~Chapter 14~

Text Complexity: A Study of STAAR Readability

Michael Lopez, Graduate Student
University of Mary Hardin-Baylor

Jodi Pilgrim, Ph.D.
University of Mary Hardin-Baylor

Abstract

A critical component of English Language Arts and Reading standards includes the student's ability to comprehend increasingly more complex text by applying a flexible range of metacognitive reading skills. This study examined the text complexity of reading passages on the 2013-2015 State of Texas Assessments of Academic Readiness (STAAR) materials for grades 3 –5. Readability formulas were used to determine the reading levels of selected passages and to analyze text complexity as it relates to the state reading tests. Findings indicated that high levels of readability found in materials may be problematic for teaching and learning and as a result, teachers must differentiate texts used in their curriculum and instruction to optimize the learning environment.

Keywords: *text complexity, readability, STAAR*

Incorporated among the strands for the 2016 Literacy Summit was *Implementing Texas State Standards/Close Reading and Text Complexity*. High stakes assessments, including the Common Core (Fisher, Frey, & Lapp, 2012) and the State of Texas Assessments of Academic Readiness, have placed much attention on text complexity, which has been associated with raising the rigor in reading achievement. Proponents of increasing the rigor of texts argue a gap exists between texts read in school versus texts required for success in college and careers (Fisher, et al., 2012). In an effort to examine the text complexity of Texas's state assessment, the study described in this paper examined the readability of the State of Texas Assessment of Academic Readiness (STAAR) released reading passages for grades 3-5.

What is text complexity? The topic itself is

> (1) vocabulary/use of language may be more varied and challenging because it is nonliteral/figurative, abstract, or academic/technical; (2) sentence structures may be more varied, dense, and sophisticated; (3) the author's use of literary elements/devices, rhetorical strategies, organizational patterns, and text features may be more nuanced or sophisticated; (4) the topic/content may be less familiar or more cognitively demanding; and (5) relationships among ideas may be less explicit and require more interpretation, reasoning, and inferential thinking to understand the subtlety, nuances, and depth of ideas. (TEA, 2013, p. 1)

Fisher et al. (2012) defined three components of text complexity, which include qualitative dimensions, quantitative dimensions, and reader and task considerations. This paper focuses on

Free Printable Form

View & Print Your Form Now FindFormsFast

length, and frequency of words used in the English language. In this study, the State of Texas Assessment of Academic Readiness was examined using readability formulas in order to further understand the difficulty levels of the STAAR.

State of Texas Assessment of Academic Readiness – STAAR

In 2012, Texas students began taking the new statewide standardized test called the State of Texas Assessment of Academic Readiness, or the STAAR, which was the successor to the Texas Assessment of Knowledge and Skills (TAKS). Significant differences marked the transition from one test to the other, and increased rigor is among one of the differences. The conversion from one test to the other began in 2007, when the Texas Legislature ordered the switch from TAKS to end of course tests for high school students, starting with ninth-graders in 2012. Two years later, the legislature mandated the test changes for grades three through eight. The new tests were supposed to be harder, implemented a four-hour time limit, and focused on one year's academic content (Weiss, 2012). In addition, STAAR was supposed to better assessment the national phenomenon of progress toward postsecondary readiness at every grade level. Increased rigor with the STAAR included lengthier tests at most grades and subjects, more rigorous test items, and the assessment of skills at a greater depth and level of cognitive complexity.

The state assessments and accountability system have caused much controversy in Texas. For example, parents, teachers, and 23 school districts in Texas are in the midst of protesting current testing and moving ahead with plans to create a new accountability system that doesn't depend on standardized tests (Stanford, 2014). Stanford (2013), who blogs about standardized testing, also claims that "nowhere is the movement against high-stakes testing as strong as it is in Texas where all this started" (para. 4). The blog states that 86% of Texas school boards have adopted resolutions which oppose high-stakes testing. It appears that not only is over-reliance on testing an issue, the current rigor of the test has frustrated both teachers and students.

According to TEA's (2013) website, "if we want students to do on-grade level work, we must teach them how to "tackle" increasingly complex texts each year" (slide 10). The STAAR addresses Readiness and Supporting standards in order to support the state's goal to become one of the top 10 states for producing college and career ready (CCR) students with its 2020 graduating class (TEA, 2013). Texas performance standards include Level 1, which reflects unsatisfactory performance, Level II, which reflects satisfactory academic performance, and Level III, which reflects Advanced Academic Performance. The performance standards for STAAR Reading test are relevant to this study, as the criteria to meet these standards indicates a potential problem with test rigor. This study specifically addresses the rigor of the STAAR Reading tests for elementary aged students (Grades 3-5). Our concern is that the reading level of the tests may be at the frustration level for many students. Therefore, the rigor of these tests was examined in depth. Level II attainment, or passing, requires students to answer only half or a little more than half of the rigorous test questions correctly. Table 1 presents data which exemplifies the pass rates for students since 2012.

