collection tool.

nwea

### Cognitive Lab Preparation

- Review the pilot test and data collection tool to familiarize yourself with them.
- Read the Cognitive Lab Protocol document before engaging with the students.

nwea

## Logistics

- One-on-one setting. Make sure you interact with only one student.
- Follow the school's procedures to escort the student from their classroom to an empty testing room or a corner of a quiet room.
- Materials:
  - Cognitive Lab Protocol
  - Laptop or tablet with a keyboard for both yourself and the student
  - Data collection tool (log in to your computer and have the data collection tool on Qualtrics ready to record your observations)
  - Pilot test form (let the student log in to their computer and access the form). The cog lab will be conducted using:
    - o Grade 5 Form B
    - o Grade 6 Form A

nwea

## Things to keep in mind:

- Do not tell the student if their answer is right or wrong.
- Do not ask questions that lead to a student's response.
- Do not interrupt the student to ask why they are responding in a particular way. Only prompt if they are not thinking aloud.
- Encourage the student to be more descriptive if they are not talking much, but do not put words in their mouth.
- Sit next to the student but not in their personal space.

nwea

## Things to keep in mind:

- Be as objective and non-biased as possible.
- Focus on the task at hand (the particular item or task). Do not ask the student a question in general terms.
- Ask the student to explain why the answer options are not right if they do not provide that information.
- Capture the frustration aspect of taking the test (and performing the think aloud exercise) if a student seems hung up with something.

nwea

## Cognitive Lab Protocol Process Overview

1. General Introduction: Introduce yourself and explain the purpose of the cog lab.
2. Think Aloud Practice: Explain how to think aloud and practice with the student.
3. Conduct the Cognitive Interview:
  - Have the student read the scenario aloud and answer the first question for the first science task. Document the student's responses in the data collection tool.
  - Repeat these steps until the student finishes all the questions for the first task.
4. Concluding Questions: Have the student respond to the questions at the end of the task. Ask the additional follow-up questions provided in the data collection tool and record their responses.
5. Repeat Steps 3 and 4 for the second science task.

nwea

## Step 1: General Introduction

- Introduce yourself and the purpose of the cognitive lab.
- Inform the student that their responses will be kept confidential and will not be graded.
- Answer any questions the student may have.

*Hi! My name is (YOUR NAME). I'm (YOUR ROLE). Thank you for helping me today. Today you will be participating in a special study. Its goal is to find better ways to design science tests for all students. We need your help creating a science test.*

*You will not be asked to write your name on any of the work you do, and no one in the school will see your answers. The information you provide will be kept confidential. Even though this is a science test, your responses will not be graded, and this test will not affect your grade in your class. We are really interested in figuring out how we can make the test questions better for students.*

*Do you have any questions before we begin?*

nwea

## Step 2: Think Aloud Practice

- Explain what it means to think aloud.
- Practice thinking aloud with the student.

*I am interested in capturing what you notice what you are thinking as you respond to questions. The best way to find out what you are thinking is to have you read and talk out loud while you are working on the test. As you answer the questions, I would like you to say out loud things that you would normally think to yourself.*

nwea

### Step 2: Think Aloud Practice

Use the information given about the organisms in the ocean ecosystem to predict the organism that will be MOST affected when sediment in the ocean increases.

A. dolphin  
B. triggerfish  
C. coral  
D. algae  
E. bob-baita

*Hint: Carefully read the information and question. Then I read through all the answer options. I think about sediment increasing in the ocean and how those could affect the organisms listed. I think sediment blocks sunlight and could cover organisms. I see that algae needs the Sun for energy so I think that organism. I think about how the other organisms but I think the algae will be affected the most.*

Now, tell me how you would answer this same question.

| Organism    | Picture   | How it obtains energy |
|-------------|---|-----------------------|
| Dolphin     |  | eat fish              |
| Triggerfish |  | eat coral             |
| Coral       |  | eat algae             |
| Algae       |  | from the Sun          |
| Bacteria    |  | from dead organisms   |

nwea

### Step 3: Conduct the Cognitive Interview

- Continue to practice with the student until they understand how to think aloud. Answer any questions the student may have.
- Be sure the student enters his/her 10-digit state student ID number. If the student does not, she/he will be prompted at the end of the 1<sup>st</sup> task to do so before going to the 2<sup>nd</sup> task.
- This is a screenshot of the actual Cognitive Lab Data Collection Tool with your instruction, "Ask the student to read the directions and scenario out loud."

Ask the student to enter the 10-digit state student ID number. Note that the student must enter his/her ID number to proceed to task #2 of the test.

State Student ID:

Ask the student to read the directions and scenario out loud.

nwea

### Step 3: Conduct the Cognitive Interview

- Let the student read aloud the scenario associated with the first science task.
- Typically, scenarios include text, illustrations, charts, and data like this one about Changing Coral.

Scientists have been investigating coral reef ecosystems over time. They have observed many changes. In some areas, coral reefs have disappeared. The scientists are wondering what is causing these changes.




**Information about Coral Reefs**

- Coral reefs are made from animals that live in the same place their entire lives.
- Coral reefs need shallow, clear, warm water to survive.
- Most corals live with algae which are tiny, plant-like organisms that need sunlight to grow.

nwea

### Step 3: Conduct the Cognitive Interview

- Ask the student this follow-up lab question and record their answers in the Data Collection Tool, along with any observations you made about the student's interaction with the task scenario.
- Notice the arrow in the blue box. These arrows allow you to move from page to page and back.

After they have finished ask the student:

A) Did you understand the text?  
B) What are you wondering after reading the passage?  
C) What science knowledge might you need to understand (change) coral?

Record their answers and any observations about the student's interaction with the scenario.

nwea

### Step 3: Conduct the Cognitive Interview

- Ask the student to read aloud the first task question associated with the scenario.
- Ask the student what he/she is thinking as she/he reads and answers the question. You may need to prompt the student.
- Record the student's thoughts.

Ask the student to read aloud and answer Question 1.

**Question 1**

Some scientists claim that there is a "c-o-l-d" cause for this (c-o-r-a-l) problem. Which number would be about their claim?

A) The ocean water is warmer than when it was in 1980.

B) The ocean today has more algae than it did in 1980.

C) There is more oxygen in the ocean today than there was in 1980.

D) There is more coral in the (c-o-r-a-l) today than there was in 1980.

Ask the student to describe what they are thinking as they read and answer the question. If the student is not taking you may ask, "What are you thinking?" or "Why did you answer that way?" or "What would help you answer this question?"

nwea

### Step 3: Conduct the Cognitive Interview

- There are a set of 4 lab questions for each of the task questions.
- The first lab question delves into the evidence you think the student said to you. The options are customized for the task question.
- Lab questions are in both selected response and open-ended response format.

The student's comments provided evidence that the item elicited thinking about:

The passage on water speed.  Other

The model to describe (change) coral.

If "Other" what thinking did the item elicit?

While reading and interacting with the question, the student appeared to:

Look to answer the question without careful consideration of the question, the text and possible responses.  Get stuck or understand how to generate a response to the question.

The response with the question.  Being more knowledgeable and confident.

nwea

### Step 3: Conduct the Cognitive Interview

- The 3<sup>rd</sup> lab question is probing the difficulty of just responding to the task question.
- The 4<sup>th</sup> lab question is probing the difficulty in using the given information.

Ask the student:  
Was it easy or hard to know how to respond to the question?

1 2 3 4 5

Ask the student:  
What would help you answer the question?

Ask the student:  
Was it easy or hard to use the given information to answer the question?

1 2 3 4 5

Ask the student:  
Why was it easy or hard?

nwea

### Step 3: Conduct the Cognitive Interview

- The lab questions for every task question finishes with a place to record other observations of student interactions with the task question.
- Notice how you can go back if you need to add to or adjust your responses.

Record any additional notes, comments, or observations made about the student's interaction with this question.

nwea

### Step 4: Concluding Questions

- After the student finishes the 5 to 8 questions of a task, and before he/she starts the Student Feedback section, ask these 3 lab questions about comparing the task she/he just completed to their classroom science tests.

Ask the student:  
What are the science tests like in your class?

Record their response:

Ask the student:  
How does that compare with what you just did?

Record their response:

Record any additional notes, comments, or observations made about the student's interaction with this question.

nwea

### Step 4: Concluding Questions

- After the student finishes all the questions of a task, they will respond to a Student Feedback section.
- Based on their ratings, ask follow-up questions such as "Why was the task interesting?" and record their responses in the data collection tool.

Ask the student to read about and answer the first question in the Student Feedback section.

Record their response:

Ask the student to read about and answer the second question in the Student Feedback section.

Record their response:

nwea

### Step 4: Concluding Questions

- The final Student Feedback question asks the student to rate the difficulty of each task question.
- Ask the student "Why did you rate the difficulty the questions with 1, 2, 3, 4, or 5 stars?" and record their responses in the data collection tool.

Ask the student to read about and answer the third question in the Student Feedback section.

Record their response:

Ask the student: "Why did you rate the difficulty of the questions with 1, 2, 3, 4, or 5 stars?"

Select here to go to the next page.

nwea

### Step 5: Repeat Steps 3 & 4 for the Second Task

- Repeat Steps 3 and 4 for the second science task, recording all observations in the Data Collection Tool.

nwea

 Step 5: Completing & Submitting the Data Collection Tool

- After the student completes the student feedback question for the second and final task, she/he (and you) will see a directive in a blue box. When the student or you clicks that, the test or Cognitive Lab is completed.
- If you or the student did not enter the 10-digit student state ID, the test asks you to do so before completion.



nwea

 Closing

- Thank the student for their hard work.
- Dismiss the student according to school procedures.
- Answer any questions the student may have.

nwea

 Help Desk

- Policy questions: Contact NDE
  - Phone: 402-471-2496
  - Email: [nde.stateassessment@nebraska.gov](mailto:nde.stateassessment@nebraska.gov)
- Science Pilot inquires or support: Contact NWEA
  - Phone: (855) 225-9926
  - Email: [NWEANebraska@nwea.org](mailto:NWEANebraska@nwea.org)
  - 7:00 a.m. – 5:00 p.m. Central Time (CT), Monday – Friday

nwea

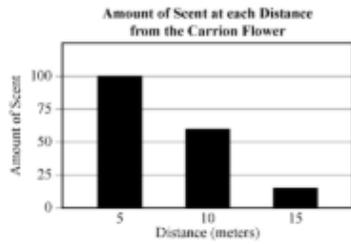
 Questions & Answers

nwea

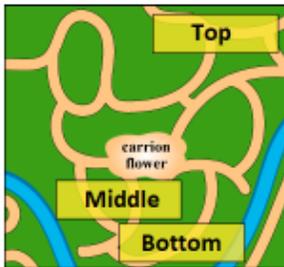


**Question 1**

Scientists can determine the amount of scent in a given area.  
The graph shows data collected for the amount of scent detected at three distances from the carrion flower.



Use the data to predict the strength of the scent at different distances from the carrion flower by moving each word into the diagram.



Drag one item into each box that matches the box in the diagram.

- Items
- Very Strong Scent
  - Strong Scent
  - Weak Scent

|        |
|--------|
| Top    |
|        |
| Middle |
|        |
| Bottom |
|        |

**Question 2**

**This Question has 3 parts. Answer Part A, then Part B, then Part C.**

**Part A**

Which model BEST represents the scent detected 5 meters from the carrion flower?

Model 1:  Nitrogen (red) and oxygen (blue) particles are clustered on the left side of the container. Scent particles (yellow) are clustered on the right side.

Model 2:  All particles (nitrogen, oxygen, and scent) are clustered at the bottom of the container.

Model 3:  All particles (nitrogen, oxygen, and scent) are randomly distributed throughout the container.

Model 4:  All particles (nitrogen, oxygen, and scent) are clustered in the center of the container.

**Part B**

Use the models and data to make a claim about the scent particles.

The particles of the scent are a

-----

**Part C**

Which evidence supports the answer to Part B.

- A) The air particles are in a tight pattern, and the scent particles are spread out.
- B) The air particles are in a random pattern, and the scent particles are spread out.
- C) The air particles are in a tight pattern, and the scent particles are close together.
- D) The air particles are in a random pattern, and the scent particles are close together.

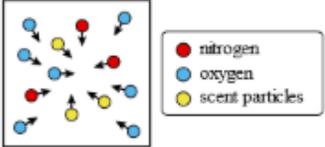
**Question 3**

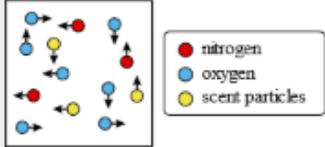
Which behavior of the scent particles explains why the students can smell the carrion flower from far away?

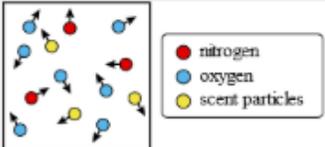
- A) The scent particles stay near the flower.
- B) The scent particles travel back into the flower.
- C) The scent particles move freely away from the flower.
- D) The scent particles flow near the ground around the flower.

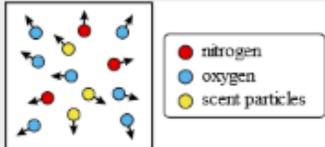
**Question 4**

Which model BEST represents the behavior of scent particles in an air sample?

○  ● nitrogen  
● oxygen  
● scent particles

○  ● nitrogen  
● oxygen  
● scent particles

○  ● nitrogen  
● oxygen  
● scent particles

○  ● nitrogen  
● oxygen  
● scent particles

**Question 5**

Construct an explanation citing evidence from the text, graph, and models to explain why the students can smell the carrion flower without seeing it. Include why the strength of the scent varies on the tour.

**Grade 5, Form B, Task 2**

**Grade 5 Form B**

**Task #2 TV Trouble**

Use the information provided in any part of this task to answer the questions.

A Nebraska student is watching a movie on TV at 9 a.m. one fall morning. At about 11 a.m., the student is having a hard time seeing the movie on the TV screen. Why is it hard to see the movie on the TV screen?



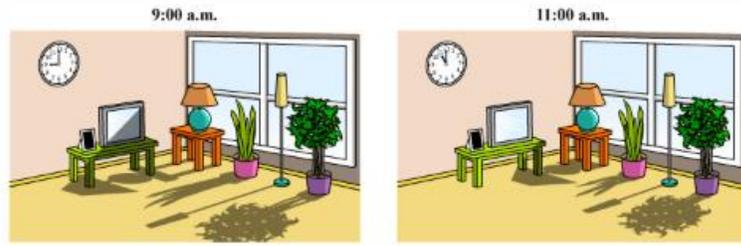
**Question 1**

What causes the shadow on the TV screen to disappear later that same morning?

- A) the position of the sun in the sky
- B) the position of the TV in the room
- C) the position of the couch in the room
- D) the position of the lamp in front of the window

Use the information provided in any part of this task to answer the questions.

A Nebraska student is watching a movie on TV at 9 a.m. one fall morning. At about 11 a.m., the student is having a hard time seeing the movie on the TV screen. Why is it hard to see the movie on the TV screen?

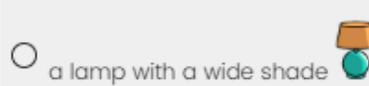
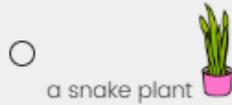
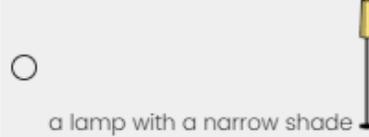


### Question 2

Select the model that explains why the Sun's position in the sky changes between 9:00 a.m. and 11:00 a.m.?

**Question 3**

Which object could be placed on the table to help the student see the movie better on the TV screen at 11:00 a.m.?



Three months later, in winter, the student is watching TV at 11:00 a.m. and having trouble seeing the screen, but the student can see the screen clearly in the afternoon. The student wonders what will happen at 11:00 a.m. in the spring.



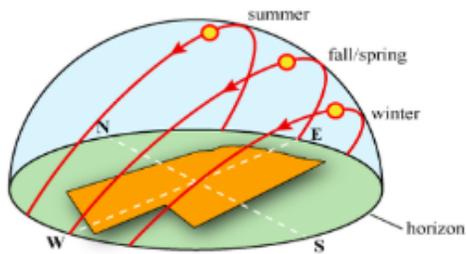
#### Question 4

Use the image to predict what will happen to the shadows in the room from winter to spring?

- A) The shadows will be wider.
- B) The shadows will be longer.
- C) The shadows will be shorter.
- D) The shadows will stay the same.

#### Question 5

##### The Sun's Path during the Year

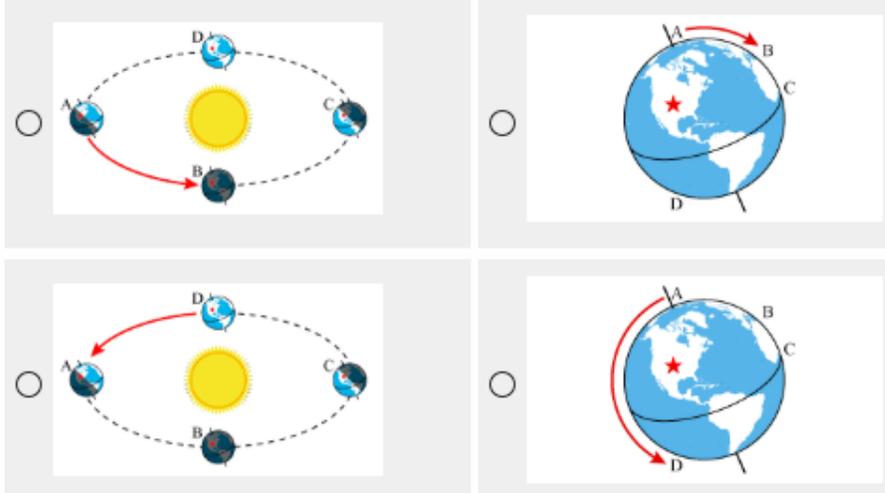


The Sun's angle is measured from the horizon, which is the line where Earth's surface and the sky meet. Using the model, what is the relationship between the season and the angle between the Sun and the horizon?

- A) From spring to fall, the angle between the Sun and the horizon at 11:00 a.m. increases.
- B) From fall to winter, the angle between the Sun and the horizon at 11:00 a.m. increases.
- C) From winter to spring, the angle between the Sun and the horizon at 11:00 a.m. decreases.
- D) From summer to winter, the angle between the Sun and the horizon at 11:00 a.m. decreases.

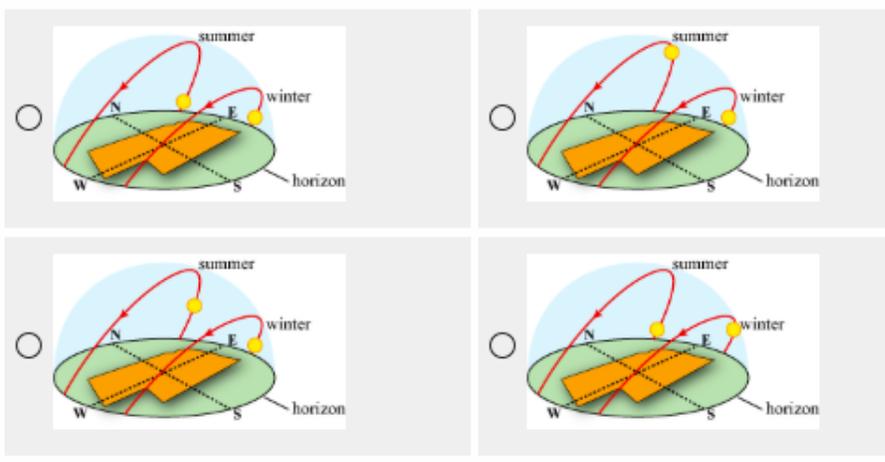
**Question 6**

What is causing the angle between the Sun and the horizon to change from fall to winter?



**Question 7**

Which model BEST represents the Sun's positions at 9:00 a.m. for summer and for winter in Nebraska?



**Question 8**

Predict the shadow length of the objects in the room at 9:00 in the summer and in the winter to explain why it is difficult to see the TV screen at different times of the year.

- A) The shadows are the same in summer and winter.
- B) The shadows in winter are longer than shadows in summer.
- C) The shadows in summer are longer than shadows in winter.
- D) The shadows in summer are shorter than shadows in the winter.

