

Examining Confusing and Unfamiliar Words on the STE MCAS

Stephani Burton

University of Miami, FL

Author Note

The research reported here was conducted as part of the Learning about Open Response science test items and English Learners (LOREL) project and conducted in collaboration with Tracy Noble, PI and Ann S. Rosebery, co-PI. This project was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A150218 to the Technical Education Research Centers, Inc. The opinions expressed are those of the author and do not represent views of the Institute or the U.S. Department of Education.

Examining Confusing and Unfamiliar Words on the STE MCAS

Although the Grade 5 STE MCAS ostensibly measures students STE knowledge and skills, it becomes an English language proficiency assessment in part because key words in the text of items are confusing and unfamiliar to EL students. The report herein describes a set of words found in the constructed-response test items on the Grade 5 STE MCAS from 2003-2015 that are identified by teachers of Grade 5 ELs as likely to be unfamiliar to their EL students, *and* by Grade 5 EL students as confusing to them. In the sections that follow, we describe our study and the set of words identified in this way.

Method

We collected three sets of data to explore ELs' understandings of the words included in the STE OR test items. The three data sets include: a list of the words (drawn from 15 MCAS STE open response test items) coded by expert teachers as most unfamiliar to 5th grade ELs; the words ELs most frequently circled as confusing in those 15 items; and ELs' definitions of selected words in those items. We established cut scores (an unfamiliar word score of at least 4.5 on a scale from 0 to 8 points) for the entire list of words coded as unfamiliar by teachers. This process narrowed the list to the top 163 most unfamiliar words. Coincidentally, there were also a total of 163 words that ELs circled as confusing. These two lists of words comprise the first two data sets.

We identified the intersection between the list of words coded by teachers and the words ELs circled, to compile a list of all the words that occurred in both data sets. The process of synthesizing the two data sets resulted in one list of the words that are confusing to students and that teachers expect their EL students will find unfamiliar. The process yielded a list of exactly 100 words that comprise the "Confusing and Unfamiliar Words" list analyzed in this report. We

did not select 100 of the most confusing and unfamiliar words by design. The cut scores and measures in our methodological approach happened to result in exactly 100 words.

In this report, we use the third data set, ELs' definitions of challenging words, to understand why ELs experienced some of these 100 words as confusing and unfamiliar. These definitions are used throughout the report to demonstrate how ELs' interpret challenging words in the OR item text. When we include ELs' definitions in this report, it is for illustrative purposes only. Appendix B also provides the average scores for unfamiliar words defined by ELs.

There are many ways to explore and analyze the words included in the Confusing and Unfamiliar Words list. We identified four major categories of words represented in this list and bulleted below, in order to better understand how these words impact ELs' understandings of and interactions with the STE OR items. We identified these categories by reviewing the words in the list and examining their functions and features in relation to research in the field, feedback from our Educators' Working Group (EWG), a group of teachers and specialists who work with 5th grade ELs in MA schools, and the ways they appear and are used in the item text. The categories themselves are not measures of word difficulty but the categories can help explain why ELs experienced certain words as confusing. Collectively, our four categories draw from research-based and practice-based perspectives that describe the ways ELs often encounter and experience these types of words.

- Low frequency non-technical words (LFNT)
- Technical terms
- Pre-stem text
- Direction words

LFNT words

LFNT words are words that occur infrequently in 5th grade texts, and do not have a primarily scientific meaning (Kachchaf et al., 2016). We expected the Confusing and Unfamiliar words list to include many LFNT words because ELs encounter these words, in English, with minimal frequency in material written for their grade level. Additionally, LFNT words are often not explicitly taught because they are not technical, STE terms. For these two reasons, this portion of the Confusing and Unfamiliar Words list may be problematic for ELs specifically.

Technical terms

Technical terms are individual words or compound terms that, have a have a primary scientific meaning or are used in a technical way when they appear in the Massachusetts STE Frameworks. In contrast to LFNT words, technical terms refer to STE and technical concepts, so they are often included in 5th graders' academic texts and therefore taught explicitly in science class. Including this category allows us to explore ELs' understanding of the technical terms that represent the state assessment's target STE content. We expected the Confusing and Unfamiliar words list to include very few technical terms because these words are often selected as vocabulary words that are taught explicitly in science class.

Pre-stem text

The pre-stem text is relatively consistent across the items in which it appears. The pre-stem text is a set of instructions that includes expectations for the test taker. Each of the words included in this narrative comprise the pre-stem text. ELs in our study frequently encountered this text because it appears as an introductory text to many of the STE OR items. The EWG explained that ELs may interact with these words in two distinct ways: ELs would either 1) ignore the pre-stem text completely or 2) spend a significant amount of time working to decode

and comprehend the pre-stem text. We were eager to examine the representation of these words in the Confusing and Unfamiliar list given the feedback provided by the EWG. An example of pre-stem text appears below:

Question 21 is an open-response question.

