Not So Common

Comparing Lexile® Measures With the Standards’ Other Text Complexity Tools

by Malbert Smith III, Ph.D., Metametrics President and Co-founder and Research Professor, The University of North Carolina at Chapel Hill
Executive Summary

Since the Common Core State Standards were published last year, much national attention has focused on the importance of text complexity in evaluating college- and career-readiness. Common Core authors David Coleman and Sue Pimentel have stated that understanding and measuring text complexity is a major shift in the new English Language Arts Standards and that these criteria are key to determining if students are adequately prepared for the academic and professional reading demands they will likely face after high school (The Hunt Institute, 2011). Subsequent reports, “Publishers’ Criteria for the Common Core State Standards in English Language Arts and Literacy” (for grades K-2 and Grades 3–12) and the more recent, “Measures of Text Difficulty: Testing Their Predictive Values for Grade Levels and Student Performance,” have echoed these same text complexity themes.

MetaMetrics® focuses on the importance of matching individual readers with targeted texts that provide the right level of challenge to support continued reading growth. Long before the Common Core movement, The Lexile® Framework for Reading played an important role in articulating the reading demands typically encountered in first grade through college and careers. In fact, MetaMetrics’ research on K-12 reading demands and ultimately those of the postsecondary world are annotated in the text complexity “staircase” in the Standards’ English Language Arts Appendix A. This staircase approach to text complexity is designed to help guide students’ reading comprehension development through their school years.

The subsequent reports noted earlier were intended to provide policy makers, educators and publishers with additional information and guidance on the value of measuring text complexity. While the “Publishers’ Criteria” reports primarily reiterate much of the research contained in the Common Core, the “Measures of Text Difficulty” report offers a detailed analysis of the text complexity landscape in terms of the tools commonly found in the marketplace. The report compares six text complexity tools—Carnegie Mellon University’s and the University of Pittsburgh’s REAder-specific Practice(REAP), Renaissance Learning’s ATOS, Questar Assessment’s Degrees of Reading Power® (DRP®), Pearson’s Reading Maturity Metric, ETS’s SourceRater and MetaMetrics’ Lexile® measure—using various criterion outcomes. A description of a seventh tool, Coh-Metrix’s Text Easability Assessor, is also mentioned. In summary, the report states that “there is no agreed upon gold standard” for evaluating text complexity. Its comparisons of the text complexity tools demonstrate that while they share some commonalities, there are also distinct differences (Nelson, Perfetti, Liben, & Liben, 2011). Building upon the report’s findings, this document provides a contextual framework for how these similarities and differences could be interpreted and used by the educational and publishing communities when selecting a text complexity tool.
As state departments of education, and the districts and schools within those respective states, transition from adopting the new Common Core State Standards to the more difficult task of implementing them, the challenge now becomes how to translate these higher standards into tangible, practical and cost-effective curricula. Implementing the Common Core will require districts and schools—and the educational publishers who supply curricular materials—to develop new instructional strategies and complementary resources that are not only aligned with the national college- and career-readiness standards, but also utilize and incorporate validated and budget-friendly tools that are universally accessible to all stakeholders.

The Relevance of Text Complexity

The Standards for English Language Arts focus on the importance of text complexity. As stated in Standard 10, students must be able to “read and comprehend complex literary and informational texts independently and proficiently” (Common Core State Standards for English Language Arts, College and Career Readiness Anchor Standards for Reading, NGA Center and CCSSO, 2010, p.10). The Common Core notes the following reasons for incorporating these more rigorous standards:

1. The text complexity of K-12 textbooks has become increasingly easier over the last 50 years (Chall, Conrad, & Harris, 1977; Hayes, Wolfer, & Wolfe, 1996).
2. The text demands of college and careers have remained consistent or increased over the same time period (Common Core State Standards for English Language Arts, Appendix A, NGA Center and CCSSO, 2010, p. 2).
3. As a result, there is a significant gap between students’ reading abilities and the text demands of their postsecondary pursuits. The Common Core states, “Being able to read complex text independently and proficiently is essential for high achievement in college and the work-place and important in numerous life tasks” (Common Core State Standards for English Language Arts, Appendix A, NGA Center and CCSSO, 2010, p. 4).

