METAMETRICS RESEARCH BRIEF

The 2020 Text Complexity Continuum in Grades 1-12

Xavier J. Webb Ed.D., Lead Researcher; Sue Steinkamp, Ph.D., Senior Vice President Research and Development; Heather Koons, Ph.D., Director of Research and Product Integration; Eleanor E. Sanford-Moore, Ph.D., Research Advisor; Robert Baker Ph.D., Principle Psychometrician; Amber Hinson B.A., Manager, Research Analysis Services

July 2022

OBJECTIVE

Since 2005, MetaMetrics has conducted a number of research studies to describe the typical reading demands that students encounter in the classroom. These studies were implemented in stages: high school (Grades 9-12) in Fall 2005; middle school (Grades 6-8) in Spring 2007; and, elementary school (Grades 1-5) in Spring 2008. The results of these studies were combined to establish a systematic continuum of text complexity from Grades 1 through 12 (see Williamson, Koons, Sandvik, & Sanford-Moore, 2012). Since the completion of the study by Williamson et al. (2012) many states adopted and implemented the Common Core State Standards (CCSS), Next Generation Science Standards, National Curriculum Standards for Social Studies, and have undergone other curricular changes. In light of these changes, the objective of the current study was to replicate the research conducted by Williamson et al. (2012) and compare the results from each study.

The following research questions were investigated:

1) Does median text complexity increase with grade level for the texts sampled?
2) Does text complexity vary within grade?
3) Does text complexity vary across content areas?
4) Has textbook complexity in Grades 1 through 12 changed since the text demand study completed in 2012 (Williamson, Koons, Sandvik, & Sanford-Moore)?

METHODS

Participants (Units of Analysis)

The units of analysis in the study were textbooks used in U.S. public schools in Grades 1 through 12. A total of 309 textbooks across Grades 1-12 were used in the study which included texts from the content areas of reading (n = 77), mathematics (n = 87), science (n = 77), and social studies (n = 68). *Table 1* provides the sample sizes for the text collections for each grade level.

Table 1. Sample sizes for the Grades 1-12 text collections.

	Subject				
Grade	**Reading**	**Mathematics**	**Science**	**Social Studies**	**Total**
1	21	7	4	8	40
2	7	7	4	8	26
3	7	7	4	4	22
4	5	7	4	5	21
5	5	7	5	5	22
6	8	6	6	8	28
7	4	6	6	8	24
8	4	6	6	4	20
9	4	11	13	6	34
10	4	11	9	4	28
11	4	11	8	4	27
12	4	1	8	4	17
Total	77	87	77	68	309

Procedure

From 2012 to 2020, MetaMetrics systematically collected commonly used reading, mathematics, science, and social studies textbooks across Grades 1 through 12 and measured their text complexity using the Lexile® Framework for Reading and the Lexile® Text Analyzer. To guide textbook selection, MetaMetrics consulted the Education Market Research (EMR) database to determine which publishers and textbooks had the largest market share in each content area. At each grade, the target sample for each content area was four textbooks. When the textbooks occurred in a series, all books in the series were included in the study. To ensure balanced representation across publishers, at least three of the four textbooks identified were from different publishers. When a publisher specified a grade level for a text, that grade level was used. However, some middle and high school texts were designed for grade bands instead of individual grades by the publisher so MetaMetrics developed a schema to provide a consistent course-to-grade correspondence to apply when a text-grade assignment was not available from

the publisher. The design of the schema was informed by examining the standards and curriculum for the 20 most populous states and using the most common course-to-grade correspondence observed.

Even though the sample sizes for the current study appear to be modest, they in fact represent a large proportion of the textbooks used in K-12 education because there are relatively few textbook publishers and states have limited options for textbook adoption. The current study was limited to textbooks that were designated for use in one particular grade. While this decision may have reduced potential sample sizes, it facilitated the ability to more finely ascribe text complexity exposure to a particular grade-grouping of students. Including textbooks that are designed for a sequence of grades would increase sample size, but would also tend to homogenize the grade-level text distributions and could cause one to infer less increase in text complexity across grades. In addition, it would introduce statistical dependency between the grade-level text distributions.