- *BE SURE TO ANSWER AND LABEL ALL PARTS OF THE QUESTION.*
- *Show all your work (diagrams, tables, or computations) in your Student Answer Booklet.*
- *If you do the work in your head, explain in writing how you did the work.*

Write your answer to question 21 in the space provided in your Student Answer

Booklet

Direction words

Like the pre-stem text, direction words are also prevalent across items. Direction words tell the test takers what they need to do to answer the OR question. The pre-stem text and direction words both instruct the test taker, but direction words are more specific to the item. Direction words most frequently appear in the item's writing prompt but sometimes appear in the body of the OR item text. When direction words appear in the writing prompt, they typically instruct the test taker to *explain, identify, or describe* the targeted STE content. When found in the body of the OR item text, direction words instruct the test taker to step into the role of a character in the narrative who is trying to accomplish an STE-related task. Teachers in the EWG often coded these words as unfamiliar. Therefore, we expected these words to appear in the Confusing and Unfamiliar words list.

Results

Table 1.1 below illustrates the percentage of Confusing and Unfamiliar Words accounted for by each of the four aforementioned categories. The LFNT and Technical categories are mutually exclusive but both categories may overlap with the Pre-Stem and Direction categories. It is also possible to categorize the same words as both Pre-Stem and Direction words. Table 1.2 lists the words included in each category and quartile. Quartiles are determined based on frequency. The frequency for each word is the frequency with which teachers coded it as unfamiliar and students circled it as confusing. Some words appear in the texts of test items more often than others. The words that appear more often have more opportunities to be circled by students and coded by teachers. The 25 words that are most often circled by students and coded by teachers are included in the fourth quartile. There are 25 words in each quartile and as you move from the fourth Quartile (Q4) to the first quartile (Q1), you move from higher frequency to lower frequency words.

Table 1.1. Percentage of Confusing and Unfamiliar Words by Category

Category	Percentage of Confusing and Unfamiliar Words
Low Frequency Non-Technical Terms (LFNT)	25%
Technical Terms	7%
Pre-Stem Text	7 %
Direction Words	9%

Table 1.2. List of Confusing and Unfamiliar Words by Category and Quartile

Quartile	Low Frequency Non-Technical Terms (LFNT)	Technical Terms	Pre-Stem Text	Direction Words
Q4	computations booklet droplets spout teakettle tidepool fasteners hike illustrated	diagrams	computations booklet diagrams provided tables	identify describe include
Q3	construct rearranged forecast oily power plant primary seashore transformed battery	chemical energy electrical energy environment predicted		construct attach
Q2	brackets classify burrow factors		head*	build classify explain improve
Q1	open-response suspended tape measure	organisms stream	open-response	

* From the pre-stem instructions: “If you do the work in your head, explain in writing how you did the work.” We expect that the word “head” may have been included here because of the unfamiliarity of the phrase it is part of.

The Abundance of confusing and unfamiliar LFNT words

LFNT words account for 25% of the words included in the Confusing and Unfamiliar words list. The findings here suggest that ELs have difficulty with the volume of infrequently used words that do not convey the STE content included in the OR items.

Though these words do not have primarily scientific technical definitions, they often represent important information needed to comprehend the item text. Many of the words in this

list describe the context that frames the STE content. This is extremely concerning as the words in this list do not represent the STE content being assessed, but can create a barrier for ELs to access the STE content. For example, the LFNT words: *droplets*, *spout*, and *teakettle* are all used to describe a scenario presented in the “Teakettle” OR item. It is nearly impossible to comprehend the item text without understanding the LFNT words students found to be confusing and unfamiliar. An example of this is provided below. This is the text for the “Teakettle” OR item when the confusing and unfamiliar words are obscured:

A cook notices a [REDACTED] full of water on a stove. There is a cold window close to the [REDACTED] of the kettle. The water begins to boil and [REDACTED] droplets begin to form on the window.

a. Describe in detail what is happening to the water inside the kettle.

b. Why do the water [REDACTED] form on the window? Be sure to explain in detail.

This is the text for the “Teakettle” OR item in its original form:

A cook notices a teakettle full of water on a stove. There is a cold window close to the spout of the kettle. The water begins to boil and water droplets begin to form on the window.

a. Describe in detail what is happening to the water inside the kettle.

b. Why do the water droplets form on the window? Be sure to explain in detail.