The Common Core recommends a three-part model for evaluating the complexity of a text that takes into account its qualitative dimensions, quantitative measure, and reader and task considerations. It describes text complexity as “the inherent difficulty of reading and comprehending a text combined with consideration of reader and task variables...a three-part assessment of text [complexity] that pairs qualitative and quantitative measures with reader-task considerations” (NGA Center and CCSSO, 2010, p. 43). In short, text complexity is a transaction between text, reader and task. The Lexile® Framework for Reading is based on this transaction and is significant because it allows for matching individual readers with specific texts on the same developmental scale.

MetaMetrics® research on the importance of the reader-with-text match and the typical reading demands of college and careers contributed to the Common Core as a whole and, more specifically, to the Lexile-based bands in Table 1.

Table 1. Text Complexity Grade Bands and Associated Lexile Ranges

<table>
<thead>
<tr>
<th>Text Complexity Grade Bands</th>
<th>Lexile Ranges Aligned to College- and Career-Readiness Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td>K–1</td>
<td>N/A</td>
</tr>
<tr>
<td>2–3</td>
<td>450L–790L</td>
</tr>
<tr>
<td>4–5</td>
<td>770L–980L</td>
</tr>
<tr>
<td>6–8</td>
<td>955L–1155L</td>
</tr>
<tr>
<td>9–10</td>
<td>1080L–1305L</td>
</tr>
<tr>
<td>11–CCR</td>
<td>1215L–1355L</td>
</tr>
</tbody>
</table>

Note: There is an infinity of ways of connecting a single starting point in grade 1 to a single ending point in grade 12. Each of these ways may find a supportive constituency and a constituency in opposition.

Text Complexity and The Lexile Framework for Reading

Since its inception in 1984, MetaMetrics has focused on the development and refinement of a scientific scale that measures text complexity and, more importantly, places individuals on that same scale to evaluate reading ability. Today, The Lexile Framework for Reading is used by nearly half of U.S. state departments of education and tens of millions of individuals worldwide (primarily because ETS’s TOEFL iBT® test, TOEFL® Junior™ test and TOEIC® test report a Lexile measure). Consistent with the Common Core’s definition of text complexity as the transaction between reader, text and task, the originating psychometric method used to develop the Lexile Framework was based on the relationship between individuals’ actual reading comprehension (for a given task) while reading specific texts. The Lexile Framework is not a readability formula. Readability formulas are based only on the features and/or characteristics of a text. As stated in the Common Core, “…[A]n important difference between the Lexile system and traditional readability formulas is that traditional formulas only assign a score
to texts, whereas the Lexile Framework can place both readers and texts on the same scale” (NGA Center and CCSSO, 2010, p. 7).

MetaMetrics’ vision of a common scale to measure both text complexity and reading ability was shared by policy makers and researchers at the National Institutes of Health, which supported the organization’s work through five grants over a ten-year period. As the education community transitions to the Common Core, stakeholders will need to be cognizant of both text complexity and reading ability, and how they can use this information to accurately match individual readers with targeted texts. Accelerating reader growth and scaffolding support to the more rigorous college- and career-readiness levels can be accomplished most effectively when a common scale connects readers with appropriate texts.

Comparing the Standards’ Text Complexity Tools

In addition to the Common Core, two more supporting documents were published last year on the importance of text complexity. The “Publishers’ Criteria for the Common Core State Standards in English Language Arts and Literacy, GradesK-2” and “Publishers’ Criteria for the Common Core State Standards in English Language Arts and Literacy, Grades 3-12” reinforce the value of measuring text complexity for educational publishers. A third report, “Measures of Text Difficulty: Testing their Predictive Value for Grade Levels and Student Performance,” released in December 2011 describes the various text complexity tools available to the education community: Carnegie Mellon University’s and University of Pittsburgh’s REAder-specific Practice (REAP), Renaissance Learning’s ATOS, Questar Assessment’s Degrees of Reading Power (DRP), the Pearson Reading Maturity Metric, ETS’s SourceRater, MetaMetrics’ Lexile measure, and Coh-Metrix’s Text Easability Assessor.