Table 1
Raw Score Conversion Data for the STAAR Reading for Level II Attainment

Year/Grade	Raw Score	Converted Percentage
2012/Grade 3	20/40	50%
2012/Grade 4	23/44	52%
2012/Grade 5	25/46	54%
2013/Grade 3	20/40	50%
2013/Grade 4	24/44	55%
2013/Grade 5	26/46	57%
2014/Grade 3	21/40	53%
2014/Grade 4	23/42	55%

Source: TEA STAAR Raw Score Conversion Tables, 2011-2015

TEA initially planned to "phase-in" higher standards/expectations for these levels. In other words, with changes in phases, an increased score for the levels would be required. However, level II attainment, or passing, was initially set very low and the state has been in Phase 1 of the plan for the past four years. TEA recently announced that performance standards have been scheduled to move to phase-in 2 passing standards this year, but instead of the rigorous advances in standards every few years, the new proposed progression includes smaller, predictable increases every year through the 2021-2022 school year (TEA, 2015). As presented in Table 1, student pass rates do reflect the rigor of the test. A test that requires a pass rate equivalent to a 50% may be too difficult for Texas students. The present study developed from the work of Szabo and Sinclair (2012), who analyzed the Texas Education Agency's sample pilot test questions released prior to the spring of 2012. Szabo and Sinclair (2012) used readability formulas and determined the passages to be written at a level too high for the tested grade levels. The purpose of this present study was to further investigate the readability of the tests and student performance on the tests to determine text complexity and reading levels.

Readability Formulas

Fisher, Frey, and Lapp (2012) argue that readability impacts text complexity. During the 1920s, it was discovered that there was a way to use the difficulty of a word and the length of the sentence to estimate the difficulty level of the text. By the 1950s, Rudolf Flesch, Edgar Dale, and Jeanne Chall brought their readability formulas into general usage. By the 1980s, there were 200 formulas and over a thousand published studies on readability formulas (DuBay, 2004). The formulas that were chosen for the present study to determine the grade level of the STAAR reading passages are described below.

Lexile Measure

A Lexile measure for a text is the analysis of the word level difficulty and the complexity of the sentence. The measure is a numerical value from 200L to 1700L. The lower the number, the easier the text is for readers. The Lexile measure was developed in 1989 and is currently used across the

Flesch-Kincaid Readability

The formula is based upon Flesch's reading ease formula that was developed in 1943. J. Peter Kincaid, in 1975 while under contract with the US Navy, expanded upon Flesch's original work. The formula uses the word and sentence length per 100 words to calculate a United States grade level (Kincaid, Fishbone, Rogers & Chsisom, 1975). The formula can now be found within many popular word processing software programs, including Microsoft Word.

SMOG Readability

Developed by Dr. McLaughlin in 1969, the SMOG grade is obtained by counting the first 10 consecutive sentences near the beginning of a text, ten in the middle and ten near the end. Within those thirty sentences, count only the polysyllable words and round up to the nearest square root number then add three to determine the number of years of education a reader would need in order to fully comprehend the text (McLaughlin, 1969).

Gunning Fog Readability

In 1952, Robert Gunning was an American businessman when he developed the Fog index. The index is an estimation on the number of formal educations years a reader needs to have in order to comprehend a text during the initial reading. He helped editors and writers of newspapers and popular magazine write for their audience by eliminating the "fog" (DuBay, 2004).

Fry Readability

Edward Fry developed the Fry graph, while working to help educators in Uganda teach English as a second language (DuBay, 2004). He would later expand the graph to include primary and college grade levels. The reader selects a 100 word passage and calculates the average number of sentences and plots the number on the y axis. The average number of syllables in the 100 word sample is placed on the y axis. The intersection of the two axis provides an estimate of the grade level.

Raygor Readability

Alton Raygor readability index was

with six or more letters are counted and averaged for the three samples. The results are then plotted onto the Raygor graph (Szabo & Sinclair, 2012).

Methodology

The readability formulas used in this study included the Lexile, Flesch Kincaid, SMOG, Gunning Fog, Fry, and Raygor. These were the formulas used by Szabo and Sinclair (2012), with the addition of the Lexile formula. The inclusion of the Lexile measure enabled the comparison of the student's current grade level with their Lexile range measure as well as the Lexile measures which have grade level ranges aligned with college and career readiness expectations (Daggett & Pedinotti, 2014; Williamson, 2004).