**Grade 8, Form A, Task 1**

**Grade 8 Form A**

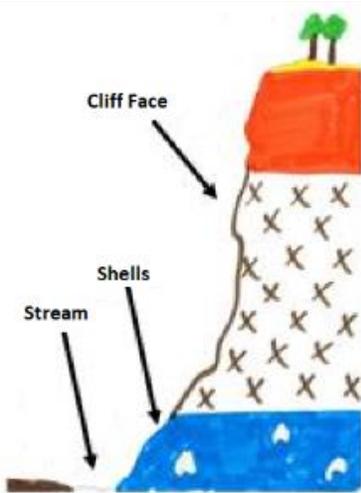
**Task #1 Fossil Discovery**

State Student ID

Use the information provided in any part of this task to answer the questions.

Shortly after a field trip to Ashfall Fossil Beds State Historical Park, a student finds a rock with shells embedded in it while hiking near a stream. After looking around, the student notices a layer of rock in the cliff with more shells embedded. The student sketches the rock layers seen in the cliff.

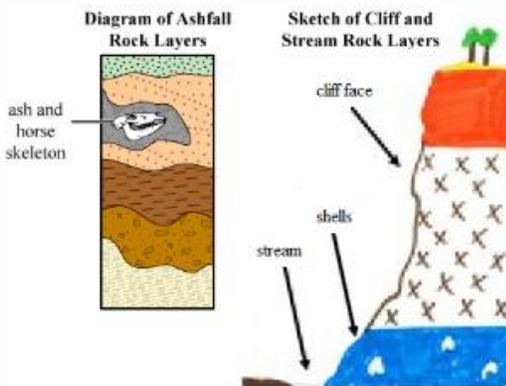




At Ashfall Fossil Beds State Historical Park, the student saw fossilized skeletons of horses and other mammals. The student wonders which is older: the rock with the shell or the fossil horse skeleton.



The student compares their sketch of the cliff with the shell layer to a diagram of rock layers at Ashfall Fossil Beds State Historical Park.



### Question 1

**This question has two parts. Answer part A, then answer part B.**

#### Part A

Make a claim about the age of the rock with shells compared to the age of the fossil horse skeleton.

- A) The rock with shells is likely older than the fossil horse skeleton.
- B) The rock with shells is likely younger than the fossil horse skeleton.

#### Part B

What evidence from the rock layers supports the claim in part A?

- A) The rock layers are in the same order in both locations.
- B) The volcanic ash layer is not seen in the cliff near the stream.
- C) The shells are in a lower rock layer than the fossil horse skeleton.
- D) The shells are in a younger rock layer than the fossil horse skeleton.
- E) The shells were exposed by a stream, while the fossil horse skeleton was not.

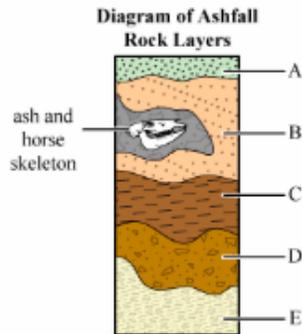
### Question 2

What evidence would support a claim that the shell in the rock is a fossil?

- A) It looked very old.
- B) It was easily broken.
- C) It was found inside a rock.
- D) It was found near a stream.
- E) It was found in the oldest rock layer.
- F) It was found in the youngest rock layer.

**Question 3**

Use the diagram of rock layers at Ashfall Fossil Beds State Historical Park to predict which layer might contain fossil shells like the one found near the stream.



Which layer is MOST LIKELY to have fossil shells like the ones near the stream?

A

B

C

D

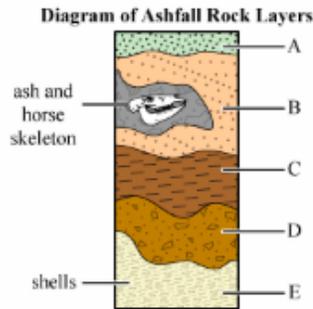
E

**Question 4**

**This question has two parts. Answer part A, then answer part B.**

The diagram shows other fossils found in the shell layer.

|                  |  |   |
|------------------|--|---|
| <b>Paleozoic</b> | <b>Devonian</b><br>359–419 million years ago   |  |
|                  | <b>Silurian</b><br>419–443 million years ago   |  |
|                  | <b>Ordovician</b><br>443–485 million years ago |  |
|                  | <b>Cambrian</b><br>485–541 million years ago   |  |



**Part A**

Make a claim about the local environment when the shell layer formed.

- A) The environment was marine and similar to layer B.
- B) The environment was similar to the environment today.
- C) The environment was marine and different from layer B.
- D) The environment was the same as the environment in layer B.

**Part B**

What evidence supports the claim made in part A?

- A) Environments change over time.
- B) Environments remain the same over time.
- C) Shells are found in Nebraska's streams today.
- D) The shell layer contains many marine organisms.

Use the information in any part of this task to answer the questions.

The diagram shows the geologic timescale and some common Nebraska organisms at those time periods.

| Era       | Period        | Millions of Years Ago (approx.) | Nebraska Organisms   |   |
|-----------|---------------|---------------------------------|--|---|
| Cenozoic  | Quaternary    | 2.5                             |   |   |
|           | Tertiary      | 66                              |  |   |
| Mesozoic  | Cretaceous    | 145                             |   |   |
|           | Jurassic      | 201                             |  |   |
|           | Triassic      | 251                             |  |   |
|           | Permian       | 299                             |  |   |
| Paleozoic | Carboniferous | Pennsylvanian                   | 323  |  |
|           |               | Mississippian                   | 359  |   |
|           |               | Devonian                        | 419  |   |
|           | Silurian      | 443                             |   |   |
|           | Ordovician    | 485                             |  |   |
|           | Cambrian      | 541                             |  |   |

### Question 5

Based on the diagram, which statement BEST describes the environmental changes that occurred in Nebraska?

- A) Nebraska changed from warm to cold.
- B) Nebraska changed from cold to warm.
- C) Nebraska changed from ocean to swamp to prairie.
- D) Nebraska changed from prairie to swamp to ocean.

### Question 6

**This question has two parts. Answer part A, then answer part B.**

#### Part A

Scientists at Ashfall Fossil Beds State Historical Park estimate that the horse skeleton fossils were formed about 11 million years ago.

Make a claim about the age of the shells found by the student.

- A) The shells are younger than 11 million years.
- B) The shells are between 11 million and 200 million years old.
- C) The shells are between 200 million and 541 million years old.
- D) The shells are older than 541 million years.

#### Part B

What evidence about the age of the shells supports the claim made in part A?

- A) The shells are marine in origin.
- B) The shells are in the bottom rock layer.
- C) The shells are similar in age to the horse skeleton fossils.
- D) The shells are found in Nebraska's streams today.

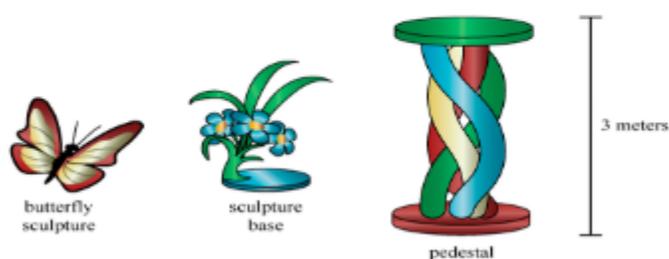
## Grade 8, Form A, Task 2

### Grade 8 Form A

#### Task #2 Sculpture Setup

Use the information provided in any part of this task to answer the questions.

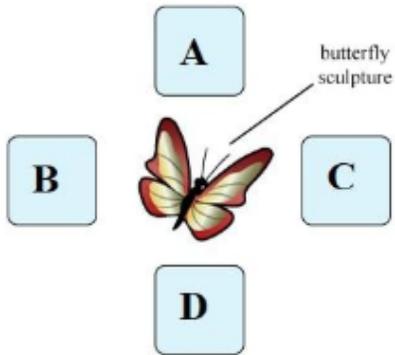
A two-part sculpture of a butterfly was donated to a school and needs to be placed onto a pedestal. Students will program a robot to place the sculpture into the pedestal safely without breaking the sculpture or the pedestal.



- The sculpture base has a mass of 40 kilograms.
- The butterfly sculpture has a mass of 10 kilograms.
- The robot applies an unbalanced force of 1000 newtons.
- The pedestal is 3 meters tall.
- The safe acceleration rate for the sculpture is 2 meters per second squared.

**Question 1**

The students think about how forces will affect the motion of the sculpture. Move one force arrow into each box to show how forces cause the sculpture to be lifted upward.



| Items   | A | B |
|---|---|---|
| Small Up     |   |   |
| Large Up     |   |   |
| Small Left   | C | D |
| Large Left   |   |   |
| Small Right  |   |   |
| Large Right  |   |   |
| Small Down   |   |   |
| Large Down   |   |   |

**Question 2**

**This question has two parts. First answer Part A, then Part B.**

To program the robot, students do an investigation to determine how much applied force is needed to lift objects with different masses. The students lift the objects with a spring scale and record the results in the data table.



| Mass of Objects (kilograms) | Applied Force (newtons) |
|-----------------------------|-------------------------|
| 1                           | 10                      |
| 2                           | 20                      |
| 3                           | 30                      |
| 4                           | 40                      |

**Part A**

Which relationship do the students find between the mass of an object and the applied force?

- A) The object's mass is 10 times the applied force.
- B) The mass of the objects causes the applied force to increase.
- C) The less the applied force, the greater the mass of the objects.
- D) The greater the mass of an object, the greater the applied force.

**Part B**

Calculate the minimum applied force needed (in newtons) to lift the sculpture base and the butterfly together onto the pedestal. Enter your answer in the box.

### Question 3

Students do another investigation to determine the relationship between acceleration and force. They lift a 10-kilogram object repeatedly with increasing unbalanced forces. They measure the resulting accelerations and record the results in the data table.

| <b>Mass of Object (kilograms)</b> | <b>Unbalanced Force (newtons)</b> | <b>Acceleration of Object (meters per second squared)</b> |
|-----------------------------------|-----------------------------------|---|
| 10                                | 10                                | 1   |
| 10                                | 50                                | 5   |
| 10                                | 100                               | 10  |
| 10                                | 150                               | 15  |

Which relationship do the students find between acceleration and force?

- A) As the unbalanced force increases, the mass of the object stays the same.
- B) As the unbalanced force increases, the acceleration of the object increases.
- C) As the acceleration of the object increases, the unbalanced force decreases.
- D) As the acceleration of the object increases, the mass of the object stays the same.

**Question 4**

This question has two parts. First answer Part A, then Part B.

**Part A**

Use patterns in the data as evidence to make a scientific claim about the relationship between the unbalanced force applied to an object, the mass of the object, and the change in motion.

| Mass of Objects<br>(kilograms) | Applied force<br>(newtons) |
|--------------------------------|----------------------------|
| 1                              | 10                         |
| 2                              | 20                         |
| 3                              | 30                         |
| 4                              | 40                         |

| Mass of Object<br>(kilograms) | Unbalanced Force<br>(newtons) | Acceleration of Object<br>(meters per second squared) |
|-------------------------------|-------------------------------|---|
| 10                            | 10                            | 1   |
| 10                            | 50                            | 5   |
| 10                            | 100                           | 10  |
| 10                            | 150                           | 15  |

Which conclusion can the students make from the patterns of data from their investigations?

- A) If mass increases from 1 kilogram to 4 kilograms, the motion stays the same.
- B) If the force increases from 2 newtons to 8 newtons, the motions stays the same.
- C) If the force increases by 10 newtons for each 1-kilogram increase in mass, then the change in motion will stay the same.
- D) If the mass increases by 2 times, then the force must also increase by 2 times to keep the change in motion the same.

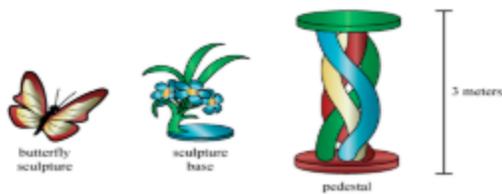
**Part B**

Which mathematical formula BEST represents the data?

- A)  $a = m / f$
- B)  $f = ma$
- C)  $m = fa$
- D)  $m = a / f$

Use the information provided in any part of this task to answer the questions.

A two-part sculpture of a butterfly was donated to a school and needs to be placed onto a pedestal. Students will program a robot to place the sculpture into the pedestal safely without breaking the sculpture or the pedestal.



- The sculpture base has a mass of 40 kilograms.
- The butterfly sculpture has a mass of 10 kilograms.
- The robot applies an unbalanced force of 1000 newtons.
- The pedestal is 3 meters tall.
- The safe acceleration rate for the sculpture is 2 meters per second squared.

| Mass of Objects (kilograms) | Applied force (newtons) |
|-----------------------------|-------------------------|
| 1                           | 10                      |
| 2                           | 20                      |
| 3                           | 30                      |
| 4                           | 40                      |

| Mass of Object (kilograms) | Unbalanced Force (newtons) | Acceleration of Object (meters per second squared) |
|----------------------------|----------------------------|--|
| 10                         | 10                         | 1  |
| 10                         | 50                         | 5  |
| 10                         | 100                        | 10   |
| 10                         | 150                        | 15   |

### Question 5

The students have two options to set up the sculpture.

- Assemble the 10-kilogram butterfly sculpture on the 40-kilogram sculpture base and move them together onto the pedestal.
- Move the 40-kilogram sculpture base onto the pedestal, then move the 10-kilogram butterfly sculpture onto the pedestal.

Use the equation to determine the acceleration of the parts and the assembled sculpture when the robot moves them with an unbalanced force of 1,000 newtons.

$$F = ma$$

Enter the acceleration (in meters per second squared) of each part of the sculpture and the assembled sculpture into the boxes.

Acceleration of the  
40-kilogram butterfly  
sculpture:

Acceleration of the  
10-kilogram sculpture  
base:

Acceleration of the  
50-kilogram  
sculpture and  
sculpture base  
together:

**Question 6**

How could the sculpture be safely placed onto the pedestal using the robot? Make a claim for what is the best solution to the problem. Use the data and your calculations as evidence to support your claim.



**Appendix I: Cognitive Lab Responses**

**1. After they have finished ask the student:**

- D. Did you understand the text?**
- E. What are you wondering after reading the passage?**
- F. What science knowledge might you need to understand smells?**

**Record their answers and any observations about the student’s interaction with the scenario.**

**Table I.1. Question 1 Responses**

| Question 1 Responses  |
|---|
| <b>Grade 5, Form B, Task 1</b>  |
| Pay attention to the picture because that can help - it has range of distance and how far the flowers are apart. We have to look back about the areas and the 50 meters - keeping that in mind as well as the amount of scent in each area. If the carrion flower in in middle the top is further away - so not strong, middle is right by flower so it would be strongest, the middle would be the most between the bottom since it is closest |
| Yes. Understood the text. Wondering why the flower smells bad. Why did they stop in front of a bad smelling flower? The picture shows the flower is really big. That might matter. Most of the plants look 50 meters from the Carrion. Maybe it gives off fumes that can harm the other flowers. Need to know senses.   |
| Yes, I didn't know the flower name, but it didn't matter. Wondering why the flower smells like meat. Smells can show you how to identify things, it's a way of observing (5 senses).  |
| Yes. How does the flower grow? How does that one certain flower grow? I can't think of any.   |
| yes. B. What the smell is coming from until it told us. C. What type of plant, and why it might be making that smell. How smells travel through the air.  |
| yes. B. I am thinking that the plant breathes in the air and lets the bad air out. C. What makes the smell and how far smells can travel  |
| Kind of understand text. The rotted meat confused me. I don't know about smells   |
| Yes. I didn't know how to pronounce the flower name, but it doesn't matter. Need to know that smells are strong when you are closer to the source and they get weaker as you move away. There are good smells and bad smells  |
| <b>Grade 5, Form B, Task 2</b>  |
| Sun is sky is a possibility, TV didn't move or the couch in room - it doesn't show one, so is there a couch even in the room. The lamp in front of the window - the lamp is in the same place and you can see where the lamp shadow is. In the morning when the sun raises it is darker and then moves from east to west so it gets brighter. It is all about the sunrise - and there is more sun at 11 than 9                                  |
| Yes. understood text. Why would the student have a hard time seeing the screen? Need to know movement of sun. Where it sets and rises.  |
| Yes. I am wondering why they told us it was fall? Anytime there is sun it would work? I think you need to know about shadows and the earth's rotation and the suns position might make a shadow or not depending on where it is.  |
| A) A Little Bit. B) I don't know the difference between the pictures. C) Suns Rotation.   |
| A. Yes. B. I would say since the sun is moving shadows are on the lamp and then not on the lamp so the sun is on the screen. C. That the sun is moving so the shadows get shorter or even disappear at night.   |
| Yes. No. Don't Know.  |

### Question 1 Responses

Yes. No. Shadows are made by the sun.

A – yes. B - why they can't see the movie; C - you'll need to know how a shadow effects a TV and how it might change it in a different way

#### Grade 8, Form A, Task 1

A. Yes. understandable. B. Height of each layer? Age of each layer? Location of the park? C. Examples of animals that are fossilized here.

A. Yes. Understood the text. B. I wonder if the horse fossil is lower than the shells. Where are they positioned? What happens to fossils depending on their environment, example oceans vs other climates?

A. Yes. B - How you can tell the difference between the two. C - animal science and chemical science to see how long it has been there.

Yes. No. What they evolved from. How they stay fossil or disintegrate.

Yes. No. You might need to know what time period they are from, age of layers, when they are from.

A – yes. B - What the question will be based on? Will it go by years? C - Earth Science and how the layer are created and the age of the layers

#### Grade 8, Form A, Task 2

Yes, I understood the text. I am wondering the length of the robot arm and the mass of the robot. Surface areas of the pedestal and sculpture base. Need to know: Newton's Laws. What a Newton is. How mass and velocity plays into the equations. The role of gravity in the different equations

I understood the first part of the text but when it got to Kg and N I became more confused. I wonder how the N that the robot applies relates to the kg in the sculpture. Need to know how the mass of something and the force relate to each other.

A – Yes. B - Nothing at this time. C - Force and motion

Yes. I have no clue what a pedestal is, but the picture helped. It gives you mass and Newtons sop you will have to know what those mean

Kind of. The data was confusing, and I don't know how to use it. Need to know how much force to use so it doesn't break.