Digital versions of all textbooks were obtained and analyzed with MetaMetrics' Lexile Text Analyzer. When text is analyzed by MetaMetrics, all electronic files are initially edited according to established guidelines used with the Lexile Text Analyzer software. These guidelines include the removal of all incomplete sentences, chapter titles, and paragraph headings and the running of a spell-check. The Lexile Text Analyzer first examines the features of a piece of text and attempts to determine if it is written for early readers (early-reader texts) or for more advanced readers (upper-level texts). Based on the results of the examination, the Lexile Text Analyzer applies the most appropriate text complexity indicators to the measurement process. There are four text complexity indicators (semantic demand, syntactic demand, decoding demand, and structure demand) and the Lexile Text Analyzer employs estimates of two indicators (semantic and syntactic demand) for upper-level texts and estimates of all four indicators for early-reader texts. The Lexile Text Analyzer then reports a Lexile text measure for the text. The resulting Lexile text measures for all textbooks in the study were then collated and statistically summarized by grade.

Measures

The Lexile® Framework for Reading is a scientific way to match readers with text using the same developmental scale. Lexile text measures (MetaMetrics, 2022; see also Fitzgerald, Elmore, Hiebert, Koons, Bowen, Sanford-Moore & Stenner, 2016; Fitzgerald, Elmore, Koons, Hiebert, Bowen, Sanford-Moore & Stenner, 2015; and Stenner, H. Burdick, Sanford & D.S. Burdick, 2007) are measures of text complexity based on quantitative factors. Independent psychometric studies of the Lexile scale indicate that it is a valid and reliable measure of reader ability and text complexity (Mesmer, 2008; White & Clement, 2001). A Lexile measure is the numeric representation of an individual's reading ability or a text's complexity (or, difficulty) followed by an "L" (for Lexile). The Lexile scale is a developmental scale for measuring person ability and resource complexity, with values ranging from below 200L to above 1600L. Measures can be below 0L for beginning reader materials (e.g., BR150L) to above 1800L for advanced materials. Lexile text measures were used for all analyses. Extensive information about the development of the Lexile Framework for Reading can be found on the Lexile website.

ANALYSES

Each grade-level text collection was analyzed separately. Selected percentiles (5^{th}, 25^{th}, 50^{th}, 75^{th}, 95^{th}) of the text complexity distributions were calculated using SAS PROC UNIVARIATE. Next, the percentiles and grade-level ranges were summarized (see *Table 2*) and used to construct box-and-whisker plots for the grade-level distributions of text complexity (see *Figure 1*). Research Questions 1 and 2 examined whether median text complexity increases with grade level and the extent to which text complexity varies within grade (see *Table 2*). Next, for Research Question 3, to examine whether text complexity varies across content areas, mean text complexity by grade for each content area was plotted (see *Figure 2*). To understand which content area differences were statistically significant, Tukey's HSD Test for multiple comparisons was performed. Research Question 4 investigated whether textbook complexity in Grades 1 through 12 had changed since the 2012 text demand study conducted by Williamson et al. To examine differences between the two studies, percentile ranges from the current study and the 2012 study (Williamson et al.) were used to construct box-and-whisker plots for the grade-level distributions of text complexity (see *Figure 3*). To see whether differences between mean Lexile text measures at each grade for the two studies were statistically significant, Welch's unequal variances *t*-tests were performed.