Online readability calculators were used in this process. Each digital (state released) reading passage found on TEA's website was copied and pasted into a word document. The passages were checked for word spacing, spelling, and removal of non-ASCII characters (ie., quotes, ellipses, em-and en-dashes). Each passage was calculated using readability formulas from four different free, online websites: www.lexile.com, www.webpagefx.com/tools/read-able/, https://readability-score.com, www.readabilityformulas.com. The formula average is the grade level average of the FleschKincaid, SMOG, Gunning Fog, Fry, and Raygor readability results. (NOTE: Due to copyright the 2014, 4th Grade reading passage 1 was not published with the released assessment. The lexile measure was obtained by using the author and title of the reading passage).

Findings

Since 2012, approximately 25 percent of elementary school students assessed in reading have failed to meet satisfactory achievement (TEA, 2012, 2013, 2014, 2015). Study of the 2015 STAAR Grades 3 – 5 reading passages reveals a high readability for those grade levels. Overall, for each grade level, the reading passages were one to three grade levels above the students' current grade level. Of the nine STAAR reading assessments analyzed, only the 2015 5th grade reading assessment was at the students' current grade level. Findings are

The students in the 3rd grade have been assessed at least two to three grade levels higher than their current grade. Over the past three years, the Lexile measure has ranged from 380L to 980L which places them in the upper band of the 5th grade to low range of 6th grade. The 4th grade students have been assessed two grades above their current grade, however in 2015 it was reduced to one grade level above their current grade. Overall, their Lexile measure has ranged from 380L to 1050L, which puts the text at the upper band of the 8th grade to low range of 9th grade band. The 5th grade student has been assessed one grade level above their current grade in two of the three years. As mentioned previously, in 2015 the students were assessed at a passage average of the 5th grade. The Lexile measure over the past three years has ranged from 350L to 1050L which, like the 4th grade student, is located at the upper band for 8th grade and beginning band for the 9th grade. The quantitative analysis of each grade level can be found in Table 2.

Table 2
Readability of STAAR Passages

3rd Grade

2011 (SAMPLE)

	Lexile	Formula Avg	Overall Avg
1st Passage	650	3	
2nd Passage	840	6	5

2013

	Lexile	Formula Avg	Overall Avg
1st Passage	830	6	
2nd Passage	380	2	
3rd Passage	910	7	
4th Passage	520	3	
5th Passage	680	5	5

2014

	Lexile	Formula Avg	Overall Avg
1st Passage	830	6	
2nd Passage	470	3	
3rd Passage	850	7	
4th Passage	700	4	
5th Passage	950	7	5

2015

	Lexile	Formula Avg	Overall Avg
1st Passage	840	7	
2nd Passage	900	6	
3rd Passage	980	4	
4th Passage	560	4	
5th Passage	940	7	6

4th Grade

2011 (SAMPLE)

	Lexile	Formula Avg	Overall Avg
1st Passage	890	7	
2nd Passage	300	2	
3rd Passage	880	7	5

2013

	Lexile	Formula Avg	Overall Avg
1st Passage	1000	7	
2nd Passage	580	5	
3rd Passage	380	3	
4th Passage	1050	8	
5th Passage	950	9	
6th Passage	420	3	6

2014

	Lexile	Formula Avg	Overall Avg
1st Passage	430		
2nd Passage	900	7	
3rd Passage	440	2	
4th Passage	1040	8	
5th Passage	930	7	
6th Passage	640	4	6

2015

	Lexile	Formula Avg	Overall Avg
1st Passage	340	3	
2nd Passage	950	8	
3rd Passage	850	5	
4th Passage	790	4	
5th Passage	930	6	
6th Passage	920	7	5

5th Grade

2011 (SAMPLE)

	Lexile	Formula Avg	Overall Avg
1st Passage	680	4	
2nd Passage	1060	9	7

2013

	Lexile	Formula Avg	Overall Avg
1st Passage	520	3	
2nd Passage	990	8	
3rd Passage	350	2	
4th Passage	930	7	
5th Passage	660	5	
6th Passage	1050	9	6

2014

	Lexile	Formula Avg	Overall Avg
1st Passage	810	7	
2nd Passage	680	5	
3rd Passage	660	4	
4th Passage	890	7	
5th Passage	990	7	
6th Passage	910	8	6

2015

	Lexile	Formula Avg	Overall Avg
1st Passage	470	3	
2nd Passage	880	6	
3rd Passage	790	5	
4th Passage	730	4	
5th Passage	990	7	
6th Passage	510	2	5

[ISSN 2168-0019 online]