A - Ya because of the weight and the robot. B - What the robot needs to do and what the questions will be. C - Physical science - Newton Laws

**2. Ask the student to describe what they are thinking as they read and answer the question. If the student is not talking, you may ask "What are you thinking?" or "Why did you answer that way?" or "What would help you answer this question?".**

**Table I.2. Question 2 Responses**

| Question 2 Responses  |   |
|---|---|
| Grade 5, Form B, Task 1   |   |
| Item 1  | See previous notes  |
|   | The student described the graph incorrectly but then knew that the closer it was, the stronger the smell.   |
|   | Graph: The 5 is really strong scent, very close. 10 not as much and 15 is far away so not as much smell. Question: Top is a weak scent because it is farthest away, Middle is very strong because it is so close to the flower and the bottom is strong because it is still close. She had everything that she needed to answer the question.   |
|   | I think the middle would be into the flower. Wasn't sure how or where to drag the items. Top is weak scent because it is too far away. Bottom would be strong because it is close.  |
|   | Well the farther away from the plant the weaker the scent, because eventually, well it just weakens as it goes farther in the air   |
|   | I was thinking if I move far would smell but if I moved closer it would be stronger. Maybe because the flower is letting the bad air near the plant. If there is air it might go other places   |
|   | The weak scent is at the bottom because as they were moving the scent go weaker. Strong scent in the top not sure why. Very strong scent in middle it is close to the flower  |
|   | Doesn't understand the x axis variable and numbers on the graph? Weak at the bottom because at the bottom of a chart you aren't doing very well, middle is strong because middle is doing good, strong is at the top of the chart that means you are the best. See was reading the diagram incorrectly.   |
| read the table, estimated some mathematics. did the math 75-50 then divided by 2... it's around in the center of the two lines; it will show me the exact amounts |   |
| Item 2  | Do you need to go back to the first pictures? It doesn't seem to give you information you need - but the bar graph shows you the five meters. You have to use that to see which has the most scent - but they seem all equal (counting the scent dots). They didn't talk about the nitrogen or oxygen - but I know we give those to plants - the information is confusing. There are equal particles, so it must be the form the particles are in that they want me to look at - I think it might be the one grouped since it is strong - but it is confusing. I know it's not a solid or a liquid because you can see them, and you can't see scents, so it is a graph. Also, the way they represented it as particles - it's my background knowledge that helps me. I'm thinking about each option and I am going back to the graph that showed the particles and how they are located. The air would be the nitrogen and the oxygen. I can eliminate them by which looks like my answer that I chose. Several have tight patterns - but the air particles are different. I feel like it is C because they are tight together, but the scent particles are still close together |
|   | All parts collaborate together but parts B and C are really connected. Student really focused on 5 meters distance rather than the model of gases. "5 meters away would be strong scent so would go with B because the scent is up in the front so more scent and there are 6 scent particles and only 3 nitrogen. The Oxygen would spread out more than scent. Q2 he described behavior of gas correctly. Q3 looked back at the model and said the one he picked is a tight pattern. When thinking through #3 he went back to Q1 and changed it to C. He knew his reasoning was wrong by working through the questions.  |

## Question 2 Responses

|        |   |
|--------|---|
|        | <p>There is a lot of scent at 5m, oxygen is a type of gas, they are gas particles, the flower is a solid or living thing, so the particles should be close to each other, there. The graph said that the smell was bad, the diagram shows that the scent particles are on the top and sides where you can smell them, they should be a lot of oxygen, so the plant can push out the scent, less nitrogen in the plant. The last chapter was states of matter, solids are tightly packed, and a flower is a solid, so the particles should be close together. Claim: Part C: NO to A because the air particles are tight but not spread out, No to B because the air particles are random but are not spread out, No to C because the air particles are tight, but D gives the best explanation because they are random and tight.</p> |
|        | <p>Upper right diagram is the answer because I remember the flower is in the middle and the gases should be 5 meters away. Gas, because the scent is like air and you can smell that. I think A because particles are always tight together</p>   |
|        | <p>I honestly don't understand...oh I see the distance now and need to look at the diagram. Oh each one has the same amount. If this is the test, then it will be hard. The particles are a gas or solid and if I choose B then the answer to C is D. I think the diagram choices are confusing him even though he know the scent particles are close together.</p>   |
|        | <p>The fact that the 5 meters is close to the flower can't be found as the student can't go back and check the graph, so they chose answer B because nitrogen and scent particles spread once they are out of the plant. I chose gas because the scent was spreading all around. I chose B on C cause the air meets the smells and they mix together.</p>   |
|        | <p>Part A: Scrolled up to look at graph, trying to figure out what the pictures are trying to tell me. I picked upper left choice because the bottom middle and bottom of the garden picture weren't that spread out. Not using the correct diagram from above for reference. Part B: Gas, because oxygen is a gas. Part C: trying to distinguish tight pattern from random A picked because they are tight and close together</p>  |
|        | <p>Part A: Top right because I not sure I think the scent particles will be close together when you are so close to the flower. Part B: Gas because you can't see it. Part C: Scent particles are in a tight pattern, so I choose C.</p>  |
|        | <p>A - red are nitrogen, blue are oxygen, yellow are scent particles. 5 m is 100, so... I probably think... explained thinking of each choice. when particles go together they stay in a clump depending on how much space they have; more spread out particles would be farther away because they are more spread out. B - if solid, wouldn't really be able to smell; if liquid would stay on ground and not get to your nose; probably gas because you could sniff it; easier question because liquid or solid would change completely how our nose would function; C - went through each choice answering yes or no to each. Decided on ""B"" other choices were talking about particles being close together and they don't really have a pattern; previous answer to question (B) helped to answer question (C).</p>            |
| Item 3 | <p>I know I can take out A because they were far away and could smell it - so it didn't stay near the flower. Can rule out B since it travels back into the flower and wouldn't be able to smell it then. So, I need to focus on C and D. The scent particles move freely away may be why they could smell it because they would travel. D says it stays by the flower, so I feel I can eliminate that as well.</p>   |
|        | <p>Since they could smell it far away, the scent particles wouldn't stay near the flower. "B" would still be close to the flower. "C" means they could smell it from away.</p>  |
|        | <p>All the answers have a good explanation, some scent stays near, some moves away, it depends on how strong the scent is. B is wrong because the particles don't go back to where they came from (used a example of dust). D is wrong because the particle spread out not just to the ground. to smell is far away the particles have to spread out and so C is correct.</p>   |

| Question 2 Responses |   |
|----------------------|---|
|                      | I think D, because when the students got closer they could still smell the flower   |
|                      | He reasoned is choice and talked about the particle movement...knowing that the scent particles move father away from each other as distance increases  |
|                      | The student read the questions twice and first answered D then changed to C. The words far away helped you to know the particles move freely.   |
|                      | C because the they can smell it far away, so the scent must be all over.  |
|                      | C because to smell far away the particles have to move and none of the other choice make sense  |
|                      | I would think C because the most they can smell was around 15 m; if it were close, it would be 5 or maybe 10 m; if the particles move freely then they could go the 15 m  |
| Item 4               | We know they could smell the scent particles, so they are going outward not inward - so can take out the first one pointing in. Also, could take out second as the particles are getting mixed in and not on the outside. bottom right is a good one because they all point outward and have nitrogen and oxygen all going out. bottom left - they are all going in different directions - even though they all point outward the bottom right is the better answer   |
|                      | The student described what each model was showing. Since C shows the particles spread farther out, I am choosing C.   |
|                      | This where my information comes in handy: if the flower is farther away the particles would be leaving. I disagree with top right because the particles are too close together, the particles of a gas spread not stay together. I disagree with top left the particles are all going the same direction which wouldn't allow them to spread out. In the lower right the particles are all going different directions, so this is the answer. The lower left is wrong because the nitrogens are moving together and the oxygens are moving away.                                  |
|                      | I think the lower right because all the particles are moving away from the flower   |
|                      | The all gathered as one so the smell.   |
|                      | I would think about how they move the air moves them in different places to the same way.   |
|                      | Picked lower right because the arrows are pointing all different directions so that means that you can smell it from far away   |
| Item 5               | Top right is a pattern. Bottom left because all the air particles are going different way and the first one has a pattern which doesn't make sense. Bottom left doesn't make sense  |
|                      | identified each type of particle; they are either going (described details in each image) ... I would think... if it's an air sample, it's probably a small compact area and they would be bouncing off the walls. selected bottom left choice; evidence - it says an air sample, the particles wouldn't all be going to the center; since it's probably small, they are going random directions  |
| Item 5               | It says to construct an answer using evidence - so we have to remember to use the information that was given to us before. I think it is because the flower scent is a gas. You can't see it but can smell it. Why it varies on the tour - because the different places are in the garden and the distances from the flower and its scent. Farther away from the scent it is harder to smell than when you are closer. It varies because of where you are - I also feel it varies depending on how well you are able to smell - maybe my nose is stuffed up and my friends is not |
|                      | I have to make an explanation using evidence from the text and I also have to say why they could smell it without seeing it and why it varies on the tour. The student went back to each part of the task and described the kind of evidence each part gave. It's a gas that spreads out, the graph shows it can be smelled from away and the farther away the less strong the scent. The particles moved freely away from the flower. Went back to the scenario and pulled out that the smell changed as they moved through. He paid close attention to his writing skills.      |

| Question 2 Responses    |   |
|-------------------------|---|
|                         | Particles spread out so they can be smelled far away. No evidence was cited.  |
|                         | Farther away you can still smell the flower   |
|                         | To me this was just common sense. He talked through that the particles move randomly and reach us before we see the plant but get stronger the closer we get to the plant   |
|                         | Even though we can't see the flower the air makes the scent travel to different places.   |
|                         | the scent travels in different directions   |
|                         | The scent particles spread far so they can smell without seeing, they have three senses and they were only using scent.   |
|                         | rereads; goes back to open images; give where it is, how far away, how strong smell is from where they are... that's mostly it... ok... writes: they can smell the flower from longer distances because the air and scent particles move freely throughout the (what was it called; looks back) the garden, but not completely through it. They could smell it from around the top half of the garden not the bottom. (rereads the directions). When they get closer, the smell gets stronger because they are next to the object that is producing the scent particles. When they get farther, there are less scent particles. utilized the images and graphs to help answer |
| Grade 5, Form B, Task 2 |   |
| Item 1                  | See previous box - entered there  |
|                         | It could be A because the pictures don't match any of the other choices.  |
|                         | I disagree with BC and D because the furniture wasn't moving in the picture and there isn't even a couch in the room. I say A because the sun is rotating and the shadow will get bigger or smaller for the lamp wasn't blocking the tv at 11.  |
|                         | A because if you look at the screen it is darker at 9:00 than at 11:00 because the sun has changed position.  |
|                         | They reasoned this out in the last question that the sun is moving and causing the shadow to change and shorten.  |
|                         | A because the sun isn't fully out at 9 but by 11 it is up in the sky  |
|                         | A because at 9 it was morning still so the sun was low in the sky and not all the way into the window but at 11 the sun was higher and there was no shadow  |
|                         | Which lamp because there are 2 lamps. Why does it say a couch because there is no couch in the picture. In the pic it shows that the shadows shifted at 11 a.m. meaning that the sun moved to a different place. Mostly like A since all of the shadows shifted, there is glare on the TV, whereas before that there was a shadow so there was no glare.  |
| Item 2                  | The first one shows how the sun moves over Nebraska in the different seasons, the second one shows how the earth goes around the sun and how the earth moves, the third one has the sun and the earth moving showing where Nebraska is, the last is the orbit and where the earth is on the orbit. I feel like it is between the 2nd and 3rd because they both show information that is right. The first because it talks about the seasons and they said in Fall.  |
|                         | The first model shows fall. The next model shows the Earth going around the sun. I don't think that would make a difference. The last model shows the moon and it wouldn't make a difference. The first one shows Nebraska and that is where the student lives. I think it is A. Wait, on second thought. I will go with C because it shows rotation and you can see dark or light.   |
|                         | No to the first one because you don't know where the house is. I disagree with the 2nd one because the earth doesn't move around the sun that quickly. I agree with the 3rd one it shows how the axis of sun spins in circles instead of going all the way around the sun.  |

| Question 2 Responses |   |
|----------------------|---|
|                      | I choose the last one: you can see the earth rotating around the sun making it darker and lighter.  |
|                      | I chose the third one cause the pictures show the shadows moving and choice C should the red dot where the tv is. The more rotating of the earth the more shadow changed.   |
|                      | C because it shows what the earth is moving around so the sun stays and the earth spins and hit different spot on it,   |
|                      | I am stuck between the second one and the bottom one because they both show the earth circling the sun. I am choosing the second one because it shows how the earth is shaded different in different seasons.   |
|                      | Observes the images. This one shows there the sun is during the season. next shows difference between night and day. next shows how it goes from night to about 11; last one shows rotation. I think it is (3rd choice) because it shows the position of how Nebraska changes over time toward the sun  |
|                      | The thickest plant is a possibility and the wide shade is a possibility - the narrow items are slimmer. I chose the thick plant because it is taller and wide.  |
|                      | Student read ahead to the text for the next one, didn't see a break between questions. I need to look back at the diagrams. I'm looking at the shadows and I see them get shorter. I don't think the lamp would be big enough. The plant would be the biggest shadow.   |
|                      | NO to a picture frame because it is too small to block the light. No to lamp with a narrower shade because it is skinny, so it won't block as much. No, the plant because it's too small and narrow and the sun won't get blocked and the leaves have a lot of opening to pass light through. No to tree because the leaves will make shade, but the light can still pass through because there are little openings. The answer is wide shade lamp it will block more light and will make a big shadow and th |
| Item 3               | Lamp with wide shade. It will take more of the sun away. The sun would go through the plants and not block the sun as well.   |
|                      | I chose the tree because it is taller so if the sun is higher it will make more shadow on the tv  |
|                      | Scrolled back up to first picture. Chooses the tree because the leaves can block the screen   |
|                      | THE LAMPS ARE THE SAME. FROM THE DIAGRAM ABOVE THE FICUS PLANT COVERS THE TV BEST IT IS WIDER AND WILL PROVIDE MORE SHADE. PICTURE FRAME IS TOO SMALL   |
|                      | top picture shows that the lamp is already on the table; can't be the narrow lamp because it wouldn't fit and probably hit the roof. snake plant wouldn't work because it's too narrow and leaves are spread out, there would be gaps where the glare would still be on TV. ficus plant would probably work because it is wider (and talked about the similarities and differences in both initial pictures.  |
|                      | I feel that D can come out because they will not be the same - there is not as much sun in winter. I feel like it would be shorter because there is more sun in spring than winter  |
| Item 4               | The student looked back at the first model and connected spring to fall. He observed the shadows in the different pictures. He chose C. "I think this would be the correct length because it is like fall".   |
|                      | The shadows will change because the earth will change position and change the shadows. No to shorter because in spring there is more sun, so the shadows will be wider.   |
|                      | C, because a new season means the earth has rotated and a new shadow forms. The spring is rainier so there is less sun.   |

| Question 2 Responses |   |
|----------------------|---|
|                      | Hmmm, I will be C shadows will be shorter because there is less light in winter and lighter in spring.  |
|                      | In the spring the shadows will be wider because the sun will hit them in a different position in the spring   |
|                      | IM NOT SURE ON THIS ONE. MAYBE C BECAUSE THE WINTER HAS CLOUDS TO BLOCK OUT THE SUN   |
|                      | Q2 shows where the Spring will be; Fall and Spring are the same, so selected choice C   |
| Item 5               | A - fall and spring are the same - so not this one. B - Fall to winter - angle increase - so not that one. C - Winter to spring angle decrease - not that one. D - this one because going from summer to winter it decreases.   |
|                      | Analyzed the diagram for the horizon. He used the diagram to see the patterns.  |
|                      | She doesn't reference the diagram much. No to A because summer is super sunny and fall is not as hot. NO to B FROM FALL TO WINTER it would decrease because fall is sunnier than winter. No to C because the spring is more sunny and nicer. D because from summer to winter it gets darker and the earth is rotating around the sun and it gets darker and then lighter.   |
|                      | I don't know which angle the sun has on the horizon. The choices are hard to understand. I think it might be C but I am really not sure   |
|                      | I pick D. I was looking and thinking that the summer sun is higher, so the angle decreases in the winter.   |
|                      | A because the picture shows that the earth is slanting so the sun will hit a different spot of the earth  |
|                      | D summer to winter, in the winter the sun isn't as close  |
|                      | Spring to Fall is the same, can't be A. compares other choices to the graphic. selected D. image shows how far apart each season. so you know if the difference between the horizon and the sun decreases or not. In winter the sun is closer to the horizon, Fall/Spring a little bit farther, Summer it's the farthest.   |
| Item 6               | First one has A to B or fall to winter - but is showing night and day and we are talking about horizon - same with 3rd one - so can take those out. On the other two we are choosing the way the Earth is orbiting. It wouldn't be A to D because I don't think it would orbit that way - Used the season graph to help.  |
|                      | Fall is brighter than winter. I think A on the pics on the left is fall and I think I agree on the upper left because it gets darker and darker as you go around, Fall is first in the question, so it is A and winter gets darker. The right side diagrams are only showing the earth so it's not enough.  |
|                      | It's hard to understand because I don't know what is fall and what is winter in the pictures. I am trying to find the horizon and where fall and winter are. I think lower left because it makes sense, but I don't know how to explain it or what the seasons are.   |
|                      | I chose the lower right and looked at the pictures and imaged the earth going down and it would get darker.   |
|                      | I am confused by the pictures I don't know what to look for. I just picked an answer  |
|                      | Top left. A to B to the sun and then to the dark the earth is moving around   |
|                      | describes each image choice; each image is showing that ""it will be there at that time"". If it's going around and it's spinning... I don't know. Probably need to figure out... (rereads question) ... where the sun will be at that time; one image just shows from day to night... oh, no it's not, because it shows the sun is tilted; it's showing the seasons... so I think the 3rd choice (bottom left); re-explained each model to confirm his thoughts. |

**Question 2 Responses**

|        |   |
|--------|---|
| Item 7 | The first one they similar so I feel like they should be lined up (referred back to the room pictures that showed the two seasons exactly the same) - so it would have to be the same   |
|        | Summer is brighter, so the sun is closer to earth because in winter it is up there more. No to upper left because the sun is at the same height, so the temp would be the same. NO, the lower left because the winter one is closer which would make it warmer and that is wrong. No, the lower right because the sun is getting closer to its destination. I agree with the top right: because the winter sun is farther away and colder.          |
|        | I choose the lower left. Because the summer is closer to the horizon line.  |
|        | ..The lower left More daylight in the summer so it has to be a bit higher in the summer than the winter.  |
|        | lower right because the summer has higher sun and is warmer and winter is cold, and the sun is farther down   |
|        | UPPER RIGHT BECAUSE THE WINTER SUN RISES SLOWER AND IS LOWER BUT IN THE SUMMER THE SUN IS HIGH AND BRIGHT AT 9  |
|        | reads again for clarification (9 a.m.). referred to earlier question, then said can't be ___ these because they (the yellow dot on the red line) aren't lined up; continued to  |
| Item 8 | We are going to have to look back at the different graphics to see how it is in each season - so when we look back we find they are exactly the same  |
|        | I disagree with A the shadows are not the same because there is more sun int he summer. I disagree with B because the shadows in winter are shorter because the sun isn't as bright. I disagree with D because in winter the shadows should be small. I pick C because in summer the sun is hotter and bigger so there is more sun and more light on earth  |
|        | I think B shadows in winter are longer because of the amount of sun and the shadows in the summer are lower because I don't know.   |
|        | I chose C because in the summer the sun goes up more and winter there is not much daylight.   |
|        | C in summer the sun is out longer and shines more while in winter it isn't out as long  |
|        | C because is summer the sun is out so the shadows will be longer and in the winter, there is clouds and little sun so little shadows  |
|        | doesn't say a.m. or p.m.; scrolled back to Q7 and the choice he selected to help answer the question. thinking through... continued to question whether it was am. or p.m. thinking either the shadows are the same or the shadows are longer. I'm thinking the same, because the sun is in the same position, but later on it would be different. A because the sun is in the same position but at other times the sun isn't in the same position. |

**Grade 8, Form A, Task 1**

|        |  |
|--------|--|
| Item 1 | Because the fossils are of shells, I would assume that it had to be an oceanic environment. I would, therefore, think that the layer would be older. If the shells are in a lower rock layer they would have had to form before the higher rock layer.   |
|        | Part A: I notice the shells and horse are found in different locations. I wonder if one location has been around longer than the other location. The fossil horse skeleton probably takes longer to become a fossil whereas the shells could happen more quickly. Part B: I can see the shells are lower and that must be a lower rock layer. I think the stream helped uncover the rock. C could be true because it shows them lower. |
|        | I think the shells are older, because in the diagram the cliff face is where the horse is while the shells are at the bottom - which seems older. I think part B is C and E - C because lower means older and can be exposed like in E where the water exposed the shells.   |

| Question 2 Responses |  |
|----------------------|--|
|                      | Looking at the layers in the fossil horse diagram and trying to determine where the shells would be. Chooses shell layer because it is down lower. Shells were exposed to stream and the fossils weren't because they were up farther in the layers.   |
|                      | A because the diagrams show that the shells are lower than the horse in the layers. Lower = older  |
|                      | Looking at where the items are in the pictures - he figures the shells are older because they are lower. It wouldn't be D - because I think the shells are older. Going with C because the shells were lower than the horse skeleton.  |
| Item 2               | A is subjective. Brittle doesn't necessarily mean a fossil. Found in a rock is good evidence of a fossil. Found near a stream could apply to anything. Found in the oldest rock layer would have a good chance of a fossil, so would youngest, but there wouldn't be as much time to fossilize.                                |
|                      | The horse fossil was found in a rock layer and if the shell is a fossil it would be found in a rock layer as well. E and F not sure which one is the youngest or oldest, so I don't want to pick those. I can rule out the others.   |
|                      | I think it is C but could also be E. A fossil is on a rock or a sheet of Ash - an oldest rock layer would be a harder form and wouldn't break easy. it would be helpful to have the diagram right beside the question and the diagram labeled the ages   |
|                      | It was found inside a rock and it can be seen in the layers on the diagram   |
|                      | Found in oldest rock layer.... farther down means things are older which means they are fossils  |
|                      | Not old - not easily broken. I want to say C - it was found inside a rock  |
| Item 3               | B would have a chance to have other fossils because that is where the ash is. Layer A might be too young. Looking back, I saw where the layer with shells was and this is asking about shells, so my answer is E.  |
|                      | E is the lowest one and would be most close to the stream. The stream plays a part in the shells being in the rock.  |
|                      | I'm thinking that A is water and B is under the water and it keeps getting denser towards the bottom. I think it is A because shells don't usually live on the bottom - more on the top  |
|                      | I think D because it is farther down by the shells in the other diagram, so they are compared across to be in the same place   |
|                      | E because when I scrolled up it showed how far the stream wore the cliff down, so I could match the up   |
|                      | If the horse was found in B and I said shells older I will go with C because it is lower   |
| Item 4               | Prior knowledge of time periods. Those four listed were all oceanic. From what i know I will make the claim, ""C"". Part B: Both choice A and D seem reasonable, but I will choose D because the fossils were marine and now they are land-based.  |
|                      | Part A: The diagram shows that they were ocean animals, so it wouldn't be similar to B. The best answer would be C that it was different. The picture shows that there were marine organisms in the same layer where the shells were.  |
|                      | C best describe question 4 because mammals are not marine animals. The evidence that supports it is A because they change over time from the sediment stirring up over time, pollution and water levels  |
|                      | Trying to decide if I am thinking up and down in the ocean or up and down through time. Layer B is a fossil layer above ground and the new picture is above ground in the ocean, so it is marine but different than layer B. Environments are always changing where you are at so new stuff shows up there at different times. |