RESULTS & DISCUSSION

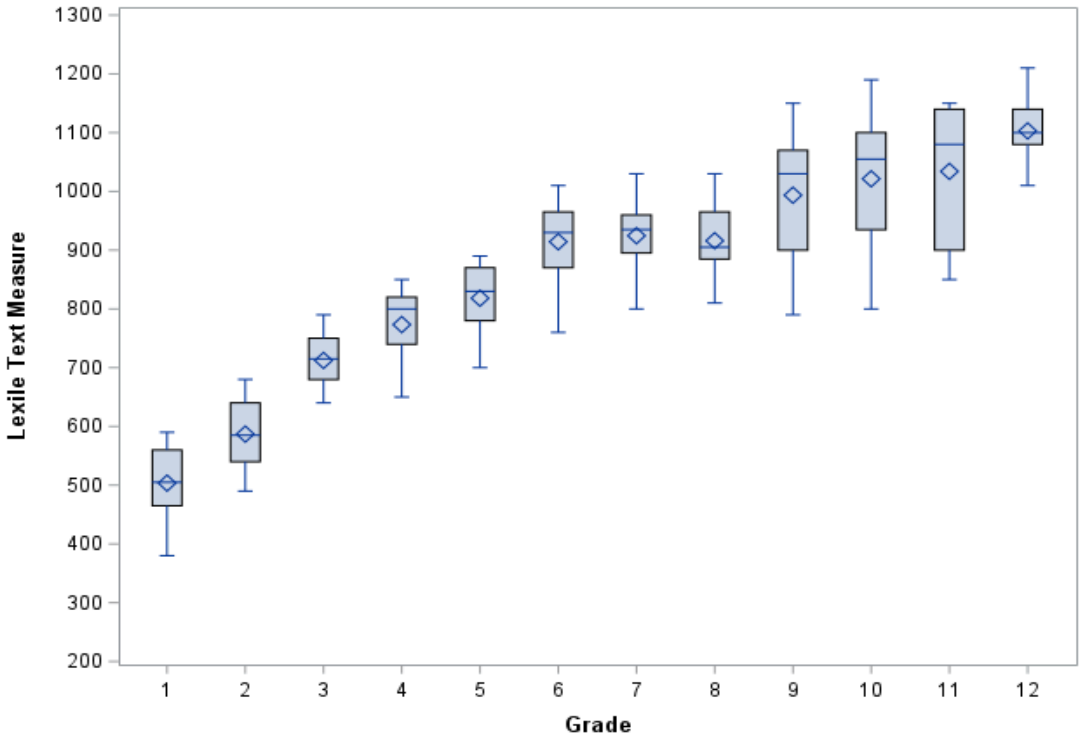
Research Question 1 examined whether median text complexity increases with grade level for the texts sampled. As seen in *Table 2*, median text complexity generally increases with grade. The only point where the increase is not monotonic is between Grades 7 and 8 which may be a reflection of how publishers sometimes designate middle school texts as appropriate for Grades 6-8 instead of a specific grade level. The decrease from Grade 7 to Grade 8 is only 30L and the sample size in Grade 7 (n = 24) is larger than that in Grade 8 (n = 20). As such, it is also possible that this decrease is due to the effects of sampling variability. Furthermore, given the relatively modest sample sizes, some departure from strict monotonicity might be expected. The fact that there is so little departure supports the alternative hypothesis that text complexity generally increases from Grades 1 to 12. Lastly, grade-to-grade increases tend to be larger during the earlier grades than the later ones, with the largest single grade increase (145L) occurring between Grades 2 and 3.

Table 2. Median text complexity measures and interquartile range boundaries, by grade.

Grade	5th Percentile	25th Percentile	Median	75th Percentile	95th Percentile	Range (5th to 95th Percentile)
1	375L	465L	505L	560L	595L	220L
2	490L	540L	585L	640L	680L	190L
3	640L	680L	715L	750L	790L	150L
4	650L	740L	800L	820L	850L	200L
5	700L	780L	830L	870L	890L	190L
6	760L	870L	930L	965L	1010L	250L
7	800L	895L	935L	960L	1030L	230L
8	790L	885L	905L	965L	1035L	245L
9	790L	900L	1030L	1070L	1150L	360L
10	800L	935L	1055L	1100L	1190L	390L
11	850L	900L	1080L	1140L	1150L	300L
12	1010L	1080L	1100L	1140L	1210L	200L