During interviews, we asked ELs we to define some of the words used in the teakettle text. The table below shows some of the definitions ELs provided. These illustrative examples provide insight into these ELs’ misinterpretations of the text.

Table 1.3 Definitions Given in Interviews for Confusing and Unfamiliar Words

EL (pseudonym)	Word from Teakettle Text	Definition Provided
Noemi	Teakettle	<i>“I think a little like, a big like um, a big example of like, a cup?”</i>
Esther	Spout	<i>“It means to grow like in growing places...like a tree is growing a spout”</i>
Eduard	Droplets	<i>“Droplets is when water falls on the floor or something”</i>

Technical terms

Technical terms account for 7% of the words included in the Confusing and Unfamiliar words list. These words represent vocabulary with technical and primarily scientific definitions. This percentage is relatively low when compared to the percentage of LFNT words in the Confusing and Unfamiliar Words list. This is important to note because the LFNT and technical terms are mutually exclusive.

Pre-stem text

The pre-stem text terms account for 7% of the Confusing and Unfamiliar words. 5 of the 7 pre-stem terms appear with high frequency in the fourth quartile. The pre-stem text appears in several items with nearly identical wording. As a result, the high representation of words from the pre-stem text in the fourth quartile is in part due to the frequency with which students in this study encountered the text.

The pre-stem text provides instructions for the test taker and test takers may presume the text is relevant to the STE OR items and what they are expected to do. But the text is not always relevant to the item. The text instructs the test taker to show all of his or her “diagrams, tables, or computations” even when diagrams, tables, and/or computations are neither needed nor appropriate for the written response.

The pre-stem text provides generalized instructions for the test taker and is the same text used for the Mathematics MCAS test. This means that some of the instructions provided may be relevant for the STE MCAS test takers, while others are not. In some cases, a majority of the text may be completely unnecessary and unhelpful for the test taker. Additionally, some of the problematic terms, like *computations*, are not words readily associated with STE concepts in the OR items on the STE MCAS, most of which require no computations at all.

Direction words

Direction words account for 9% of the Confusing and Unfamiliar words list. Direction words communicate what the test taker is expected to do in ways that are specific to a particular item. We hypothesize that the most frequently used direction words: *identify*; *describe*; and *explain*, are confusing and unfamiliar because the same words are used in different ways in different test items. In some items, the word *identify* signals a task in which the test taker must select an object or set of objects from what is provided in a given item. In other instances, *identify* signals a task in which the test taker must generate their own response absent of information provided in the text. The words *identify*, *describe*, and *explain* are often included in the items, but represent a range of tasks. The interchangeable use of polysemous terms can make them challenging for ELs to accurately define and apply in their responses. For example, the word *identify* appears in the prompt for the “Rosa’s Well” OR item as shown below:

a. Identify which of these objects Rosa can use with her tape measure to find the depth of the well.

After reading this text, interviewers asked 5th grade ELs who participated in the study, “*What does the word “Identify” mean in this sentence: Identify which of these objects Rosa can use with her tape measure to find the depth of the well?*” ELs responded with different definitions for the same word in the context of the item. One student said identify means, “like, like to pick, to pick a, like a object that Rosa needs to, needs to fix the well.” However, another student offered a different definition for identify saying, “identify is to like find out.” Despite encountering the word in the same context, ELs offered different definitions of the word. There are multiple different senses in which a polysemous word is used. Students accustomed to seeing it in one way may not be able to easily pivot to another sense of the word in another test item.

Appendix A

Confusing and Unfamiliar Words List by Category

LFNT	Technical Terms	Pre-Stem Text	Direction Words
Computations	Diagrams	Computations	Identify
Booklet	Chemical energy	Booklet	Describe
Droplets	Electrical energy	Diagrams	Include
Spout	Environment	Provided	Construct
Teakettle	Predicted	Tables	Attach
Tidepool	Organisms	Head*	Build
Fasteners	Stream	Open-response	Classify
Hike			Explain
Illustrated			Improve
Construct			
Rearranged			
Forecast			
Oily			
Power plant			
Primary			
Seashore			
Transformed			
Battery			
Brackets			
Classify			
Burrow			
Factors			
Open-response			
Suspended			
Tape measure			

* From the pre-stem instructions: *“If you do the work in your head, explain in writing how you did the work.” We expect that the word “head” may have been included here because of the unfamiliarity of the phrase it is part of.*

Appendix B

Average EL 1-4 Word Definition Scores = Average Definition Score* for all ELs asked to define this word in Interviews.