| Question 2 Responses    |   |
|-------------------------|---|
|                         | <p>C because the first diagram shows marine creatures and it not the same as the other diagram which shows land animals. Environments change because I'm guessing that the continents have broken up yet and there used to be an ocean where we are now</p>   |
|                         | <p>Read through items... I want to say C because it shows marine animals - so it is not the same as B because the horse doesn't live in the water. Between A and D - because the shells are close to a stream, which is marine kind of. I went with A because it showed the shells in the bottom.</p>   |
| Item 5                  | <p>A and B can be ruled out because nothing in this task has been related to temperature. D is incorrect, but you would need to read the diagram correctly, from bottom to top. The oldest is on the bottom so the correct answer is C.</p>   |
|                         | <p>I can see the bottom is the most years ago and they are ocean organisms, then I see it change as I go up the chart. I will pick C because it shows ocean, to swamp, to prairie.</p>  |
|                         | <p>C describes the changes because the largest number is on the bottom and it looks like it is getting closer to current days at the top</p>  |
|                         | <p>Ocean, to swamp to prairie. Fish first than alligator than dinosaurs so getting more land based.</p>   |
|                         | <p>Ocean, swamp to prairie because in the diagram it show aquatic creature and then reptiles and dinosaurs, so I figured out their environment</p>  |
|                         | <p>Diagram shows geologic time, read items, animals at bottom are more marine up to alligator to horses. I'm going with C</p>   |
| Item 6                  | <p>D is ruled out because the chart stops at 541 million years. B is when the dinosaurs etc. lived. C has to be the answer. Part B: Choosing A rules out the other choices because the organisms at that time have to be ruled out. Older time periods wouldn't have as much life like the shells since not much oceanic life (it was unicellular at that time)</p>   |
|                         | <p>They were more than 11 million years old based on the diagram. They are probably between 200 million and 541.</p>  |
|                         | <p>C is the answer because it approximates from 541 (the longest) to 200 million years ago. B supports that because the longer the time the more soil that goes on top of it.</p>   |
|                         | <p>Re-reading question to determine if it says found or formed, they are younger than 11 million years because they were formed 11 million not found so they aren't that old. Found in Ne streams today because they are still finding stuff at Ash falls, we just went there in 6th grade :)</p>   |
|                         | <p>It took a while between dinosaur age and mammal age so there was a long time to make this change, many years. Shells are at the bottom, so I know they are older so that's why I picked the older dates. the rock layers build on tops.</p>  |
|                         | <p>Going back to the chart of where they found the fossils - so can eliminate younger. Would say it is not over 541 million - that is too much. I would go with 11-200 million because of the graph that showed that horse was 11 million - 541 is too far - so eliminated and thought it seemed about right based on chart. The shells are in the bottom rock layer and older than the horse layer - so I am going with B.</p> |
| Grade 8, Form A, Task 2 |   |
| Item 1                  | <p>Small up and large up are subjective depending on how fast you want to lift it. Student is working on the math behind solving the problem. I'm going to take the large out but I'm unsure because of how it is worded.</p>   |
|                         | <p>For A small up because it is closer to the pedestal. For D large up because I assume it is farther down from the top of the pedestal. For B and C, I have to assume that it is moving it inward.</p>   |
|                         | <p>I don't know - It is supposed to be lifted upward - but I'm not sure how to figure that out. I think all of them should be a small up. It says it needs to go upward - so down arrow doesn't make</p>  |

| Question 2 Responses |   |
|----------------------|---|
|                      | sense. Maybe if large on A, small down on D in case you lifted it too high, so you can bring it down some, then small up on B and small right on C, because it needs to move on sculpture, so it needs to go over to sit on base  |
|                      | D will push up toward the air large. A will push up small. It take more to push you off the ground than it does to keep you up. B small left. C small right. This will create balance so it doesn't fall.   |
|                      | D is large down because the sculpture is heavy and will push down on the pedestal. A is small up because is won't put as much force up because weight pulls you down. B small left because it will put a little pressure over there but not a lot. C small right because it will put some pressure but not a lot  |
|                      | A would be large up because you need more force to pick it up, so it takes more energy, B would be large left because you have to move the whole thing, C would be large right to get it there, D would be small down because you have gravity helping you  |
| Item 2               | Gravity is pulling down on the mass which accounts for the newtons. When the kg increase the n increases. gravity affects the mass, the mass doesn't affect the gravity. It is a constant. Unless you are great distance. A can't be correct mathematically. D has to be correct. Part B: Mentally calculated. Used the chart to know to multiply by 10 to get the Newtons.   |
|                      | I am saying D is correct because it shows that as the mass gets greater, the applied force does as well. Part B I would add the 40 plus 10 kg and them multiply that by 10 like they did in the data table.   |
|                      | Two of the answers are good because they seem to match it but I think D is better because you wouldn't take a KG and multiply by Newton because Newton is a force   |
|                      | B or D, because the mass increases with the forces so it does go up but the greater the mass increase the force. Chooses D because it sounds more like a scientific explanation. The base is 40 and sculpture is 10 so together that is 50 then times 10 to get Newtons so 500N   |
|                      | A mass is 10x's the force in the chart (the data is multiplied by 10). He didn't read it very well! I scrolled up to the find the force of the robot was 1000 N, so I know to take that times 10 so the answer is 10,000  |
|                      | A - would be D because it is getting heavier, so it has to apply more force (pointing to data chart). B - looked back to weight of the butterfly and base (50 kg). 50 X 10 is 500 newtons   |
| Item 3               | A is correct but not what we are looking for. B is correct. D can automatically be eliminated. C the table does not support this and that is not how physics works.   |
|                      | the kg is 10 and the N is 10 and acceleration is 1. I notice that the N and the M/s/s have a relationship. They are the same without the ""0"" at the end. The answer is B because you can see that they both increase. A and D show mass but the question doesn't ask that.  |
|                      | I think B because each time the unbalanced force goes up the acceleration of the object goes up as well.  |
|                      | Reading numbers and thinking about relationship. I am picking b unbalanced increases acceleration according to the chart and the kg stays the same  |
|                      | acceleration increases the mass stays the same on the chart is shows this relationship. Not C the answer didn't make sense to me, so I didn't pick it. The force increases  |
|                      | The more force they added the faster it got (looking at the data chart). I'm going with B because every time they added newtons it went faster  |
| Item 4               | $F = ma$ . you have the mass with the correlating N. D is correct because it shows the relationship between force, motion, and mass. Used the formula to reason choice C. Both C and D seem correct mathematically. I don't know what I would choose.   |
|                      | I will rule out A and B because the motion doesn't stay the same. It changes. I think the answer is D because if the mass and force are increasing by the same then it should balance out the change in the motion. Part B: The answer is B because I plugged in the data from the chart to the formula and if the mass was 10 and acceleration was 5, then the force was 50. |

| Question 2 Responses |   |
|----------------------|---|
|                      | I think it is C because newtons for each kg increase and it makes sense. In the top one it changes the force. There are not 8 newtons and it would be different anyway. It is also not multiplied by 2. Part b is A because acceleration is mass divided by force   |
|                      | Reading through choices very carefully. Trying to figure out what the motion is? Acceleration? Thinking force is motion. Lots of confusion. Chooses A but explanation is unclear. Part B: Thinks a in the equation is area? $10 \times 5 = 50$ for B works  |
|                      | Trying to compare all three charts to make relationship. A doesn't work with the charts. B doesn't work because motion is less than the newtons, so it wouldn't stay the same. Only consulting one chart at a time? D because in the table if it increases by the force must increase by two to keep it the same one goes up and the other goes up according to the table. Part B: $F = m \times a$ because the force is equal to mass times acceleration according to chart $1 \times 10 = 10$ |
|                      | A - I am going with D because if it increased by 2 times it would have to increase by 2 times to keep it balanced. B - A because you are trying to find the acceleration of the object  |
| Item 5               | We know $F = ma$ and the force is 1000 N. The force is constant and always has to stay 1000. It doesn't seem right because it said previously the safe acceleration is 2m/s/s. The answers don't seem reasonable in the situation but using the formula this is what I get.   |
|                      | if the force = the mass times the acceleration and the force is 1000, then I just have to put in the mass to figure it out. The safe is acceleration rate is 2 m/s/s.   |
|                      | Mass X acceleration would make the 1st one 40 m/s squared, then 100 m/s squared and 500 m/s squared   |
|                      | This is tough. Trying to determine where i get data, trying to use the chart instead of the equation. She's not doing the math. Using the chart data in the equation instead of the info in the question. Kg in chart are 10 and force changes? 40 isn't on the chart? I don't know how to do it. $40 \text{ kg} \times 10 \text{ kg} = 400$ . $10 \text{ kg} \times 10 \text{ kg} = 100$ . $50 \times 10 = 500$ .  |
|                      | He is using the chart not the equation. I put 400000, I took 1000N unbalanced and took it times 400 because of the 40.... He got lost in this thinking and couldn't finish is explanation. I said 10 for the second part because if $f = m \times a$ . Take $10 \times 1 = 10$ ., comes from the chart. I said 600000 I took $50 + 10$ and then multiplied by 10,000 to get the answer.   |
|                      | Every 10 is 1 so it would be 4 S squared. 1 second squared. 5 second squared for the last. using the data chart to help   |
| Item 6               | Increasing the mass to lower the acceleration to a safe point. I could use the formula. I used the information that the robot always applies 1000 N of unbalanced force.  |
|                      | I think it is safer to place the sculpture and the base on together because the acceleration rate is slower that way. I used my calculations from prior questions and if they were loaded separately it wouldn't be as safe.  |
|                      | I am wondering what the best solution is because the acceleration for the 40 kg is slow and I'm not sure if that is what they want, or if they want it to go quick.   |
|                      | robot can do 1000N and the object is 500N so it can easy to pick up because this is half the weight   |
|                      | Formulated a lengthy response. Need right amount of force for robot and pedestal to make sure they don't break. Calculate max force for the robot. Make sure all the right angles for robot to get it up to base and make sure it doesn't break. No evidence  |
|                      | If I was doing it I would put the sculpture base on the pedestal first and then put the butterfly on after. I would do that because it would be the safest was to move it. Otherwise you would risk moving two pieces together and risk it falling in the transition. 1 at a time would be safer.   |

**3. The student's comments provided evidence that the item elicited thinking about:**

1. [varies by item]
2. [varies by item]
3. Other

**Table I.3. Question 3 Responses**

| Item #                         | Response #  | N | %*   |
|--------------------------------|---|---|------|
| <b>Grade 5, Form B, Task 1</b> |   |   |      |
| Item 1<br>(9 students)         | 1. Scent can be measured scientifically.  | 1 | 11.1 |
|                                | 2. Relate the amount of scent to the distance from the source.  | 8 | 88.9 |
|                                | 3. Other  | 2 | 22.2 |
| Item 2<br>(9 students)         | 1. Gas particles are widely spread.   | 3 | 33.3 |
|                                | 2. Use models to describe phenomena.  | 4 | 44.4 |
|                                | 3. Other  | 6 | 66.7 |
| Item 3<br>(9 students)         | 1. Gas particles randomly spread from a concentrated area.  | 8 | 88.9 |
|                                | 2. Construct an explanation of an observed relationship between a model and a phenomenon.   | 2 | 22.2 |
|                                | 3. Other  | 1 | 11.1 |
| Item 4<br>(9 students)         | 1. Gas particles move randomly in all directions.   | 7 | 77.8 |
|                                | 2. Relate a model to a phenomenon.  | 5 | 55.6 |
|                                | 3. Other  | 1 | 11.1 |
| Item 5<br>(9 students)         | 1. Scent particles are a gas that can be invisible, widely spaced, and move randomly in all directions becoming less dense the further from the source. | 4 | 44.4 |
|                                | 2. Use evidence to construct an explanation with supporting evidence.   | 3 | 33.3 |
|                                | 3. Other  | 4 | 44.4 |
| <b>Grade 5, Form B, Task 2</b> |   |   |      |
| Item 1<br>(8 students)         | 1. There are daily changes in the length and direction of shadows.  | 5 | 62.3 |
|                                | 2. Related patterns in observations to support an explanation.  | 6 | 75.0 |
|                                | 3. Other  | – | –    |
| Item 2<br>(8 students)         | 1. The orbits of Earth around the Sun together with the rotation of Earth about an axis cause the daily changes in the length and direction of shadows. | 5 | 62.5 |
|                                | 2. Determine the best model to describe a phenomenon.   | 4 | 50.0 |
|                                | 3. Other  | 1 | 12.5 |
| Item 3<br>(8 students)         | 1. Determining a solution to a problem.   | 2 | 25.0 |
|                                | 2. Use patterns in observations to predict a solution to a problem.   | 7 | 87.5 |
|                                | 3. Other  | 1 | 12.5 |
| Item 4<br>(8 students)         | 1. The orbits of Earth around the Sun cause changes in the length and direction of shadows at different times of the year.                              | 4 | 50.0 |
|                                | 2. Predict changes based on patterns and cause and effect relationships.  | 3 | 37.5 |
|                                | 3. Other  | 2 | 25.0 |
| Item 5<br>(7 students)         | 1. The orbits of Earth around the Sun cause observable patterns from different times of the year.   | 5 | 71.4 |
|                                | 2. Determine the best model to predict changes in a phenomenon.   | 2 | 28.6 |
|                                | 3. Other  | 2 | 28.6 |

Appendix I: Cognitive Lab Responses

| Item #                         | Response #   | N | %*    |
|--------------------------------|--|---|-------|
| Item 6<br>(7 students)         | 1. The orbits of Earth around the Sun together with the rotation of Earth about an axis between its North and South poles cause observable patterns.                                   | 2 | 28.6  |
|                                | 2. Models can be used to support an explanation.   | 3 | 42.9  |
|                                | 3. Other   | 3 | 42.9  |
| Item 7<br>(6 students)         | 1. The orbits of Earth around the Sun together with the rotation of Earth about an axis between its North and South poles cause observable patterns.                                   | 1 | 16.7  |
|                                | 2. Determine the best model to explain a phenomenon.   | 3 | 50.0  |
|                                | 3. Other   | 2 | 33.3  |
| Item 8<br>(7 students)         | 1. The orbits of Earth around the Sun causes changes in the length and direction of shadows at different times of the year.  | 1 | 14.3  |
|                                | 2. Predict changes based on patterns and cause and effect relationships.   | 4 | 57.1  |
|                                | 3. Other   | 2 | 28.6  |
| <b>Grade 8, Form A, Task 1</b> |  |   |       |
| Item 1<br>(6 students)         | 1. A fossil record can indicate the chronological order of fossils.  | 6 | 100.0 |
|                                | 2. Determine empirical evidence to support a claim.  | 3 | 50.0  |
|                                | 3. Other   | 1 | 16.7  |
| Item 2<br>(6 students)         | 1. Fossils are usually located of sedimentary layers where they are formed.  | 3 | 50.0  |
|                                | 2. Determine scientific reasoning to support identifying fossils.  | 2 | 33.3  |
|                                | 3. Other   | 3 | 50.0  |
| Item 3<br>(6 students)         | 1. Analyses of rock strata can indicate likelihood of finding similar fossils.   | 3 | 50.0  |
|                                | 2. Use a graphical display of a large data set to identify spatial relationships.  | 6 | 100.0 |
|                                | 3. Other   | 1 | 16.7  |
| Item 4<br>(6 students)         | 1. The fossil record, documents the existence, diversity, extinction, and change of many life forms throughout the history of life on Earth.   | 5 | 83.3  |
|                                | 2. Use graphical displays to identify temporal and spatial relationships.  | 3 | 50.0  |
|                                | 3. Other   | 1 | 16.7  |
| Item 5<br>(6 students)         | 1. The fossil record, documents the existence, diversity, extinction, and change of many life forms throughout the history of life on Earth.   | 3 | 50.0  |
|                                | 2. Use graphical displays to identify temporal and spatial relationships.  | 6 | 100.0 |
|                                | 3. Other   | – | –     |
| Item 6<br>(6 students)         | 1. Analysis of rock strata and the fossil record provide relative dates.   | 2 | 33.3  |
|                                | 2. Construct a scientific explanation based on valid and reliable evidence.  | 5 | 83.3  |
|                                | 3. Other   | 1 | 16.7  |
| <b>Grade 8, Form A, Task 2</b> |  |   |       |
| Item 1<br>(6 students)         | 1. The motion of an object is determined by the sum of the forces acting on it. For any given object, a larger force causes a larger change in motion.                                 | 4 | 66.7  |
|                                | 2. Use a model to describe a phenomenon.   | 4 | 66.7  |
|                                | 3. Other   | 1 | 16.7  |
| Item 2<br>(6 students)         | 1. The motion of an object is determined by the sum of the forces acting on it. The greater the mass of the object, the greater the force needed to achieve the same change in motion. | 4 | 66.7  |
|                                | 2. Determine causal relationships in data.   | 3 | 50.0  |
|                                | 3. Other   | 1 | 16.7  |

Appendix I: Cognitive Lab Responses

| Item #                 | Response #   | N | %*    |
|------------------------|--|---|-------|
| Item 3<br>(6 students) | 1. The motion of an object is determined by the sum of the forces acting on it, a larger force causes a larger change in motion. | 3 | 50.0  |
|                        | 2. Determine causal relationships in data.   | 6 | 100.0 |
|                        | 3. Other   | 1 | 16.7  |
| Item 4<br>(6 students) | 1. Newton's second law ( $F = ma$ ) accurately predicts changes in the motion of objects.  | 4 | 66.7  |
|                        | 2. Determine causal relationships in data.   | 3 | 50.0  |
|                        | 3. Other   | 2 | 33.3  |
| Item 5<br>(6 students) | 1. Newton's second law ( $F = ma$ ) accurately predicts changes in the motion of objects.  | 3 | 50.0  |
|                        | 2. Apply mathematical concepts and processes to scientific and engineering questions and problems.                               | 5 | 83.3  |
|                        | 3. Other   | 1 | 16.7  |
| Item 6<br>(6 students) | 1. Newton's second law ( $F = ma$ ) accurately predicts changes in the motion of objects.  | 2 | 33.3  |
|                        | 2. Apply scientific ideas or principles to design, construct, and/or test a design of an object, tool, process, or system.       | 4 | 66.7  |
|                        | 3. Other   | 2 | 33.3  |

\*Percentages are out of the number of students who responded per item. Proctors could indicate more than one answer for a student, so the percentages may not add up to 100.0.