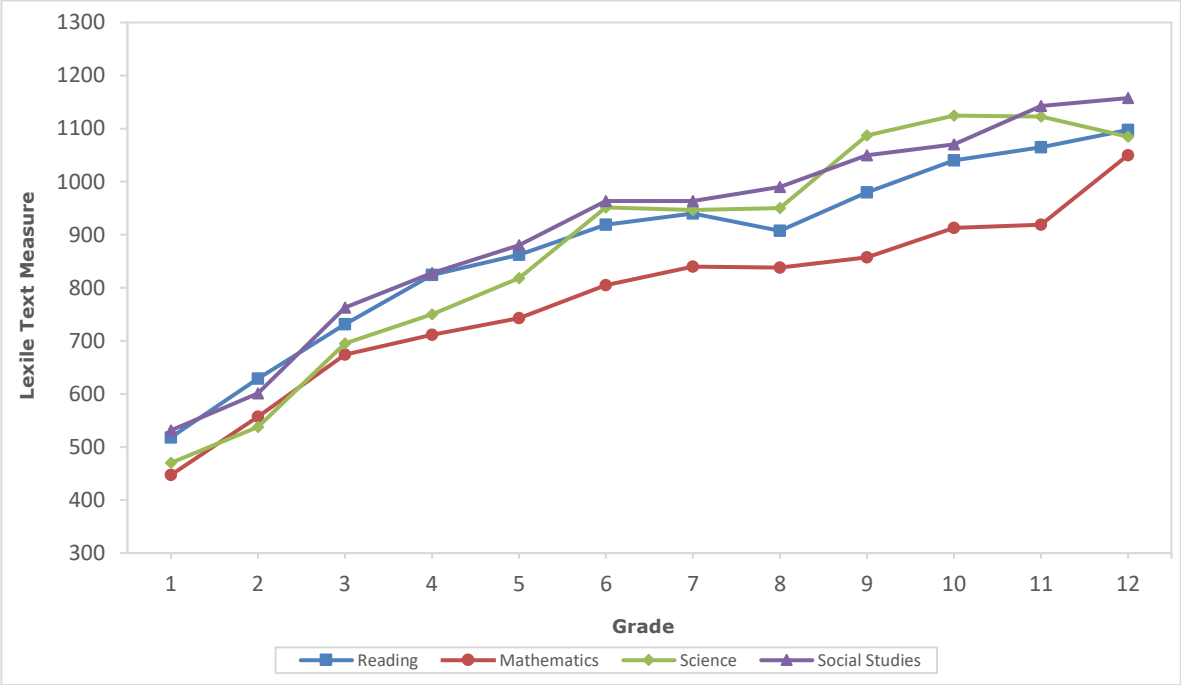
Research Question 2 investigated whether text complexity varies within grade level for the texts sampled. F*igure 1* illustrates the distribution of Lexile text measures (the *y*-axis) by grade (the *x*-axis). In the figure, the boxes refer to the interquartile range of Lexile text measures. The line within the box indicates the median and the diamond symbol identifies the mean. The end of each whisker represents the 5th percentile at the low end and the 95th percentile at the high end of the distribution of Lexile text measures for each grade. As seen in *Figure 1*, it is clear that text complexity varies within grade level. The range of text complexity (5th to 95th percentile) observed at each grade level (see *Table 2*) tends to increase from elementary school (~200L) to the middle school (~250L) to high school (~325L). Within-grade variability appears to be greatest in the high-school grades where the order of courses taken may not be as strictly aligned with specific grade levels. These findings are consistent with results from the 2012 text demand study conducted by Williamson et al. Next, less variability was observed for Grade 12 than the other high school grades; however, Grade 12 not only had the smallest sample of any grade level (*n* = 17), it also had the fewest mathematics textbooks (*n* = 1), which is important to note because mathematics textbooks consistently had the lowest text complexity across the content areas (see results and discussion related to Research Question 3 in the next section). Including more mathematics textbooks may have extended the Grade 12 distribution at the lower end to be more comparable to the other high school grade-level distributions.

Figure 1. Text complexity distributions (whiskers represent 5th and 95th percentiles), by grade.