(Score Categories: 0 = Not Accepted, 1 = Partially Accepted, 2 = Accepted as Correct Definition)

Tide Pool	0	Unfamiliar	LFNT
Main Characteristic	0	Unfamiliar	
Most Likely	0	Unfamiliar	
Explain the Reasons	0	Unfamiliar	
Diagrams	0	Unfamiliar	
Spout	0	Unfamiliar	LFNT
Average	0	Unfamiliar	
Front	0	Unfamiliar	
Webbed	0	Unfamiliar	
Burrowing	0	Unfamiliar	LFNT
Damp	0	Unfamiliar	
Seashore	0	Unfamiliar	LFNT
Brackets	0	Unfamiliar	LFNT
Best	0.14		
Chemical Energy	0.17	Unfamiliar	
Turbine	0.18	Unfamiliar	LFNT
Coils	0.19	Unfamiliar	LFNT
Predicted	0.20	Unfamiliar	
Forecast	0.20	Unfamiliar	LFNT
Safety Factors	0.20	Unfamiliar	LFNT
Prairie	0.23	Unfamiliar	
Burrows	0.25	Unfamiliar	LFNT
Kettle	0.25	Unfamiliar	
Deposits	0.25	Unfamiliar	
Fasteners	0.29	Unfamiliar	LFNT
Thick	0.29	Unfamiliar	
Paddles	0.29	Unfamiliar	
Cover	0.29	Unfamiliar	
Not Unusual	0.33	Unfamiliar	LFNT
Features	0.43	Unfamiliar	
Classifying	0.50	Unfamiliar	LFNT
Castings	0.50	Unfamiliar	LFNT
Earthworms	0.50		LFNT
Identify	0.57	Unfamiliar	
Open Response	0.57	Unfamiliar	LFNT

Help an Ecosystem	0.67		
Droplets	0.67	Unfamiliar	LFNT
Conditions	0.67	Unfamiliar	
Oily	0.71	Unfamiliar	LFNT
Generator	0.72	Unfamiliar	LFNT
Shaped	0.75	Unfamiliar	
Surface	0.75	Unfamiliar	
Natural Processes	0.75	Unfamiliar	
Organisms	0.85	Unfamiliar	
Precipitation	0.86	Unfamiliar	
Primary Source	0.92	Unfamiliar	LFNT
Explain Why	1.00	Unfamiliar	
Backyard	1.00		LFNT
Tape Measure	1.00	Unfamiliar	LFNT
Battery Powered	1.17	Unfamiliar	LFNT
Marker	1.50		LFNT
Weather	1.50		
T-Storm	1.57	Unfamiliar	LFNT
Depth	1.60	Unfamiliar	
Battery	1.67	Unfamiliar	LFNT

*Each word definition score is an average of at least 3 EL 1-4 student responses

Appendix C: Alphabetical Confusing and Unfamiliar Words List

1. Identify
2. Diagrams
3. Occur
4. Several
5. Columns
6. Computations
7. Depth
8. Webbed
9. Booklet
10. Describe
11. Droplets
12. Features
13. Prairie
14. Provided
15. Spout
16. Tables
17. Teakettle
18. Tide pool
19. During
20. Fasteners
21. Hike
22. Illustrated
23. Include
24. Kettle
25. Materials
26. Processes
27. Rearranged
28. Supplies
29. Transferred
30. Another
31. Attach
32. Chemical energy
33. Construct
34. Electrical energy
35. Environment
36. Forecast
37. Oily
38. Physical
39. Piece
40. Power plant
41. Predicted
42. Primary
43. Seashore
44. Shelter
45. Transformed
46. Appear
47. Area
48. Average
49. Based
50. Battery
51. Between
52. Brackets
53. Build
54. Burrow
55. Cart
56. Classify
57. Clearing
58. Club's
59. Conditions
60. Damp
61. Deposits
62. Directly
63. Driven
64. Either
65. Energy
66. Entire
67. Explain
68. Factors
69. Fog
70. Form
71. Gathered
72. Head*
73. Highest
74. However
75. Improve
76. Key
77. Layer
78. Likely
79. Natural
80. Northeast
81. Open-response
82. Organisms
83. Paddles
84. Properties
85. Safety
86. Shape
87. Source
88. Stacked
89. Steel
90. Stream
91. Sure
92. Survive
93. Suspended
94. Tape measure
95. Thick
96. Those
97. Tools
98. Type
99. Unusual
100. Upright

* From the pre-stem instructions: *“If you do the work in your head, explain in writing how you did the work.” We expect that the word “head” may have been included here because of the unfamiliarity of the phrase it is part of.*