## 4. If “other” what thinking did the item elicit?

Table I.4. Question 4 Responses

| Question 4 Responses   |   |
|--|---|
| Grade 5, Form B, Task 1  |   |
| Item 1   | no other  |
|  | The student knows the scent is stronger but does not talk about it mixing with non-scent particles  |
|  | Related to distance but explanation was unsure  |
|  | Confusion of chart and diagram  |
| Item 2   | Looked back at graphs and looked at shape of particles and thought must be the grouped because the scent is strong and they all have the same number of particles   |
|  | Reasoning   |
|  | She was thinking the smell was from the flower and the flower is a solid instead of thinking that the smell was a gas and how those particles distributed. Struggled to determine the state of matter for the smell but settled on gas, was really hung up on the word particles. |
|  | Lacks understand of the nature of gases and how the diagram relates   |
|  | He stated it should be a gas as the particles were moving around but then he thought not maybe a solid...the model diagrams were more confusing than helpful. as it didn't fit his view of a model of this  |
| Confusion on the gas particle model not a lot of coherent thought on this question |   |
| Item 3   | confusion about smell and distance  |
| Item 4   | Particles move outward so scent can be noticed  |
| Item 5   | No citation of evidence   |
|  | The student thinks of the air moving the scent particles.   |
|  | No evidence or support did not address all areas of the question, no science behind her explanation   |
|  | Scent particle movement only with no evidence   |
| Grade 5, Form B, Task 2  |   |
| Item 1   | <i>No answers provided.</i>   |
| Item 2   | Understands that the earth is moving relative to the sun but lacks some information on path, tilt and process   |
| Item 3   | didn't connect the shadow size the object   |
| Item 4   | The student chose C, but they were thinking about the amount of sun not the length and direction of the sun.  |
|  | SHE DIDN'T KNOW   |
| Item 5   | Was basically unable to answer the question.  |
| Item 6   | Little connection to rotation in the diagrams.  |
|  | I think the student was just looking at the arrows and she said the earth went down (maybe angled down?)  |
| Item 7   | Used the pictures from previously that showed the two different seasons with the shadows at 11:00 am exactly the same   |
|  | Confusion and lack of explanation   |

| Question 4 Responses    |  |
|-------------------------|--|
| Item 8                  | using pictures from before that showed shadows the same in the different seasons   |
|                         | No connection to sun and earth relationship  |
|                         | The student was giving a cause and effect but did not talk out the orbits and changes in length and direction.   |
| Grade 8, Form A, Task 1 |  |
| Item 1                  | Considering the oceanic environment.   |
| Item 2                  | eliminating other responses.   |
|                         | Used depth of layers - but did not talk about how they build on each other found in rock   |
| Item 3                  | Using prior knowledge.   |
| Item 4                  | Background knowledge of past time periods.   |
| Item 5                  | <i>No answers provided.</i>  |
| Item 6                  | Utilizes the graphics  |
| Grade 8, Form A, Task 2 |  |
| Item 1                  | Didn't have a way of identifying what to do here - talked her way through it some  |
| Item 2                  | mathematical reasoning.  |
|                         | Utilized elimination of some and then determined between them  |
| Item 3                  | Test taking strategy of understanding the question.  |
| Item 4                  | using mathematical thinking  |
|                         | elimination of answers   |
| Item 5                  | Multiplied by the force rather than reworking the formula to calculate   |
| Item 6                  | Understood the numbers on this one - but wanted it to move faster, did not recognize the fragile nature of the work that would require the slower speed - this affected the answer and choose a mid-level acceleration |
|                         | cause and effect relationship  |

- 5. While reading and interacting with the question, the student appeared to...**
- 1. rush to answer the question without careful consideration of the question, the text, and how to respond.**
  - 2. be engaged with the question, taking time to consider and respond.**
  - 3. be able to understand how to generate a response to the question.**

**Table I.5. Question 5 Responses**

| <b>Item #</b>                  | <b>Response #</b> | <b>N</b> | <b>%*</b> |
|--------------------------------|-------------------|----------|-----------|
| <b>Grade 5, Form B, Task 1</b> |                   |          |           |
| Item 1<br>(9 students)         | 1                 | 2        | 22.2      |
|                                | 2                 | 6        | 66.7      |
|                                | 3                 | 3        | 33.3      |
| Item 2<br>(9 students)         | 1                 | –        | –         |
|                                | 2                 | 9        | 100.0     |
|                                | 3                 | 2        | 22.2      |
| Item 3<br>(9 students)         | 1                 | 3        | 3.33      |
|                                | 2                 | 4        | 4.44      |
|                                | 3                 | 4        | 4.44      |
| Item 4<br>(9 students)         | 1                 | 2        | 22.2      |
|                                | 2                 | 7        | 77.8      |
|                                | 3                 | 2        | 22.2      |
| Item 5<br>(9 students)         | 1                 | 2        | 22.2      |
|                                | 2                 | 8        | 88.9      |
|                                | 3                 | 2        | 22.2      |
| <b>Grade 5, Form B, Task 2</b> |                   |          |           |
| Item 1<br>(8 students)         | 1                 | 2        | 25.0      |
|                                | 2                 | 6        | 75.0      |
|                                | 3                 | 3        | 37.5      |
| Item 2<br>(8 students)         | 1                 | –        | –         |
|                                | 2                 | 8        | 100.0     |
|                                | 3                 | 3        | 37.5      |
| Item 3<br>(8 students)         | 1                 | 1        | 12.5      |
|                                | 2                 | 5        | 62.5      |
|                                | 3                 | 5        | 62.5      |
| Item 4<br>(8 students)         | 1                 | 1        | 12.5      |
|                                | 2                 | 6        | 75.0      |
|                                | 3                 | 3        | 37.5      |
| Item 5<br>(7 students)         | 1                 | –        | –         |
|                                | 2                 | 7        | 100.0     |
|                                | 3                 | 2        | 28.6      |
| Item 6<br>(7 students)         | 1                 | 1        | 14.3      |
|                                | 2                 | 6        | 85.7      |
|                                | 3                 | 2        | 28.6      |
| Item 7<br>(7 students)         | 1                 | –        | –         |
|                                | 2                 | 7        | 100.0     |
|                                | 3                 | 2        | 28.5      |

| Item #                         | Response # | N | %*    |
|--------------------------------|------------|---|-------|
| Item 8<br>(6 students)         | 1          | 1 | 16.7  |
|                                | 2          | 5 | 83.3  |
|                                | 3          | 2 | 33.3  |
| <b>Grade 8, Form A, Task 1</b> |            |   |       |
| Item 1<br>(6 students)         | 1          | – | –     |
|                                | 2          | 5 | 83.3  |
|                                | 3          | 5 | 83.3  |
| Item 2<br>(6 students)         | 1          | – | –     |
|                                | 2          | 5 | 83.3  |
|                                | 3          | 5 | 83.3  |
| Item 3<br>(6 students)         | 1          | 1 | 16.7  |
|                                | 2          | 5 | 83.3  |
|                                | 3          | 4 | 66.7  |
| Item 4<br>(6 students)         | 1          | – | –     |
|                                | 2          | 5 | 83.3  |
|                                | 3          | 5 | 83.3  |
| Item 5<br>(6 students)         | 1          | 1 | 16.7  |
|                                | 2          | 5 | 83.3  |
|                                | 3          | 4 | 66.7  |
| Item 6<br>(6 students)         | 1          | – | –     |
|                                | 2          | 6 | 100.0 |
|                                | 3          | 4 | 66.7  |
| <b>Grade 8, Form A, Task 2</b> |            |   |       |
| Item 1<br>(6 students)         | 1          | – | –     |
|                                | 2          | 6 | 100.0 |
|                                | 3          | 3 | 50.0  |
| Item 2<br>(6 students)         | 1          | 1 | 16.7  |
|                                | 2          | 5 | 83.3  |
|                                | 3          | 3 | 50.0  |
| Item 3<br>(6 students)         | 1          | – | –     |
|                                | 2          | 6 | 100.0 |
|                                | 3          | 4 | 66.7  |
| Item 4<br>(6 students)         | 1          | – | –     |
|                                | 2          | 6 | 100.0 |
|                                | 3          | 3 | 50.0  |
| Item 5<br>(6 students)         | 1          | – | –     |
|                                | 2          | 6 | 100.0 |
|                                | 3          | 3 | 50.0  |
| Item 6<br>(6 students)         | 1          | – | –     |
|                                | 2          | 6 | 100.0 |
|                                | 3          | 3 | 50.0  |

\*Percentages are out of the number of students who responded per item. Proctors could indicate more than one answer for a student, so the percentages may not add up to 100.0.

**6. Ask the student: Was it easy or hard to know how to respond to the question?**

1. Easy
2. Medium
3. Hard

**Table I.6. Question 6 Responses**

| Item #                         | Response # | N | %*   |
|--------------------------------|------------|---|------|
| <b>Grade 5, Form B, Task 1</b> |            |   |      |
| Item 1<br>(9 students)         | 1          | 7 | 77.8 |
|                                | 2          | 2 | 22.2 |
|                                | 3          | – | –    |
| Item 2<br>(9 students)         | 1          | – | –    |
|                                | 2          | 3 | 33.3 |
|                                | 3          | 6 | 66.7 |
| Item 3<br>(9 students)         | 1          | 6 | 66.7 |
|                                | 2          | 3 | 33.3 |
|                                | 3          | – | –    |
| Item 4<br>(9 students)         | 1          | 5 | 55.6 |
|                                | 2          | 3 | 33.3 |
|                                | 3          | 1 | 11.1 |
| Item 5<br>(9 students)         | 1          | 5 | 55.6 |
|                                | 2          | 4 | 44.4 |
|                                | 3          | – | –    |
| <b>Grade 5, Form B, Task 2</b> |            |   |      |
| Item 1<br>(8 students)         | 1          | 7 | 87.5 |
|                                | 2          | 1 | 12.5 |
|                                | 3          | – | –    |
| Item 2<br>(8 students)         | 1          | 1 | 12.5 |
|                                | 2          | 5 | 62.5 |
|                                | 3          | 2 | 25.5 |
| Item 3<br>(8 students)         | 1          | 6 | 75.0 |
|                                | 2          | 2 | 25.0 |
|                                | 3          | – | –    |
| Item 4<br>(8 students)         | 1          | 4 | 50.0 |
|                                | 2          | 2 | 25.0 |
|                                | 3          | 2 | 25.0 |
| Item 5<br>(7 students)         | 1          | 3 | 42.9 |
|                                | 2          | 3 | 42.9 |
|                                | 3          | 1 | 14.3 |
| Item 6<br>(7 students)         | 1          | – | –    |
|                                | 2          | 3 | 42.9 |
|                                | 3          | 4 | 57.1 |
| Item 7<br>(7 students)         | 1          | 5 | 71.4 |
|                                | 2          | 1 | 14.3 |
|                                | 3          | 1 | 14.3 |

| Item #                         | Response # | N | %*    |
|--------------------------------|------------|---|-------|
| Item 8<br>(6 students)         | 1          | 4 | 57.1  |
|                                | 2          | 1 | 14.3  |
|                                | 3          | 2 | 28.6  |
| <b>Grade 8, Form A, Task 1</b> |            |   |       |
| Item 1<br>(6 students)         | 1          | 3 | 50.0  |
|                                | 2          | 2 | 33.3  |
|                                | 3          | 1 | 16.7  |
| Item 2<br>(6 students)         | 1          | 2 | 33.3  |
|                                | 2          | 4 | 66.7  |
|                                | 3          | – | –     |
| Item 3<br>(6 students)         | 1          | 5 | 83.3  |
|                                | 2          | – | –     |
|                                | 3          | 1 | 16.7  |
| Item 4<br>(6 students)         | 1          | 2 | 33.3  |
|                                | 2          | 3 | 50.0  |
|                                | 3          | 1 | 16.7  |
| Item 5<br>(6 students)         | 1          | 6 | 100.0 |
|                                | 2          | – | –     |
|                                | 3          | – | –     |
| Item 6<br>(6 students)         | 1          | 4 | 66.7  |
|                                | 2          | 2 | 33.3  |
|                                | 3          | – | –     |
| <b>Grade 8, Form A, Task 2</b> |            |   |       |
| Item 1<br>(6 students)         | 1          | – | –     |
|                                | 2          | 2 | 33.3  |
|                                | 3          | 4 | 66.7  |
| Item 2<br>(6 students)         | 1          | 3 | 50.0  |
|                                | 2          | 1 | 16.7  |
|                                | 3          | 2 | 33.3  |
| Item 3<br>(6 students)         | 1          | 4 | 66.7  |
|                                | 2          | 2 | 33.3  |
|                                | 3          | – | –     |
| Item 4<br>(6 students)         | 1          | – | –     |
|                                | 2          | 4 | 66.7  |
|                                | 3          | 2 | 33.3  |
| Item 5<br>(6 students)         | 1          | 1 | 16.7  |
|                                | 2          | 1 | 16.7  |
|                                | 3          | 4 | 66.7  |
| Item 6<br>(6 students)         | 1          | 2 | 33.3  |
|                                | 2          | 3 | 50.0  |
|                                | 3          | 1 | 16.7  |

\*Percentages are out of the number of students who responded per item.

## 7. Ask the student: What would help you answer the question?

Table I.7. Question 7 Responses

| Question 7 Responses    |  |
|-------------------------|--|
| Grade 5, Form B, Task 1 |  |
| Item 1                  | There is a lot of graphics and pictures to help you. If you really think about how it is organized, you can figure it out by how far you are apart from the flower.  |
|                         | It gave a graph and a diagram so that helped. The directions to drag and drop were pretty easy.  |
|                         | Nothing  |
|                         | Nothing  |
|                         | Diagram helped with top, medium and bottom   |
|                         | You could see that close it was strong and farther it was not on the graph.  |
|                         | You had to think about it a lot  |
|                         | I thought of it as a chart top to bottom, bottom is the worst best is the top  |
| Item 2                  | the diagram showed how far everything is apart to give more information to answer the question   |
|                         | It would help to know if we are using the graphs above or if there was a graph that I didn't have to scroll to. The nitrogen and oxygen confused me because there is no other information about that - so it just makes it harder.   |
|                         | You need a lot of knowledge to know how to find the answer.  |
|                         | I didn't understand the patterns.  |
|                         | The map and measurement graph helped the model choices were confusing.   |
|                         | The pictures   |
|                         | I didn't understand the pictures or the questions  |
| Item 3                  | I want to know where the flower is in the gas particle diagrams.   |
|                         | it wants you to go back to previous information instead of just what is in the question; you have to know how particles work together in order to answer. enough info to answer the questions, but the questions are a bit difficult to think about how the particles are together or aren't, what they are, how they should be. it would help if they gave the space that the particles were in... is the space 5m all around, in only two directions... we don't know about that given the information |
|                         | Not really - The question gives the information that I need. Don't need a picture, information is in the item, so I can think it in my head  |
|                         | There isn't a lot of information in the chart about where the particles went to, students have to know about gases to answer this.   |
|                         | Nothing  |
|                         | Nothing else needed  |
| Item 3                  | The diagram and the words far away.  |
|                         | gave the information on where they are smelling it from and exactly how far away they are  |

| Question 7 Responses    |  |
|-------------------------|--|
| Item 4                  | It was kind of easy - but not the hardest one. It was not easy because my background was bigger on some of the other items like matter -   |
|                         | There weren't very many scent particles shown in the model. If there had been more, he thinks he could see the pattern of movement better.   |
|                         | I have a lot of knowledge about this already.  |
|                         | The picture helped me to think about it.   |
|                         | talks about an air sample instead of a whole room; answer would be completely different for a whole room; since it is a sample, it would be more compact and tighter.  |
| Item 5                  | They told us we could construct from the graphics - so we have a lot to help us. It goes back to all of the questions before to answer this one. Don't need anything else - there is a lot of evidence from before that helps. |
|                         | Good directions.   |
|                         | Tools (graphs, pics...) provided helped me explain the answer  |
|                         | It was easy to explain about the scent particles but not the strength of the scent.  |
|                         | I just knew it.  |
|                         | I have had the experience and the distance and map helped.   |
|                         | It wants you to look back at everything that has been given and tell how and why the smell gets stronger when people get closer to the plant.  |
| Grade 5, Form B, Task 2 |  |
| Item 1                  | If there was a couch in the picture. Knowing the position of the room.   |
|                         | Easy because we have learned about shadows and we have the knowledge.  |
|                         | The pictures.  |
|                         | have to know about how the sun changes and how a TV will glare or not  |
| Item 2                  | I feel like the information is there - it is just confusing  |
|                         | More of a description of each model. Labels on the first one but not on the others.  |
|                         | You have to know about space and the earth and the sun and how fast it all goes.   |
|                         | We learned about this in class and how it get darker and lighter.  |
|                         | I had to reason out each picture.  |
|                         | have to know about how Earth rotates and how the sun goes across the sky over time. It doesn't stay in one place because the Earth is rotating around the sun.   |
| Item 3                  | Nothing really   |
|                         | have to know where the shadow is going to land on the TV. might be helpful to know where the sun is in the background of the pictures  |
| Item 4                  | The graphic looks the same as the one in the fall - so maybe another graphic showing the difference there -  |
|                         | Spring can be sunny but not always so that is tricky.  |
|                         | if you go to info previously given, it will give you your answer a bit   |
| Item 5                  | Nothing needed   |
|                         | If you know how the earth moves, then you know that in NE the weather is colder and less sunny in the winter because you are farther away from the sun.  |

| Question 7 Responses    |   |
|-------------------------|---|
|                         | have to know about how far apart they (horizon and sun) are... have to use your information; the model and the question, 11:00 a.m. -- all important pieces of information in order to answer |
| Item 6                  | Maybe putting a sun on the two that don't have it so we can see which way it is orbiting  |
|                         | Maybe...I don't know.   |
|                         | have to know when the Earth tilts, how it tilts, where each season is; the question wants to know between Fall and Winter; have to know how each model shows each season                      |
| Item 7                  | nothing needed  |
|                         | previous information from earlier question about what it should look like at a later time; so you should know that it will look similar but at a different position                           |
| Item 8                  | nothing   |
|                         | very situational question -- it depends on where the shadow would be, where the room is; the time (a.m. or p.m.) -- they don't give you enough info to make it an easier question             |
| Grade 8, Form A, Task 1 |   |
| Item 1                  | There were lots of different pictures and I wasn't sure how to connect them to each other.  |
|                         | The questions make sense but some of the answers were different - like they don't explain why   |
|                         | If you look at the diagrams and re-read it's easy to understand what they are talking about   |
|                         | More charts - I like those. How they were found might help  |
| Item 2                  | I understood the question and the pictures provided good help, but I didn't know how to use them to link the fossil and the shell.  |
|                         | If it came out directly and told you the ages of the layers   |
|                         | No  |
| Item 3                  | Less choices to choose from, more specific answers with better description of why it was found inside a rock  |
|                         | Having access to the previous question resources. Stating to myself over and over what I am trying to answer.   |
|                         | I understood the question and there was only 1 answer that I could respond to making it easy to eliminate choices.  |
|                         | If it told you which layers were what   |
|                         | Tell me to compare the diagrams and see how they go together.   |
| Item 4                  | Picture of how far the stream went on those layers where the horse was found  |
|                         | Clearly state what we are looking for (the environment, or type of organism).   |
|                         | I understood the question and it was easy to understand the choices.  |
|                         | Marine life changes over time and B is different from E. If they had a table showing how things change, or if they change over time.  |
|                         | I need more explanation about the new marine life in the second diagram. How should I compare these?  |
| Item 5                  | A diagram of where the new fossils would be found in the rock layers with the horse   |
|                         | The pictures along the side and time periods of the eras helped me to answer. What might help other students is to indicate you need to read the diagram from bottom to top.                  |
|                         | The question made sense. The answers that went along with it, pertained to the question. The diagram included how many years ago and I wouldn't have known by just the era names.             |

| Question 7 Responses    |   |
|-------------------------|---|
|                         | Put the largest number on the top and move down from there  |
|                         | Show the landscape with each animal at the timeframe  |
| Item 6                  | Using the prior information and evidence from the previous questions. Using the diagrams and charts.  |
|                         | If the answers gave more information  |
|                         | Everyone thinks about the diagram, so I would have liked to write my own evidence explanation   |
|                         | What the layers looked like and a timeline showing that along the dates that they gave us   |
| Grade 8, Form A, Task 2 |   |
| Item 1                  | the student knows the concepts but the way it was worded and how you put the arrows in the boxes was unclear. Kind of between medium and hard. There had to be a force acting downward too because gravity is the constant. |
|                         | I did not understand how to answer it and what exactly the picture was showing.   |
|                         | Said to move up, but only let you use each arrow once. If you could use the same arrows more than once  |
|                         | Lifted upward, was it going left or right once it is lifted? I was confused about where it is going.  |
|                         | it could have given us amounts of force with the arrows does it lean?   |
|                         | more of an explanation of what is happening with the arrows. A picture of where it is now and where going   |
| Item 2                  | You can directly see the correlation between the data in the table and access to the given information about each part of the sculpture.  |
|                         | The chart they provided was easy to read and the chart made sense.  |
|                         | The greater mass the greater force needed. I would like a simulation to actually do what I am thinking.   |
|                         | I had to remember to scroll and know the math but without reading this passage you would be lost  |
|                         | You really had to think and go back to resources  |
|                         | It gave me what I needed  |
| Item 3                  | The relationship you were trying to find out was stated clearly   |
|                         | Instead of using the chart there I would like a line graph instead  |
|                         | Not much else needed - the data chart gave the information  |
| Item 4                  | Two answers could be correct. I can't find any way to eliminate one of them. A and B were easy to eliminate.  |
|                         | The first part was hard - I didn't understand quite what the question was asking, and I wasn't 100% sure on any of the choices. The second part was easy because I could plug in the numbers and check.                     |
|                         | If it had the two tables in a line graph instead of tables  |
|                         | Could have given us more information  |
|                         | Not much really   |
| Item 5                  | I knew how to do everything, but I doubted myself   |
|                         | It gives you the formula but doesn't give you the number for the acceleration   |
|                         | I didn't know what it was talking about it gave mass of 40 but the chart is only 10   |

| Question 7 Responses |  |
|----------------------|--|
|                      | We needed more information in the charts that include the numbers in the math problems, so like 40 kg wasn't on the chart.               |
|                      | More description - I was confused at first of what I was to try to do  |
| Item 6               | Previous knowledge and the information given through each previous question. The data tables were helpful.                               |
|                      | The question was easy to understand.   |
|                      | to make a good claim you would need to know accuracy of information. the specific calculation to check and see if that is what they want |
|                      | Explain how the base and butterfly connected to each other. If they connect well, maybe it could be done at once.                        |