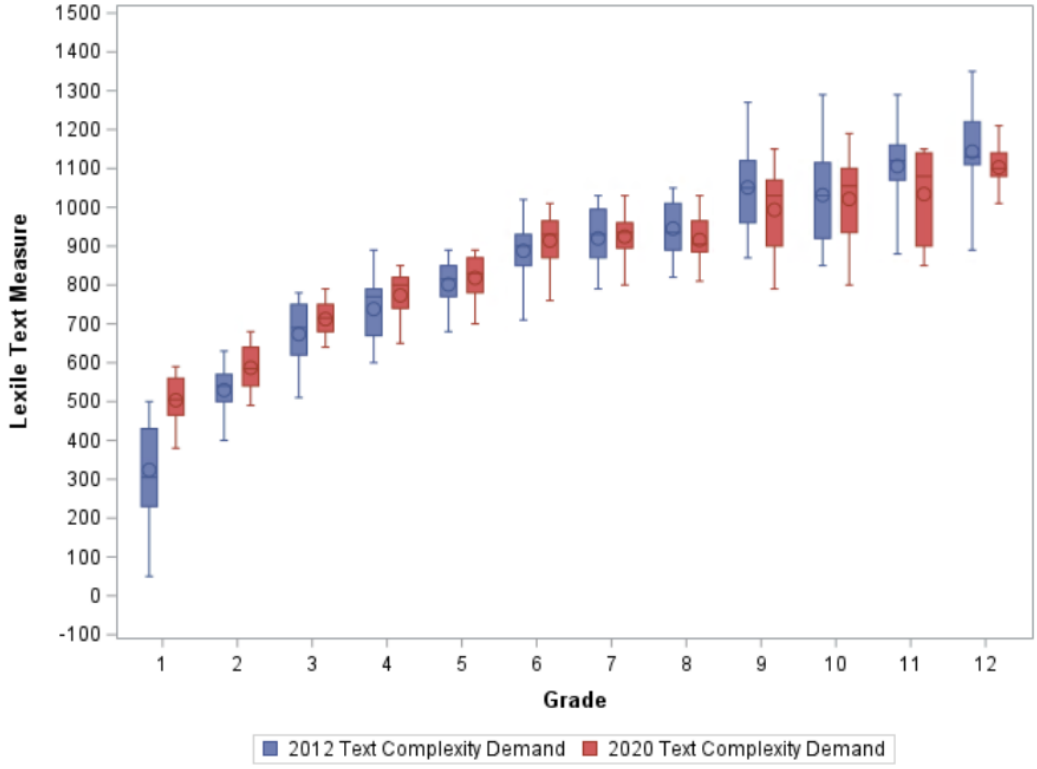
Research Question 3 examined whether text complexity varied across content area. *Figure 2* plots mean text complexity by grade for each content area. As seen in *Figure 2*, text complexity does vary across content area, with the largest differences mostly in the upper grades. Across the content areas, social studies texts generally had the highest reading demand at each grade level and mathematics the lowest. Overall, mean text complexity for each content area was within 100L for Grades 1 through 3; however, by Grade 4 mathematics begins to diverge from the other content areas and continues to be noticeably lower through Grade 11. To understand whether differences between mathematics and the other content areas were statistically significant, Tukey's HSD Test for multiple comparisons was performed. Findings indicated that the mean value of text complexity was significantly different ($p < .05$) between mathematics and reading (Grades 1, 3-7, 9-11), mathematics and science (Grades 5-11), and mathematics and social studies (Grades 1, 3-11). These findings are not surprising because in mathematics, where complex concepts need to be clear, concise, and consumable, a high level of text complexity might actually be detrimental to student understanding and learning.

Figure 2. Mean text complexity, by grade and content area.



Research Question 4 examined whether text complexity measures have changed since the 2012 text complexity study conducted by Williamson et al. The 2012 study included health and language arts textbooks; therefore, to establish comparable samples across the two studies, these textbooks were omitted. *Figure 3* illustrates the distribution of Lexile text measures (the *y*-axis) by grade (the *x*-axis) for both studies. Within *Figure 3* mean Lexile text measures across Grades 1 through 7 are higher for the current study than the 2012 study, but lower for Grades 8 through 12. To examine whether differences between mean Lexile text measures at each grade for the two studies were statistically significant, Welch's unequal variances *t*-tests were performed. Findings indicated that the mean values of text complexity were significantly different ($p < .05$) for Grades 1-3 as well as Grades 9 and 11. These results may suggest that text complexity has increased for Grades 1-3 and Grades 9 and 11 since the 2012 study. While statistically significant, the actual differences were small or modest for all the grades except Grade 1. These small or modest differences may be due to the continued expansion of the corpus of texts used to operationalize semantic demand variables and refinements in the operationalization of syntactic demand variables. However, it should also be noted that the Lexile Text Analyzer has evolved since the 2012 study, with further precision and capacity to analyze lower level texts. As a result, analyzing the same lower level text using the current Lexile text analyzer versus the legacy Lexile text analyzer may yield a higher Lexile text measure. As such, the statistically significant difference in 2012 and 2020 text complexity at Grade 1 (180L) may have more to do with the different analyzers used than an actual increase in text complexity. This was confirmed by examining the Lexile text measures derived for all Grades 1 textbooks in the sample using both the current and legacy Lexile text analyzers. Finally, while statistically significant differences were observed for Grades 9 and 11, within-grade variability in the current study was found to be the greatest in the high-school grades where the order of courses taken may not be as strictly aligned with specific grade levels. Therefore, the differences at Grades 9 and 11 may be in part due to the effects of sampling variability. Overall, findings for Research Question 4 suggest that text complexity has not significantly changed since the 2012 text complexity study conducted by Williamson et al.

Figure 3. 2012 and 2020 text complexity distributions (whiskers represent 5th and 95th percentiles), by grade.