**8. Ask the student: Was it easy or hard to use the given information to answer the question?**

1. Easy
2. Medium
3. Hard

**Table I.8. Question 8 Responses**

| Item #                         | Response # | N | %*   |
|--------------------------------|------------|---|------|
| <b>Grade 5, Form B, Task 1</b> |            |   |      |
| Item 1<br>(8 students)         | 1          | 7 | 87.5 |
|                                | 2          | – | –    |
|                                | 3          | 1 | 12.5 |
| Item 2<br>(9 students)         | 1          | 2 | 22.2 |
|                                | 2          | 2 | 22.2 |
|                                | 3          | 5 | 55.6 |
| Item 3<br>(9 students)         | 1          | 5 | 55.6 |
|                                | 2          | 3 | 33.3 |
|                                | 3          | 1 | 11.1 |
| Item 4<br>(9 students)         | 1          | 4 | 44.4 |
|                                | 2          | 4 | 44.4 |
|                                | 3          | 1 | 11.1 |
| Item 5<br>(9 students)         | 1          | 7 | 77.8 |
|                                | 2          | 2 | 22.2 |
|                                | 3          | – | –    |
| <b>Grade 5, Form B, Task 2</b> |            |   |      |
| Item 1<br>(8 students)         | 1          | 7 | 87.5 |
|                                | 2          | 1 | 12.5 |
|                                | 3          | – | –    |
| Item 2<br>(8 students)         | 1          | 3 | 37.5 |
|                                | 2          | 3 | 37.5 |
|                                | 3          | 2 | 25.5 |
| Item 3<br>(8 students)         | 1          | 7 | 87.5 |
|                                | 2          | 1 | 12.5 |
|                                | 3          | – | –    |
| Item 4<br>(8 students)         | 1          | 4 | 50.0 |
|                                | 2          | 1 | 12.5 |
|                                | 3          | 3 | 37.5 |
| Item 5<br>(7 students)         | 1          | 3 | 42.9 |
|                                | 2          | 2 | 28.6 |
|                                | 3          | 2 | 28.6 |
| Item 6<br>(7 students)         | 1          | – | –    |
|                                | 2          | 3 | 42.9 |
|                                | 3          | 4 | 57.1 |
| Item 7<br>(7 students)         | 1          | 4 | 57.1 |
|                                | 2          | 2 | 28.5 |
|                                | 3          | 1 | 14.2 |

| Item #                         | Response # | N | %*    |
|--------------------------------|------------|---|-------|
| Item 8<br>(6 students)         | 1          | 2 | 33.3  |
|                                | 2          | 3 | 50.0  |
|                                | 3          | 1 | 16.7  |
| <b>Grade 8, Form A, Task 1</b> |            |   |       |
| Item 1<br>(6 students)         | 1          | 4 | 66.7  |
|                                | 2          | 2 | 33.3  |
|                                | 3          | – | –     |
| Item 2<br>(6 students)         | 1          | 4 | 66.7  |
|                                | 2          | 2 | 33.3  |
|                                | 3          | – | –     |
| Item 3<br>(6 students)         | 1          | 4 | 66.7  |
|                                | 2          | 1 | 16.7  |
|                                | 3          | 1 | 16.7  |
| Item 4<br>(6 students)         | 1          | 4 | 66.7  |
|                                | 2          | 1 | 16.7  |
|                                | 3          | 1 | 16.7  |
| Item 5<br>(5 students)         | 1          | 5 | 100.0 |
|                                | 2          | – | –     |
|                                | 3          | – | –     |
| Item 6<br>(6 students)         | 1          | 3 | 50.0  |
|                                | 2          | 3 | 50.0  |
|                                | 3          | – | –     |
| <b>Grade 8, Form A, Task 2</b> |            |   |       |
| Item 1<br>(6 students)         | 1          | 1 | 16.7  |
|                                | 2          | 1 | 16.7  |
|                                | 3          | 4 | 66.7  |
| Item 2<br>(6 students)         | 1          | 3 | 50.0  |
|                                | 2          | 1 | 16.7  |
|                                | 3          | 2 | 33.3  |
| Item 3<br>(6 students)         | 1          | 4 | 66.7  |
|                                | 2          | 2 | 33.3  |
|                                | 3          | – | –     |
| Item 4<br>(6 students)         | 1          | 2 | 33.3  |
|                                | 2          | 4 | 66.7  |
|                                | 3          | – | –     |
| Item 5<br>(6 students)         | 1          | 2 | 33.3  |
|                                | 2          | 2 | 33.3  |
|                                | 3          | 2 | 33.3  |
| Item 6<br>(6 students)         | 1          | 2 | 33.3  |
|                                | 2          | 4 | 66.7  |
|                                | 3          | – | –     |

\*Percentages are out of the number of students who responded per item.

### 9. Ask the student: Why was it easy or hard?

Table I.9. Question 9 Responses

| Question 9 Responses    |  |
|-------------------------|--|
| Grade 5, Form B, Task 1 |  |
| Item 1                  | There was a lot of information to use. Lots of graphics and they help me a lot.  |
|                         | The information in the scenario and graph and diagram helped.  |
|                         | The words are easy and I know them.  |
|                         | The diagram didn't give you distance between places  |
|                         | easy   |
|                         | I have seen graphs like this before  |
|                         | Easy to read   |
|                         | The pictures help me and gives me details and helps you understand.<br>gave him the right amount of information to know where to put everything                  |
| Item 2                  | No information given about the nitrogen and the oxygen so that confuses me. Part a was confusing part b was ok part c was easy                                   |
|                         | He thought it started it out hard and then as he worked through it, he started to figure it out.   |
|                         | The colors were good for me because I am a visual learner and the words were easy.   |
|                         | In the box it tells you which gas, but not where the flower is   |
|                         | Model choices were confusing.  |
|                         | She couldn't look back to see if 5 meters was close or far away.   |
|                         | Pictures were confusing, and I don't know what they are showing.   |
|                         | I didn't understand the dot pictures<br>tells how far way, how much it is on the graph at the top  |
| Item 3                  | Because even though it wasn't very long it gave information along with what we got before  |
|                         | The question was clearly written.  |
|                         | The information they gave me didn't really help with this question I used what I already knew about gases.   |
|                         | I couldn't understand how close or far it needed to be.  |
|                         | Each choice helped me think it though and I matched it to what I had in my head.   |
|                         | I needed to reread the questions to see the closer or far away.  |
|                         | Not sure   |
|                         | Only one good choice.<br>they gave information about where they were smelling it from; if they didn't give that information it would have made it hard to answer |
| Item 4                  | Information is good - but it was hard to answer the question - like graphics and not just the question   |
|                         | There were 2 that were going outward but only one made more sense.   |
|                         | We talked about this a lot in class, so I know that gases go all over and don't clump up.  |
|                         | Because the farther the kids got away from the flower they could still smell it.   |
|                         | The picture helped me think.   |

| Question 9 Responses    |  |
|-------------------------|--|
|                         | The pictures are hard to follow and they all show the same things  |
|                         | I am not sure what an air sample is.   |
|                         | didn't give a lot of info, but the info was enough to answer the question. what types of info would have made it easier - to tell whether it was a small or large space; assuming small space because of the word sample |
| Item 5                  | Because there is a lot of information before the question  |
|                         | I had to go back and look at everything. text, graph, pictures, in my answers.   |
|                         | Tools (graphs, pics...) provided helped me explain the answer  |
|                         | The pictures above all explained how the particles moved.  |
|                         | I just knew it.  |
|                         | Map and the distance on the graph helped   |
|                         | I didn't really know what to write and then I remembered the pictures, so I wrote about what I saw.  |
|                         | I got to write down what I thought instead of looking at MC that has words I don't know or understand.   |
|                         | good enough amount of information; nothing else needed; gave so much info to use in my answer  |
| Grade 5, Form B, Task 2 |  |
| Item 1                  | They gave information in text and gave a picture   |
|                         | The answers were easy because of the diagram.  |
|                         | They show you the time and the decorations in the room, so you can see how the shadows were moving at different times.   |
|                         | When I first read the question it didn't make sense, but the picture helped me.  |
|                         | Comparing the two pictures made it easy.   |
|                         | It was the only choice that made sense and I have learned about it.  |
|                         | Pictures really explains it. I can use my hand to model the sun with the picture to see where it is at   |
|                         | you don't really need any more information to answer the question. The pictures give you all the information and telling you what time it is.  |
| Item 2                  | They gave a lot of information   |
|                         | Maybe labels on the all the diagrams.  |
|                         | The pictures show where you are with the red dots and how the earth is moving you around.  |
|                         | The pictures show the sun but I can't tell how it is moving from the picture.  |
|                         | After I reasoned out the pictures I knew the right answer quickly.   |
|                         | The choices were hard to pick from but the third one was all about the earth and the dots on the earth helped me figure it out.  |
|                         | The pictures are very similar.   |
|                         | gives some info, but you also have to know how the Earth rotates   |
| Item 3                  | I have a lot of information and I can use multiple things as well as my background knowledge   |
|                         | I had to look back. It did give me pictures of each item which helped.   |

| Question 9 Responses   |   |
|--|---|
|  | They showed pictures at the top, so you had information and with that the pictures of the choices I could tell which one would work best to block the sun in the picture.             |
|  | Because if you see the plant you know the sun goes through the narrow lamp and is too skinny to block the sun.  |
|  | I could go back and see the pictures of the room.   |
|  | None of the others would work as good to block the sun  |
|  | I COULD SCROLL BACK UP TO LOOK AT THE SHADOWS AND THAT TOLD ME THE ANSWER   |
|  | immediately know 3 answers that are wrong. You know that lamp that's on the table is already on the table... can immediately narrow to two choices, snake plant and ficus plant.      |
| Item 4   | We still have the information and background knowledge  |
|  | Looking back at the diagrams helped figure it out.  |
|  | They gave you a lot of information, but they didn't say what type of spring day it was.   |
|  | I understand the question and I know the earth rotates around the sun.  |
|  | A picture of winter and fall.   |
|  | The prediction part was hard, knew the shadows would change but wasn't sure how.  |
|  | I DON'T UNDERSTAND THE SHADOWS WITH THE SEASON AND HOW THOSE CONNECT. IS THIS ASKING IN BETWEEN WINTER AND SPRING   |
| if Q2 wasn't there, you would have to think logically, but since it's there it shows you that the shadow will be shorter |   |
| Item 5   | Enough is given - so you basically test the graph given to get the answer   |
|  | I didn't understand information in the question or the diagram. I was using my background knowledge instead.  |
|  | Words and picture were hard to understand. I didn't understand what the question was asking.  |
|  | The picture.  |
|  | hard to figure out which lines I needed to look at  |
|  | At first I wasn't sure but re-read it and figured it out.   |
|  | information given is a perfect amount; not a lot, not a little; shows you exactly what you need to know   |
| Item 6   | We have some information - but I don't feel like enough so more details in the question could help  |
|  | The diagrams don't tell you which earth is which season.  |
|  | The pictures are hard to understand,  |
|  | I just don't know what I am looking for in the question and it confused me  |
|  | I CAN THROW OUT SOME, BUT I GOT STUCK ON THE LEFT TOP AND BOTTOM THEY ARE A LOT ALIKE.  |
|  | information is given, but don't tell you what the information means; doesn't tell you how to use the information. talking about seasons; two choices showed seasons                   |
| Item 7   | We had enough graphics that provided more information than just words   |
|  | The pictures you could tell where the sun is and the state. I chose the points that meet where the destination is. They show you the horizon and mark which sun in summer and winter. |

| Question 9 Responses    |  |
|-------------------------|--|
|                         | It was hard because the winter is closer to the horizon line and I don't understand how the picture tells me colder or hotter.   |
|                         | The picture compared summer and winter   |
|                         | The pictures kind of look the same and they were hard to compare.  |
|                         | THE DIAGRAM SHOW THIS REALLY WELL  |
|                         | can't use only what is in the question; have to go back to previous things. have to think harder about the information in the question. referring back to earlier models and information was very helpful. |
| Item 8                  | We have the graphics at the top  |
|                         | Needs to say 9:00 am or pm.  |
|                         | Because I don't know how the shadows size is based on the sun  |
|                         | I remembered the pictures above, so I imaged summer and winter and then answered.  |
|                         | I can't explain it   |
|                         | A was easy to throw out and B and C didn't make sense. The questions before asked this similar question and i just kept to my answer   |
|                         | Yes, you can answer with the given information, but with more information you would have known how to do it.   |
| Grade 8, Form A, Task 1 |  |
| Item 1                  | Laid out nice, easy to see the diagrams for comparing.   |
|                         | the student drawing and the museum diagram don't look the same so it is confusing. The pictures did provide more information than just having the test, but they were hard to link together.               |
|                         | It's easier using your own words to explain things rather than choosing an answer that is given for me   |
|                         | Easy because they could be compared side by side and they see where everything is located  |
|                         | The passage and pictures explained everything really well  |
| Item 2                  | The pictures are helpful   |
|                         | The evidence throughout the other parts that help you answer the first part, helped build to this part. Could use my prior thinking.   |
|                         | I knew I needed to use the pictures to answer this but I didn't know how to link them together.  |
|                         | They had easier answers to choose from   |
|                         | Used diagram to find the location of the fossils.  |
| Item 3                  | The first few questions helped me answer it and we know that the earth was worn down so we can see the layers.   |
|                         | From the information I had to think to know why it was in a rock to mean it was a fossil   |
|                         | Everything was stated multiple times and the clear resources from prior questions.   |
|                         | I liked the chart and how they labeled it A - D.   |
|                         | The layers separate instead of going into each other and they labeled them, so it was easier to choose the answers   |
| Item 3                  | You don't know if one picture can go into another, not sure if they are directly comparable.   |
|                         | The picture showed enough to tell me where things are  |

| Question 9 Responses    |  |
|-------------------------|--|
|                         | The chart gave a picture - since I said shells older it would be further down. I thought D and E would be too far down to be seen from the surface |
| Item 4                  | The information was laid out nice. Easy to see side by side. Shows the relation between when the organisms lived and the rock layers.              |
|                         | Picture and diagrams were helpful. They helped make sense of the question.   |
|                         | It helps support each of the answers   |
|                         | It is hard to compare above and below ground so you have to think more about how to compare them.  |
|                         | You really had to look at it more and think outside of the box to answer this  |
|                         | Pictures showed more marine like compared to the diagram of the rock layers  |
| Item 5                  | Diagram laid out, so you can see it with lots of information in represented by pictures, times, words.   |
|                         | See above  |
|                         | It labels each column for their period and the organisms   |
|                         | The picture has more information and is easy to understand.  |
|                         | Chart gave you everything to answer  |
|                         | The pictures made it easy as it showed what they looked like as they changed   |
| Item 6                  | Well laid out.   |
|                         | The diagram helped me know how to respond and answer.  |
|                         | Two of them are really good - but there is always one that is better and answers it spot on. These answers aren't as clear as that                 |
|                         | part B there wasn't one answer that matched my answer for part A   |
|                         | First part was harder because you had to estimate. Second Part easy because it only had one answer that made sense                                 |
|                         | I had the chart from before that helped suggest what it should be  |
| Grade 8, Form A, Task 2 |  |
| Item 1                  | Mainly because of how it is worded. The way it was presented for me to give my answer was unclear. Weird.  |
|                         | I didn't understand the diagram, nor did I understand how to use what was given to answer the question.  |
|                         | It wants to move up and you can only use the arrows once - if need to go completely up then need more up arrows                                    |
|                         | We have talked about newtons or Kg, so I have forgotten how they work.   |
|                         | Because you had to think about it, there is less info, so you had to figure it out instead of looking back to find the answer.                     |
|                         | Because the picture shows A would go up and so forth   |
| Item 2                  | See above.   |
|                         | Easy to understand the question and use the information.   |
|                         | It is what we are learning right now   |
|                         | At first I didn't understand but I went back to the top and found the masses to help me  |
|                         | You really had to think and go back to resources   |

| Question 9 Responses |   |
|----------------------|---|
|                      | Because it was easy to find the relationship and then calculate the weight and determine the answer   |
| Item 3               | The table was organized and there was plenty of information to help you.  |
|                      | I understood the question and it was easy to rule two of the choices out because they didn't pertain to what the question was asking.   |
|                      | B explains the question much better than the other 2  |
|                      | I understood the information and read the passage and I consulted the chart as I read each answer.  |
|                      | The graph gave you the answer but the choices you had to read them carefully and reference the graph so make sure you were looking at things right  |
|                      | The Table have the information to make it easy to tell what happened  |
| Item 4               | Easy to uses the information but difficult to be conclusive on one answer.  |
|                      | The first part was hard - I didn't understand quite what the question was asking, and I wasn't 100% sure on any of the choices. The second part was easy because I could plug in the numbers and check. |
|                      | It is easy to understand but the question and answers make it confusing   |
|                      | A was hard to understand didn't know the relationship or how to relate the data. B I just looked at each piece and then I decided if the math would work for each equation then I picked one.           |
|                      | First part hard: lots of thinking and use all the parts. Second part easy: just do the math   |
|                      | The chart helped show what it was like  |
| Item 5               | everything was laid out previously and the formula was given.   |
|                      | The formula was provided, and the mass came from the bullet points.   |
|                      | It tells you what to use - but doesn't come right out to give you acceleration for the formula  |
|                      | I haven't done this for a while, so I didn't understand what it was asking me.  |
|                      | A lot of math and thinking, the charts didn't help as much.   |
|                      | You had to go the data charts and calculate the math for each of them   |
| Item 6               | See above   |
|                      | At first I didn't understand how to use the information, but then I took a second look and understood.  |
|                      | I'm not sure if my calculations earlier were right - but this question made sense   |
|                      | I went back to the data and found the max amount that the robot could pick up so I had to use the information in the passage above to help me.  |
|                      | Had to go back and try to remember everything. I was easy that I could explain it my way.   |
|                      | Used information at the top to help. It told me three pieces - so easier to do one at a time.   |

**10. Record any additional notes, comments, or observations made about the student's interaction with this question.**

Table I.10. Question 10 Responses

| Question 10 Responses          |  |
|--------------------------------|--|
| <b>Grade 5, Form B, Task 1</b> |  |
| Item 1                         | This is an EL student, but they understood all the words.  |
| Item 2                         | <i>No answers provided.</i>  |
| Item 3                         | <i>No answers provided.</i>  |
| Item 4                         | <i>No answers provided.</i>  |
| Item 5                         | Nothing to add, it was worded well and easy to understand  |
| <b>Grade 5, Form B, Task 2</b> |  |
| Item 1                         | <i>No answers provided.</i>  |
| Item 2                         | She is very confused.  |
| Item 3                         | <i>No answers provided.</i>  |
| Item 4                         | <i>No answers provided.</i>  |
| Item 5                         | <i>No answers provided.</i>  |
| Item 6                         | <i>No answers provided.</i>  |
| Item 7                         | She's agitated and really lost. I think she is just guessing at this point.  |
| <b>Grade 8, Form A, Task 1</b> |  |
| Item 1                         | organized.<br>The student indicated that she likes diagrams.   |
| Item 2                         | The different shape of the click boxes let me know that it's a different response format and that I could choose more than one.  |
| Item 3                         | I thought the student missed the word "shells" in the question, but he did not.  |
| Item 4                         | Part A was easy, but Part B made it more medium difficulty.  |
| Item 5                         | While the student didn't know that lower layers formed first and were older in a previous question, she could figure it out with this diagram.   |
| Item 6                         | <i>No answers provided.</i>  |
| <b>Grade 8, Form A, Task 2</b> |  |
| Item 1                         | When you put anything in box A, then it shifts the order of the boxes.<br>Note: my Cog Lab form did not look like the student question form. I would have had no idea what the question was wanting. |
| Item 2                         | <i>No answers provided.</i>  |
| Item 3                         | <i>No answers provided.</i>  |
| Item 4                         | Strong knowledge of math and use of the equation for evidence.   |
| Item 5                         | Lots of thinking through this  |
| Item 6                         | <i>No answers provided.</i>  |

**11. Ask the student: Tell me what you learned about \_\_\_\_\_. Record their response.**

**Table I.11. Question 11 Response Options**

| Task #                 | Response Options |
|------------------------|------------------|
| <b>Grade 5, Form B</b> |                  |
| Task 1                 | changing scents  |
| Task 2                 | changing shadows |
| <b>Grade 8, Form A</b> |                  |
| Task 1                 | fossils          |
| Task 2                 | forces           |

**Table I.12. Question 11 Responses**

| Question 11 Responses  |
|--|
| <b>Grade 5, Form B, Task 1</b>   |
| It depends on where you are and where the scent comes from. it also depends on if you can smell well or if you have a stuffy nose. I learned about scents from the graphics - they help me and I can see in my head.                 |
| If you get closer to something you can smell it better. Farther away it is not as strong. The scent doesn't always stay in one place it can pretty far away.   |
| By answering the questions, I saw that particles are spread out and that you can determine scent with your senses. I also learned that flowers are living things and are solids and the particles are packed closely.                |
| I learned the farther away the less you can smell  |
| I already knew the answer, but the questions were more difficult.  |
| It helped me to know that smells can be near or far or weak or stronger.   |
| I learned that you can smell far away and you don't have to see it. I also learned that flowers can be kind of hard.   |
| I learned that close to something has a stronger smell and far away is a weaker smell. Flowers can smell bad, rotting meat?  |
| I learned how the particles can move freely and they are not only around the one object. Since there is less at a farther distance, so the smell is weaker. There are more particles at a closer distance, so the smell is stronger. |
| <b>Grade 5, Form B, Task 2</b>   |
| I learned a lot from the graphics that the different seasons are similar from the motion of the earth and how the sun is shadowing   |
| I learned the depending on the earth position the shadows are larger or smaller depending on where the sun is facing.  |
| Shadows change at different times  |
| I learned in the summer, fall, winter and spring how the shadows are different and why they change.  |
| The season are different and things are moving   |
| Diagrams show rotation and the seasons and how the sun is out and not out. Why winter is winter and summer is summer because of how we face the sun.   |

### Question 11 Responses

I learned that since the Earth is tilted and goes around the sun, the seasons change. Depending on time of day, the shadows are shortest at 9:00ish in the morning (or maybe at night) and shortest at noon. The sun changes. In winter, the sun is lower because the Earth is tilted away from it. In summer, the sun is higher because the Earth is tilted toward it.

#### Grade 8, Form A, Task 1

Most of this I already knew. Otherwise, I would learn that older organisms are in the lower layers and vice versa for the younger fossils.

In Nebraska, the marine fossils are older than fossils like the horse. Nebraska has changed from ocean to swamp, to prairie.

The longer they are from the time period the deeper they are

After reading I learned that there are horse fossils and that the cliffs were used to dig fossils

NO

The rock layers can help judge how old they are and what the environment was like at that time

#### Grade 8, Form A, Task 2

I knew all of this but if I were to learn something from this, the force is dependent of the mass and acceleration of the object.

To figure out force you could use the acceleration rate times the mass.

The faster the force the more newtons it has - which is like its speed

I learned that forces help you pick up objects and it is the amount you are picking up. I learned Newtons

I learned how the forces work and the relationship with mass.

The more force the faster the acceleration of an object. The greater the mass the more force that needs to be applied.

**12. Ask the student: What are the science tests like in your class? Record their response.****Table I.13. Question 12 Responses**

| Question 12 Responses   |
|---|
| <b>Grade 5, Form B, Task 1</b>  |
| are medium and easy - we do things called anchor logs where we write things down and use worksheets that help us. The anchor logs helps to remind us on the topic. (mnemonics - rhyme to remember. If you pay attention - I like science so am more interested  |
| Work on your own. More distractions in class. Most questions are medium. Most MC but some writing. Pictures and evidence about topics.  |
| They aren't too hard, we get a study guide, I need to see stuff before a test. They are easy to understand.   |
| The tests are easy. The questions are what vocab and process based.   |
| We do science but not many tests, but we do experiments. I don't remember any of the questions.   |
| They are fun and ABC questions, some have pictures some do and some don't. I don't remember too much but I think they are all ABC type questions.   |
| Multiple choice and study guides they are easy  |
| Multiple choice, vocab, study sheets, matching, Easy to choose the answer   |
| Usually about whatever topic we are studying. Ex: solutions and mixtures. usually multiple-choice test, a few give picture of something - which one is heavier. at end, usually write 2 responses to two questions (usually a paragraph long). ex: Susie has 2 drinks with unknown solute and water is the solvent. What steps can she take to figure out the solute? |
| <b>Grade 5, Form B, Task 2</b>  |
| They are easier because we have different things to help us like anchor logs and mnemonics.   |
| Easy  |
| Multiple choice and vocab   |
| ABC questions   |
| We did have a unit on this before and talked about a similar thing. Gave a picture of a pole and a shadow and asked what time it would be. Multiple choice.   |
| <b>Grade 8, Form A, Task 1</b>  |
| Mostly multiple choice on things we already learned. We take a pretest before the unit, then given as a midpoint and final to see what we have learned.   |
| The questions are easy to answer because they are repetitive. About 4 or 5 of the same questions repeated.  |
| Most of them are easy because we do section after sections, so I know what to study specifically  |
| 70 questions, you have to study a lot, pay attention, not everything is on the study guide, vocab in context, multiple choice and 2 essays  |
| You have to study a lot, lots of info, test every week, all types of question   |
| We would learn about a specific topic and the vocabulary for sure, we would study a lot about what was on the test. We learn a lot about the vocabulary.  |
| <b>Grade 8, Form A, Task 2</b>  |
| Same as before. Questions are repetitive, but they don't build on each other.   |
| They are somewhat like this where they ask a math equation and you make a claim about it  |
| Vocabulary - we learn a lot about that and how they work  |

**13. Ask the student: How does that compare with what you just did? Record their response.**

**Table I.14. Question 13 Responses**

| Question 13 Responses  |
|--|
| <b>Grade 5, Form B, Task 1</b>   |
| The tests are easier because I know what is coming and I can study for it - but I didn't know what was coming in this test, so I couldn't study. Our tests don't really give graphics to use because we study and don't need the pictures. This was also harder because the question were more complex than what ours are.   |
| Same kind of questions and variety of difficulty. This one gave more evidence about the topic. Usually ours only have 1-2 Q about the topic. I like this one better.   |
| The questions are different. The questions are on vocabulary. This test was new to me and I didn't know what was coming for me. This test was a little bit harder than my class tests because you had to think more, and you didn't know what the questions were, and you didn't study.  |
| Different. This test is harder, and the questions aren't as simple.  |
| I think this test was harder. I knew the answers, but the diagrams and choices make it harder.   |
| This test was sometimes hard and sometimes easy. The diagrams made it easier but sometimes the words are harder.   |
| This is kind of harder because I couldn't study for it. The questions are worded differently and hard to understand.   |
| I study for class tests and you can't study for this. This test is digital not on paper. The question weren't bad, but some were hard to know what they were asking.   |
| This is also a multiple-choice test and has a written response. Different - don't usually get as much information. This gives you the information on the question. In class given information a week or two before, then given a test on that information. On this test it gives you the information as you do it, not before (meaning not during an earlier class). |
| <b>Grade 5, Form B, Task 2</b>   |
| Our test are easier because we know what is on it - what we are tested on so we can study better   |
| This relates most to what we have done in science, I have learned about the sun and it's rotation. This one was easier than the last.  |
| This is more difficult the pictures didn't make sense.   |
| easier   |
| These are more difficult questions and you cant study for them.  |
| These questions were showing a lot of pictures and we don't always get those in class. These questions relate to real stuff. The question types are similar.   |
| This test gave a lot of models of things whereas our previous tests gave one or two. The models were very important, could answer most of the questions without the models.  |
| <b>Grade 8, Form A, Task 1</b>   |
| Most of these questions were about one subject but they worked together to support the claim that was made at the very beginning - Which is older, the shell layer or the horse layer?   |
| These questions all linked together but not repetitive. You used your knowledge from the first questions to help you with the next questions.  |
| They are similar because there was just one topic. The questions are also similar because we are doing more computer based stuff (Amplify)   |

**Question 13 Responses**

Kind of similar we have started getting more diagrams in our class but the questions are different more content based make simple conclusions.

Teachers questions are way harder, this has more diagrams and pictures.

This is more charts and comparing and contrasting things. We do comparing sometimes.

**Grade 8, Form A, Task 2**

See before.

somewhat similar

We haven't talked about this. We don't do math like this on tests. We did a lot of math last year in 7th grade though

This seems more similar to what we do in our class, you need to study this so you can do good. I haven't learned this yet. The math and typing are similar.

This had a lot of charts and math that you had to go through

Ask the student to read aloud and answer each question in the **Student Feedback** section.

**Feedback Question #1**

How interesting was the task you just completed?

1 star is not interesting and 5 stars is very interesting.

☆☆☆☆☆

**Table I.15. Question 14 Ratings**

| Task #                 | Rating | N |
|------------------------|--------|---|
| <b>Grade 5, Form B</b> |        |   |
| Task 1                 | 1      | – |
|                        | 2      | – |
|                        | 3      | 4 |
|                        | 4      | 3 |
|                        | 5      | 1 |
| Task 2                 | 1      | – |
|                        | 2      | – |
|                        | 3      | 3 |
|                        | 4      | 2 |
|                        | 5      | – |
| <b>Grade 8, Form A</b> |        |   |
| Task 1                 | 1      | – |
|                        | 2      | – |
|                        | 3      | 2 |
|                        | 4      | 3 |
|                        | 5      | 1 |
| Task 2                 | 1      | 1 |
|                        | 2      | 1 |
|                        | 3      | – |
|                        | 4      | 1 |
|                        | 5      | 2 |
| <b>Overall</b>         |        |   |
| Overall                | 1      | 1 |
|                        | 2      | 1 |
|                        | 3      | 9 |
|                        | 4      | 9 |
|                        | 5      | 4 |

**14. Ask the student: Why did you rate your interest in the task with 1, 2, 3, 4, or 5 stars?****Table I.16. Question 14 Responses**

| Question 14 Responses   |
|---|
| <b>Grade 5, Form B, Task 1</b>  |
| Most of the questions were ok with my background knowledge. I didn't rate 5 because scent isn't what I am focused on. I didn't rank less than 4 because I did learn through it  |
| Overall rating a 3. Not really interested in flowers. The question that had 3 parts and the writing question were the hardest.  |
| 4 stars, I love science and it is my favorite so it just interesting to me and I learned a lot  |
| 3 Sort of interesting it explains how scent goes in and out.  |
| Very interesting, but I was curious about how many questions, so I clicked through to see and it ended the test. oops so sorry. The student and I had a good laugh...he thinks the department of education will know about this!!   |
| I put 5 stars because of the plant and that you could smell it far away. It was interesting because I have never heard of that plant.   |
| 3 because it was interesting but I kind of knew some of it.   |
| 4: because I really didn't understand but I liked the pictures  |
| Rated it 3 because I learned some new things that I never knew before. It was an interesting topic, and partly because they gave some good examples of how the particles move (referring to questions 2A and 4).  |
| <b>Grade 5, Form B, Task 2</b>  |
| rated 3 stars because it wasn't boring but I know a lot of stuff about that so it wasn't as interesting   |
| 4 because in 4th grade he did something with Earth's rotation and that was really interesting. This is a good topic.  |
| Interesting because I love learning about earth an space.   |
| 3 because the pictures interesting to understand  |
| Cause it was interesting to know about sun and the shadows and what makes it happen   |
| 4: I didn't know a lot about this and I learned a lot   |
| I liked the pictures and the seasons.   |
| 3 - it is a topic I like; taught me some things I didn't know about how the sun is lower in winter, higher in summer; explains why the seasons are the way they are   |
| <b>Grade 8, Form A, Task 1</b>  |
| 5 stars for a test. Even though the questions all related to the original question, they all built with more and more information. They seemed like new topics that you were working through. As you progressed through the test the information you are learning helps to back up what you have already learned through the previous parts of the test. Most tests are based on memory. This one you learn as you go. If I was designing a test I would put information like this in it. If you get a job filling out tax reports, your boss isn't going to say, you can't use any of this given information. You won't be successful. |
| 4 stars. I learned something. As I kept going through the questions I understood it more. I thought it was more interesting then.   |
| 3 - it was interesting - but not as interesting as some other science topics  |
| 4, it is stuff we haven't done this year  |
| 3: I like dinosaurs, so it wasn't so boring   |

**Question 14 Responses**

4 - I thought it was pretty interesting to see how old they were and how it compares to other organisms

**Grade 8, Form A, Task 2**

5 stars. The topic was interesting, and it builds through the test.

The first one was more interesting. This one I wouldn't go out of my way to learn more about it because it was harder to understand. Fossils more interesting than forces.

2 stars because it is interesting - but I don't like force and motion much

4, because I learned a lot

1: because it was hard and not something I like

5 star - I thought it was very interesting to learn about the relationships between forces

**Feedback Question #2**

How difficult was the task as a whole?

1 star is not difficult and 5 stars is very difficult.

☆☆☆☆☆

**Table I.17. Question 15 Ratings**

| Task #                 | Rating | N |
|------------------------|--------|---|
| <b>Grade 5, Form B</b> |        |   |
| Task 1                 | 1      | 1 |
|                        | 2      | 3 |
|                        | 3      | 1 |
|                        | 4      | 1 |
|                        | 5      | – |
| Task 2                 | 1      | – |
|                        | 2      | 2 |
|                        | 3      | – |
|                        | 4      | 4 |
|                        | 5      | – |
| <b>Grade 8, Form A</b> |        |   |
| Task 1                 | 1      | 2 |
|                        | 2      | 1 |
|                        | 3      | 3 |
|                        | 4      | – |
|                        | 5      | – |
| Task 2                 | 1      | 1 |
|                        | 2      | – |
|                        | 3      | 2 |
|                        | 4      | 3 |
|                        | 5      | – |
| <b>Overall</b>         |        |   |
| Overall                | 1      | 4 |
|                        | 2      | 6 |
|                        | 3      | 6 |
|                        | 4      | 8 |
|                        | 5      | – |

**15. Ask the student: Why did you rate the difficulty of the task with 1, 2, 3, 4, or 5 stars?****Table I.18. Question 15 Responses**

| Question 15 Responses  |
|--|
| <b>Grade 5, Form B, Task 1</b>   |
| Some of the questions were really easy and some were hard because of the information that was given  |
| 2 because the questions weren't super hard or worded super goofy. It could've been harder.   |
| 4: pretty difficult the questions were hard to understand. The pictures were easy.   |
| I gave it 3 stars because some were easy some were hard.   |
| 2 it wasn't that difficult but i didn't get some of the questions  |
| 1: Pretty easy   |
| Rated it 2 because they gave us a lot of information, making the questions easier than if they gave us a little information.   |
| <b>Grade 5, Form B, Task 2</b>   |
| 2 stars overall - some were hard and some were easy - but I have background knowledge and the graphics helped  |
| 4. I had to go and look back and use background knowledge.   |
| It wasn't difficult to understand the wording and pictures.  |
| 4 difficult because of pictures and wording  |
| I gave it 2 stars some questions you have to think through their meaning and the answer.   |
| 4: hard because I didn't understand the questions and the pictures   |
| Not that difficult I could figure it out most times.   |
| 4 - gave enough information, but more information could have been useful to make it easier. Some questions were not specific enough (Q8 -- need a.m. or p.m.) (Q6 just expects you to figure it out; doesn't tell you how to figure it out; maybe if models explained what the models stood for (which season), it would be easier to answer)  |
| <b>Grade 8, Form A, Task 1</b>   |
| 1 star. Not very difficult. I like that the information was listed and then keeps adding information to help you use to back up your argument (claim). As you keep going on you begin to understand it more and more. If you only posed with the last question (like most test equivalents) you might not have a good understanding of the topic or what the question is saying. But because you answered the previous questions you could understand it better. |
| 3 The questions started out difficult because I didn't know a lot about fossils but as I went along the diagrams got better and I started to understand more about fossils.  |
| 2 - there are only a couple of questions that didn't make sense in it  |
| 3, in the middle so parts were harder than others.   |
| 1: it gave you everything and you didn't really have to think  |
| 3 - because you had to know how to compare, but the charts and diagrams made that easier   |
| <b>Grade 8, Form A, Task 2</b>   |
| 1. previous knowledge and given information  |
| 4. I was confused at times and didn't understand what I was doing most of the time.  |
| 3 stars because you have to be correct with calculations in order to have a good claim   |
| 4, it was pretty hard because I forgot most of this stuff  |
| 4: Some parts were easier than others, but lots were hard  |
| 3 star - because you had to do a lot of math - charts helped - but you had to know how to use them   |

**Feedback Question #3**

Please rate the difficulty **to understand how to respond to each question.**

1 star is not difficult and 5 stars is very difficult.

**Table I.19. Question 16 Ratings**

| Task #                 | Rating | N |
|------------------------|--------|---|
| <b>Grade 5, Form B</b> |        |   |
| Task 1                 | 1      |   |
|                        | 2      |   |
|                        | 3      |   |
|                        | 4      |   |
|                        | 5      |   |
| Task 2                 | 1      |   |
|                        | 2      |   |
|                        | 3      |   |
|                        | 4      |   |
|                        | 5      |   |
| <b>Grade 8, Form A</b> |        |   |
| Task 1                 | 1      |   |
|                        | 2      |   |
|                        | 3      |   |
|                        | 4      |   |
|                        | 5      |   |
| Task 2                 | 1      |   |
|                        | 2      |   |
|                        | 3      |   |
|                        | 4      |   |
|                        | 5      |   |
| <b>Overall</b>         |        |   |
| Overall                | 1      |   |
|                        | 2      |   |
|                        | 3      |   |
|                        | 4      |   |
|                        | 5      |   |

**16. Ask the student: Why did you rate the difficulty of the questions with 1, 2, 3, 4, or 5 stars?**

**Table I.20. Question 16 Responses**

| Question 16 Responses   |
|---|
| <b>Grade 5, Form B, Task 1</b>  |
| The questions were different. I had more background knowledge of some than others. Also, I rated some hard because I didn't have much information given and little background knowledge.  |
| Question 1: Easy-it tells you where everything is located. Question 2: Sort of difficult because I didn't know where I was or the flower was in the picture. Question 3: Easy-Simple question- no picture to look at. Question 4: I don't know what an air sample is. Question 5: I didn't know how to explain the strengths of the scent.  |
| I gave it one star because I understood the directions  |
| 1: 2 picture helped. 2: 4 I didn't understand the pictures and I didn't know what type of particles they were. 3: 2 reading the options there was one the stood out to me. 4: 3 it was hard but easy too. 5: 3 it was hard to make up my own explanation  |
| I like question 5   |
| so, Q1 was asking how strong the scent was... I just realized i did it wrong. I'm going to fix that. I would rate Q1 a 1. Q2 is rated 3 - the info was enough but wasn't a lot. You have to know a little bit yourself outside of the what is in the passage. Q3 is rated 2 - mostly because it's talking about why they smell it from far away, so you have to know how particles move. Q4 is rated 4 - mostly because there isn't a lot of information, but there is enough. Q5 is rated 3 - asking you to summarize everything you just did.   |
| <b>Grade 5, Form B, Task 2</b>  |
| 1 - 1 star because of the graphics. 2 - 3 stars because of the way the graphics were pictured, and you could only rely on graphics. 3 - 1 star because I know how big the things are - or narrow or wide. 4 - 1 star because you could tell by the pictures from earlier. 5 - 1 star because the graphic was easy to use my finger and measure the increase or decrease. 6 - 3 stars because the arrows were confusing - no sun in two of them to help. 7 - 1 star because I could rely on the graphics from before because the shadows were the same in both. 8 - 1 star because we knew they were the same from the previous questions which took out b, c and d. |
| Question 1: easy because it gave you a lot of information and pictures. Question 2: it took more thinking and was more challenging. Question 3: it was easy because the light and size of the object work together to block out. Question 4: The time of spring wasn't given. Question 5: Easy because the picture was easy to understand, and I know summer is hotter because the sun is closer. Question 6: More difficult because the pictures were harder to understand. Question 7: The explanation was good, but you need knowledge to answer. Question 8: Was easy because there is enough detail before this question to know what they were talking about. |
| She is very distracted at this point. Question 1: 3 the pictures were hard to understand. Question 2: 1 its easy to understand. Question 3: 3 can't decide on shorter or longer. Question 4: 2 seasons and shadow length are confusing. Question 5: 4 didn't understand the pictures or wording. Question 6: 4 didn't understand the pictures or wording. Question 7: 4 didn't understand the pictures or wording. Question 8: 3 a little bit easier but hard because of the pictures.  |
| I gave it one star .I knew what to do.  |
| Question 3 was easier to understand. Question 5 was harder, I didn't understand it.   |
| #2. 3 because the models were very similar. #3. 1 easy cause you could go back and look at the pictures. #4. 2 because there was already another question like it. #6. 4 diagrams were similar.   |
| Q1 = 2 - have to understand how the sun works and how shadows work. Q2 = 3 - have to understand what the models mean instead of just giving 4 choices (proctor interpretation - must be able to interpret   |

### Question 16 Responses

each model). Q3 = 2 - gives a couple options that you can immediately tell that are wrong. Q4 = 3 - have to go back for previous information (Q2) because that's the only place where you can get the information. Q5 = 4 - have to understand what the model stands for exactly; if you just look at it you might be confused about what the yellow dot is (time of day? the sun?). Q6 = 5 - have to understand what each model means so you can get your right answer. Q7 = 3 - asking you for 9 a.m., early in the morning;

#### Grade 8, Form A, Task 1

1 information keeps getting added. It gets easier and easier to explain the answers. Starts off easy and progressively gets harder but you also get more information.