Summary and Next Steps

In alignment with findings by Williamson et al. (2012), results from the current study indicate that median text complexity generally increases with grade and that text complexity varies within grade level. Results also showed that text complexity varied across content area. Additionally, when comparing results from the 2012 and 2020 text complexity studies, mean Lexile text measures across Grades 1 through 7 for 2020 were found to be higher than for 2012, but lower for Grades 8 through 12. However, findings showed that there was not a statistically significant difference in text complexity for most grades when comparing results from the two studies. These findings suggest that text complexity has not significantly changed since the 2012 text complexity study conducted by Williamson et al.

The ability to read is a critical indicator of academic success, and it is also fundamental to a student's pathway to broader literacy skills including writing, listening, and speaking (Snow, Burns, & Griffin, 1998). Furthermore, the ability to read complex text both independently and proficiently is integral for college and career readiness (NGA & CCSSO, 2010). When students read complex texts, the new language and knowledge gained supports their ability to read more advanced texts (Adams, 2011).

The limited increase in text complexity observed in this study for Grades 1-7 may not be enough to support students towards college and career reading readiness. The release of the Common Core State Standards (CCSS) in 2010 placed a strong emphasis on college and career readiness and called for increasing the complexity of texts students read in K-12 classrooms (NGA & CCSSO, 2010). In response to the release of the ELA and Mathematics CCSS and the subsequent rollouts of the Next Generation Science Standards in 2013 and the National Curriculum Standards for Social Studies in 2010, publishers developed new textbooks aligned to the standards for each content area. The limited increase in text complexity observed for Grades 1-7 may suggest that the new textbooks produced by publishers for these grade levels are still falling short in meeting expectations for increased text complexity.

This study provides educators a picture of the current text complexity demands associated with Grades 1-12. As findings from the current study suggest, while there may have been some small increases in reading text complexity in the lower grades, more needs to be done to increase the reading demands students typically encounter during Grades 1-12 in order to help them be prepared for college and career reading readiness. Targeted efforts to increase the complexity of texts students read across Grades 1-12 may help close the gap that exists between the reading demands of K-12 and higher education and other postsecondary endeavors.

Clearly, an understanding of text demands in the public schools is an evolving one. Schools have the option of supplementing regular textbooks, and both texts and auxiliary reading materials can change over time. Consequently, MetaMetrics continues its study of text complexity and updates its Titles Database as new titles are submitted for measurement.

REFERENCES

Adams, M.J. (2011). Advancing our students' language and literacy: The challenge of complex texts. *American Educator*, *34*(4), 3-11.

Fitzgerald, J., Elmore J., Hiebert, E.H., Koons, H., Bowen, K., Sanford-Moore, E.E., & Stenner A.J. (2016). Examining text complexity in the early grades. *Phi Delta Kappan*, *97*(8), 60-65.

Fitzgerald, J., Elmore, J., Koons, H., Hiebert, E. H., Bowen, K., Sanford-Moore, E. E., & Stenner, A. J. (2015). Important text characteristics for early-grades text complexity. *Journal of Educational Psychology*, *107(1)*, 4-29.

Mesmer, H. (2008). *Tools for matching readers to text: Research based practices*. Guilford Publications, Inc.

MetaMetrics, Inc. (2022). *The Lexile Framework for Reading development and validity evidence*. Durham, NC: MetaMetrics, Inc.

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common Core State Standards.* Washington DC: Author.

Snow, C., Burns, M., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Stenner, A. J., Burdick, H., Sanford, E. E., & Burdick, D. S. (2007). *The Lexile Framework for Reading technical report*. Durham, NC: MetaMetrics, Inc.

White, S. & Clement, J. (2001). Assessing the Lexile Framework: Results of a panel meeting. NCES Working Paper Series, Working Paper No. 2001-08. Washington, D.C.: U.S. Department of Education, Office of Educational Research and Improvement.

Williamson, G. L., Koons, H., Sandvik, T., & Sanford-Moore, E. (2012). *The text complexity continuum in grades 1-12* (MetaMetrics Research Brief). Durham, NC: MetaMetrics, Inc.

MetaMetrics® is focused on improving education for students of all ages. The organization develops scientific measures of academic achievement and complementary technologies that link assessment results with instruction. For more than twenty years, MetaMetrics' work has been increasingly recognized worldwide for its distinct value in differentiating instruction and personalizing learning. Its products and services for reading, mathematics and writing provide valuable insights about academic ability and the potential for growth, enabling students to achieve their goals at every stage of development.