Question 1 I didn't understand at the beginning. It made sense later. Question 2 I didn't know how to rule out some of the choices. The checkboxes looked different, so I knew it could be more than one answer. Question 3 was easier because I could infer or assume that the stream was more likely toward the bottom. Question 4 gets a 1 because it was very clear, and I understood what it was saying. Question 5 gets a 1 the pictures and years were most helpful and could make some links from prior questions. Question 6 was easy because I knew where the shells were found and where the horse was found.

1 - 2 stars - not easy, but also not really hard. 2 - 3 stars because some of the claims didn't make sense so it was hard to understand. 3 - 2 stars because you can't tell which layer are which - but question is easy. 4 - 2 stars it is still understandable, and questions and answers match up. 5 - 1 star because it is easy to understand, and diagram helps with the question. 6 - 3 stars because it is understandable but the claims in part b don't answer with part a.

1: middle because I already knew about the rock layers. 2: Harder to pick one answer to match part A. 3: Had to look back and compare. 4: Harder because it was more difficult to understand what the question was asking. 5: Easy to read chart. 6: A and B parts were hard to match, part B wasn't what I was thinking.

Last Question was hardest because of the years you had to estimate.

1 - 3 star because the graph and text explained how they are found and where. 2 - 4 star because there were so many options to think about and process why you are choosing it. 3 - 2 star because I had already processed that the shells were older, so it was easier to predict. 4 - 4 star because it was harder. A was easier to determine that marine different from the horses, but part B I had to think more about how environment changed over time. 5 - 1 star - easy because of the pictures that they gave me. 6 - 4 star because part A was more difficult - you had to eliminate the younger. part b was knowing where they were on the chart

#### Grade 8, Form A, Task 2

1 the majority of the task was easy, I just doubted myself.

This was more difficult. I didn't understand the given information and it was difficult. Not as easy to figure out as the first task.

1 - 2 stars because not very difficult. 2 - 2 stars because not difficult but I had to think a little. 3 - 3 stars because it made sense to a point but was confusing. 4 - 2 stars like 1 & 2. 5 - 4 stars because it tells you that you need the formula, but I didn't know for certain what numbers to use. 6 - 3 stars because it is easy to write about it, but hard to make a good claim when I was uncertain about the information

Can't go back to see 1-4. 5: was hard, didn't remember this math. 6: was easier because it was my own answer and what I thought

Can't scroll back to 1-4 to rank them. 5: was hardest-do all the math and do you estimates right. 6: easier.

Didn't have ability to go back and look at each of the items. 1 (if arrows put in) - 3 star - confusing to know what to do. 2 - 2 - graph shows relationship. 3 - 2 - graph made easy. 4 - 4 put graphs together made you think more and compare. 5 - 4 - understand what to do and math calculations - which formula to use. 6 - 2 - knew I needed to do one by one so thought easy

## Appendix J: Cognitive Lab Coding Results

To analyze the open-ended responses for Questions 7, 9, 12, 13, 14, 15, and 16 of the cognitive lab, each response was reviewed and given a code or multiple codes based on similarities among the responses. When referring to qualitative data, a code represents a word or short phrase that summarizes the open-ended responses. The coding began with “1” and continued for as many codes as needed to make sure the responses were appropriately considered when analyzing the data. Once the responses were coded, the frequency of each code was tallied and placed in a table, thereby allowing the qualitative data to be measured and analyzed similarly to quantitative data.

The following tables present the qualitative coding results after analyzing the comments. Some responses were associated with more than one concept and, therefore, were assigned more than one code. Therefore, the percentages are out of the total number of responses indicated in each table and do not add up to 100.0%.

### Question 7. Ask the student: What would help you answer the question?

| Code  | N          | %        |
|---|------------|----------|
| 1. Graphs/diagrams/pictures/charts were helpful                                 | 16         | 13.9     |
| 2. Nothing/item had enough info/item was easy to understand/I had the knowledge | 28         | 24.3     |
| 3. Content-specific knowledge needed  | 6          | 5.2      |
| 4. Confusing/clearer directions, descriptions, graphics needed/more info needed | 33         | 28.7     |
| 5. Referring back to previous items and info                                    | 9          | 7.8      |
| 6. Suggestions for improvements to diagrams/charts/graphics                     | 15         | 13.0     |
| 7. Other  | 19         | 16.5     |
| <b>Total</b>  | <b>115</b> | <b>–</b> |

### Question 9. Ask the student: Why was it easy or hard?

| Code   | N*         | %        |
|--|------------|----------|
| 1. Graphics were helpful   | 45         | 25.7     |
| 2. Item was easy to understand/item gave sufficient info/I had the knowledge | 39         | 22.3     |
| 3. Confusing/not enough info provided/unclear graphics                       | 37         | 21.1     |
| 4. Referring to previous information and graphics                            | 18         | 10.3     |
| 5. Process of elimination  | 13         | 7.4      |
| 6. Difficult item/required critical thinking/I didn't have enough knowledge  | 19         | 10.9     |
| 7. Other   | 15         | 8.6      |
| <b>Total</b>   | <b>175</b> | <b>–</b> |

**Question 12. Ask the student: What are the science tests like in your class? Record their answer.**

| Code   | N         | %        |
|--|-----------|----------|
| 1. Tests are easy  | 9         | 39.1     |
| 2. Mostly multiple-choice                                  | 8         | 34.8     |
| 3. Focused on vocabulary                                   | 5         | 21.7     |
| 4. Use study guides/worksheets                             | 5         | 21.7     |
| 5. A lot of studying required                              | 3         | 13.0     |
| 6. Tested on what they're studying or have already learned | 3         | 13.0     |
| 7. Other   | 7         | 30.4     |
| <b>Total</b>   | <b>23</b> | <b>–</b> |

**Question 13. Ask the student: How does that compare with what you just did? Record their response.**

| Code   | N         | %        |
|--|-----------|----------|
| 1. Classroom tests are easier                                | 7         | 25.9     |
| 2. Know what to expect on classroom tests, but not the pilot | 7         | 25.9     |
| 3. Pilot had harder/more complex items                       | 6         | 22.2     |
| 4. Pilot had more graphics                                   | 6         | 22.2     |
| 5. Pilot provided more info                                  | 3         | 11.1     |
| 6. The tests have a similar format                           | 7         | 25.9     |
| 7. Other   | 9         | 33.3     |
| <b>Total</b>   | <b>27</b> | <b>–</b> |

**Question 14. Ask the student: Why did you rate your interest in the task with 1, 2, 3, 4, or 5 stars?**

| Code  | N         | %        |
|---|-----------|----------|
| 1. The task was interesting                   | 18        | 62.1     |
| 2. I learned something                        | 7         | 24.1     |
| 3. Items build on each other                  | 2         | 6.9      |
| 4. I did not like the topic/wasn't interested | 4         | 13.8     |
| 5. I liked the pictures                       | 2         | 6.9      |
| 6. Other                                      | 5         | 17.2     |
| <b>Total</b>                                  | <b>29</b> | <b>–</b> |

Note: There was often the same explanation for different ratings.

**Question 15. Ask the student: Why did you rate the difficulty of the task with 1, 2, 3, 4, or 5 stars?**

| Code  | N         | %        |
|---|-----------|----------|
| 1. Some items were easy, some were hard                                 | 5         | 18.5     |
| 2. Task was confusing/items were hard to understand (wording, pictures) | 4         | 14.8     |
| 3. Task became easier as it progressed                                  | 2         | 7.4      |
| 4. Task wasn't difficult to understand/provided enough info             | 9         | 33.3     |
| 5. Graphics helped/made it easier                                       | 4         | 14.8     |
| 6. Other  | 8         | 29.6     |
| <b>Total</b>  | <b>27</b> | <b>–</b> |

Note: There was often the same explanation for different ratings.

**Question 16. Ask the student: Why did you rate the difficulty of the questions with 1, 2, 3, 4, or 5 stars?**

| <b>Code</b>  | <b>N</b>  | <b>%</b> |
|--|-----------|----------|
| 1. Needed outside/background knowledge/didn't know how to respond                  | 7         | 28.0     |
| 2. Item/graphic was easy to understand/provided sufficient info                    | 15        | 60.0     |
| 3. Item/graphic was complex/difficult to understand/didn't provide sufficient info | 16        | 64.0     |
| 4. Parts of the item were easy to understand, while others were confusing          | 5         | 20.0     |
| 5. Referred to previous info/graphics/items to answer the item                     | 4         | 16.0     |
| 6. Items became easier as the task progressed                                      | 3         | 12.0     |
| 7. Other   | 13        | 52.0     |
| <b>Total</b>   | <b>25</b> | <b>–</b> |

Note: There was often the same explanation for different ratings. Additionally, many responses had multiple different answers for the various items associated with a task, so many responses had multiple codes assigned to them.

**Appendix K: Prompt-Level DIF Results**

| Prompt Code     | ELL     |     |     | Gender  |       |     | SES     |       |     | Ethnicity |     |     |
|-----------------|---------|-----|-----|---------|-------|-----|---------|-------|-----|-----------|-----|-----|
|                 | N-Count |     | DIF | N-Count |       | DIF | N-Count |       | DIF | N-Count   |     | DIF |
|                 | R       | F   |     | R       | F     |     | R       | F     |     | R         | F   |     |
| g5a_T1_Q1       | 2,175   | 217 | A   | 1,283   | 1,227 | A   | 1,467   | 1,043 | A   | 1,783     | 439 | A   |
| g5a_T1_Q2       | 2,159   | 217 | C-  | 1,275   | 1,219 | B-  | 1,460   | 1,034 | A   | 1,773     | 438 | B-  |
| g5a_T1_Q3       | 2,165   | 216 | A   | 1,276   | 1,223 | A   | 1,463   | 1,036 | A   | 1,776     | 438 | A   |
| g5a_T1_Q4_Group | 2,101   | 210 | A   | 1,231   | 1,193 | B-  | 1,436   | 988   | C-  | 1,733     | 419 | C-  |
| g5a_T2_Q1_PartA | 2,096   | 208 | B+  | 1,240   | 1,181 | A   | 1,422   | 999   | A   | 1,722     | 428 | A   |
| g5a_T2_Q1_PartB | 2,090   | 207 | C+  | 1,229   | 1,180 | A   | 1,417   | 992   | A   | 1,717     | 423 | B+  |
| g5a_T2_Q2_PartA | 2,088   | 208 | B-  | 1,230   | 1,182 | A   | 1,420   | 992   | A   | 1,718     | 423 | C-  |
| g5a_T2_Q2_PartB | 2,070   | 210 | A   | 1,224   | 1,170 | A   | 1,410   | 984   | A   | 1,704     | 422 | A   |
| g5a_T2_Q3       | 1,022   | 90  | A   | 571     | 578   | A   | 729     | 420   | A   | 870       | 168 | A   |
| g5a_T2_Q4       | 2,088   | 210 | A   | 1,232   | 1,183 | A   | 1,418   | 997   | A   | 1,717     | 428 | B+  |
| g5a_T2_Q5       | 2,082   | 208 | B+  | 1,227   | 1,180 | A   | 1,412   | 995   | A   | 1,711     | 428 | A   |
| g5b_T1_Q1_Group | 1,963   | 199 | A   | 1,198   | 1,102 | A   | 1,333   | 967   | B-  | 1,578     | 449 | A   |
| g5b_T1_Q2_PartA | 1,952   | 197 | A   | 1,187   | 1,097 | A   | 1,327   | 957   | A   | 1,568     | 445 | A   |
| g5b_T1_Q2_PartB | 1,923   | 192 | A   | 1,166   | 1,083 | A   | 1,307   | 942   | A   | 1,549     | 436 | A   |
| g5b_T1_Q2_PartC | 1,947   | 198 | A   | 1,187   | 1,095 | A   | 1,322   | 960   | A   | 1,563     | 447 | A   |
| g5b_T1_Q3       | 1,935   | 197 | A   | 1,182   | 1,087 | A   | 1,321   | 948   | A   | 1,554     | 444 | A   |
| g5b_T1_Q4       | 1,952   | 195 | A   | 1,184   | 1,098 | A   | 1,328   | 954   | A   | 1,565     | 445 | A   |
| g5b_T2_Q1       | 1,967   | 198 | A   | 1,195   | 1,109 | A   | 1,332   | 972   | A   | 1,576     | 451 | B-  |
| g5b_T2_Q2       | 1,966   | 197 | A   | 1,191   | 1,109 | A   | 1,333   | 967   | A   | 1,575     | 448 | A   |
| g5b_T2_Q3       | 1,964   | 199 | A   | 1,191   | 1,108 | A   | 1,332   | 967   | A   | 1,574     | 450 | A   |
| g5b_T2_Q4       | 1,962   | 199 | A   | 1,190   | 1,108 | A   | 1,330   | 968   | A   | 1,574     | 450 | A   |
| g5b_T2_Q5       | 1,955   | 197 | A   | 1,188   | 1,101 | A   | 1,326   | 963   | A   | 1,569     | 445 | A   |
| g5b_T2_Q6       | 1,938   | 192 | A   | 1,174   | 1,089 | A   | 1,312   | 951   | A   | 1,551     | 439 | A   |
| g5b_T2_Q7       | 1,930   | 193 | A   | 1,168   | 1,088 | A   | 1,308   | 948   | A   | 1,551     | 439 | A   |
| g5b_T2_Q8       | 1,963   | 198 | A   | 1,193   | 1,105 | A   | 1,329   | 969   | A   | 1,574     | 449 | A   |
| g8a_T1_Q1_PartA | 2,461   | 231 | A   | 1,400   | 1,349 | A   | 1,779   | 970   | A   | 2,014     | 474 | A   |
| g8a_T1_Q1_PartB | 2,455   | 230 | A   | 1,393   | 1,349 | A   | 1,773   | 969   | A   | 2,010     | 472 | A   |
| g8a_T1_Q2       | 2,460   | 231 | A   | 1,398   | 1,350 | A   | 1,776   | 972   | A   | 2,013     | 474 | A   |
| g8a_T1_Q3       | 2,461   | 231 | A   | 1,399   | 1,350 | A   | 1,778   | 971   | A   | 2,015     | 474 | A   |
| g8a_T1_Q4_PartA | 2,458   | 231 | A   | 1,397   | 1,348 | A   | 1,777   | 968   | A   | 2,012     | 473 | A   |
| g8a_T1_Q4_PartB | 2,457   | 230 | A   | 1,397   | 1,345 | A   | 1,775   | 967   | A   | 2,010     | 473 | A   |
| g8a_T1_Q5       | 2,461   | 229 | A   | 1,396   | 1,351 | A   | 1,778   | 969   | A   | 2,013     | 474 | A   |
| g8a_T1_Q6_PartA | 2,453   | 231 | A   | 1,395   | 1,346 | A   | 1,773   | 968   | A   | 2,011     | 472 | A   |
| g8a_T1_Q6_PartB | 2,454   | 231 | A   | 1,393   | 1,349 | A   | 1,773   | 969   | A   | 2,010     | 474 | A   |
| g8a_T2_Q1_Group | 2,418   | 223 | B-  | 1,370   | 1,327 | A   | 1,744   | 953   | A   | 1,979     | 465 | A   |
| g8a_T2_Q2_PartA | 2,450   | 226 | A   | 1,392   | 1,341 | A   | 1,766   | 967   | A   | 2,004     | 469 | A   |
| g8a_T2_Q2_PartB | 2,270   | 200 | A   | 1,268   | 1,253 | A   | 1,657   | 864   | A   | 1,878     | 419 | A   |
| g8a_T2_Q3       | 2,444   | 226 | A   | 1,388   | 1,337 | A   | 1,763   | 962   | A   | 1,999     | 465 | A   |
| g8a_T2_Q4_PartA | 2,450   | 225 | A   | 1,392   | 1,340 | A   | 1,766   | 966   | A   | 2,003     | 468 | A   |
| g8a_T2_Q4_PartB | 2,449   | 226 | A   | 1,392   | 1,340 | A   | 1,764   | 968   | A   | 2,003     | 469 | A   |

Appendix K: Prompt-Level DIF Results

| Prompt Code     | ELL     |     |     | Gender  |       |     | SES     |     |     | Ethnicity |     |     |
|-----------------|---------|-----|-----|---------|-------|-----|---------|-----|-----|-----------|-----|-----|
|                 | N-Count |     | DIF | N-Count |       | DIF | N-Count |     | DIF | N-Count   |     | DIF |
|                 | R       | F   |     | R       | F     |     | R       | F   |     | R         | F   |     |
| g8a_T2_Q5_PartA | 2,268   | 200 | A   | 1,272   | 1,244 | A   | 1,654   | 862 | A   | 1,871     | 421 | A   |
| g8a_T2_Q5_PartB | 2,271   | 200 | A   | 1,270   | 1,248 | A   | 1,657   | 861 | A   | 1,874     | 420 | A   |
| g8a_T2_Q5_PartC | 2,272   | 201 | A   | 1,273   | 1,247 | A   | 1,656   | 864 | A   | 1,875     | 420 | A   |
| g8b_T1_Q1_Group | 2,080   | 241 | A   | 1,219   | 1,163 | A   | 1,522   | 860 | A   | 1,637     | 458 | A   |
| g8b_T1_Q2       | 2,085   | 239 | A   | 1,222   | 1,164 | A   | 1,521   | 865 | A   | 1,642     | 455 | A   |
| g8b_T1_Q3       | 2,085   | 240 | A   | 1,218   | 1,169 | A   | 1,523   | 864 | A   | 1,642     | 456 | A   |
| g8b_T1_Q4       | 2,087   | 240 | A   | 1,221   | 1,168 | A   | 1,524   | 865 | A   | 1,641     | 457 | A   |
| g8b_T1_Q5_PartA | 2,080   | 239 | A   | 1,214   | 1,166 | A   | 1,519   | 861 | A   | 1,636     | 456 | A   |
| g8b_T1_Q5_PartB | 2,086   | 239 | A   | 1,216   | 1,170 | A   | 1,523   | 863 | A   | 1,640     | 457 | A   |
| g8b_T1_Q6_PartA | 2,067   | 234 | A   | 1,208   | 1,154 | A   | 1,507   | 855 | A   | 1,630     | 448 | A   |
| g8b_T1_Q6_PartB | 2,084   | 239 | B+  | 1,219   | 1,166 | A   | 1,520   | 865 | A   | 1,639     | 455 | B+  |
| g8b_T2_Q1_Group | 2,065   | 234 | A   | 1,206   | 1,155 | A   | 1,512   | 849 | A   | 1,624     | 456 | A   |
| g8b_T2_Q2_PartA | 2,079   | 236 | A   | 1,215   | 1,162 | A   | 1,519   | 858 | A   | 1,636     | 457 | A   |
| g8b_T2_Q2_PartB | 1,912   | 207 | A   | 1,106   | 1,069 | A   | 1,429   | 746 | A   | 1,518     | 409 | A   |
| g8b_T2_Q3       | 2,067   | 235 | A   | 1,208   | 1,155 | A   | 1,509   | 854 | A   | 1,625     | 455 | A   |
| g8b_T2_Q4_PartA | 2,073   | 234 | A   | 1,209   | 1,160 | A   | 1,516   | 853 | A   | 1,633     | 452 | A   |
| g8b_T2_Q4_PartB | 2,080   | 235 | A   | 1,215   | 1,162 | A   | 1,519   | 858 | A   | 1,636     | 455 | A   |
| g8b_T2_Q5_PartA | 1,917   | 212 | A   | 1,113   | 1,075 | A   | 1,421   | 767 | A   | 1,526     | 410 | A   |
| g8b_T2_Q5_PartB | 1,919   | 214 | A   | 1,116   | 1,076 | A   | 1,423   | 769 | A   | 1,530     | 410 | A   |
| g8b_T2_Q5_PartC | 1,915   | 214 | A   | 1,111   | 1,077 | A   | 1,421   | 767 | A   | 1,527     | 409 | A   |

\*R = reference. F = focal.