Noting that “there is no clear gold standard measure for text difficulty against which to compare the various metrics,” the report authors evaluated each of the tools (except Coh-Metrix’s Text Easability Assessor) using three basic types of criterion measures (Nelson, Perfetti, Liben, & Liben, 2011). Table 2 provides an analysis of the report’s findings on one of the criterion measures. The “Reference Set” column includes the five reference sets that were used as criterion measures. Each of the six text complexity tools in the top-row headings was correlated (using Pearson Product Moment Correlations) with predetermined text complexity levels in the reference sets. None of the correlations were corrected for the measurement error or any other artifactual sources of variation.

While there are five different criterion reference sets, the sets fall into three categories:

1. CCSS (Common Core State Standards) exemplars comprised of passages in which the text complexity has been determined based on human judgment.
2. State test passages, SAT-9, and GMRT are based on empirical student test performance within a grade level.
3. EdSphere™ data is based entirely on empirical data of student responses to text passages.

The first two categories of criterion measures suffer from serious limitations which compromise their utility. The first category is comprised of a disproportionate number of passages that had a significantly different measure than the text complexity measure for the complete book. This bias calls into question its validity and utility. The second set of criterion measures while at least based on empirical student performance is compromised due to the truncation of scores into grade levels. The categorization of the criterion into grade-level-groupings restricts the correlation coefficients of text complexity metric that is a vertical scale. Even with these limitations all the tools were moderately correlated with the criterion measures.

The third criterion measure is based entirely on empirical student performance without the restriction of collapsing into grade levels. For the EdSphere criterion, all of the text complexity tools were moderately correlated to highly correlated.
It is important to note that the correlations for the text complexity tools (except for ETS's SourceRater which was not part of this reference set) were higher on this criterion than for the other four references sets. While all of the correlations were higher, the Lexile and ATOS scales had the highest correlations. MetaMetrics believes that this criterion measure is the ultimate litmus test for text complexity tools: How well does the tool estimate text complexity from actual student comprehension? A second and related point is the importance of placing students in terms of their reading ability on the same scale.

**Selecting and Using a Text Complexity Tool**

Besides the correlation coefficients, there are many other features and dimensions that should be considered when selecting a text complexity tool. MetaMetrics believes that the most important feature is the ability to place both a reader and a text on the same developmental scale, which is the essence of the Lexile Framework. Simply ordering text from easy to difficult on a scale is of limited value without measuring and ordering an individual's reading ability on that same scale.

Table 3. Text Complexity Tools

<table>
<thead>
<tr>
<th></th>
<th>REAP</th>
<th>ATOS</th>
<th>DRP</th>
<th>Lexile Measure</th>
<th>Reading Maturity</th>
<th>SourceRater</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scale Type</strong></td>
<td>Grade Level</td>
<td>Grade Level</td>
<td>Vertical</td>
<td>Vertical***</td>
<td>Grade Level</td>
<td>Grade Level</td>
</tr>
<tr>
<td><strong>Based on Student-Test-Track transaction</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Free Text Analyzer Available</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Spanish Scale and Online Analyzer</strong></td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>


Table 4. Number of TextsMeasured to Date

<table>
<thead>
<tr>
<th></th>
<th>REAP</th>
<th>ATOS</th>
<th>DRP</th>
<th>Lexile Measure</th>
<th>Reading Maturity</th>
<th>SourceRater</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Books Measured</strong></td>
<td>--</td>
<td>124,861</td>
<td>38,000</td>
<td>168,000*</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>Textbooks Measured</strong></td>
<td>--</td>
<td>--</td>
<td>7,000</td>
<td>5,000+</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>Articles Measured</strong></td>
<td>--</td>
<td>0</td>
<td>--</td>
<td>100,000,000**</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>Webpages Measured</strong></td>
<td>--</td>
<td>0</td>
<td>0</td>
<td>60,000**</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

*Updated August 2013  **As of November 2011, this information was not publically available.

**https://d1jt5u2s0h3gkt.cloudfront.net/m/uploads/positionpapers/TheReading-WritingConnection.pdf

Table 5. Number of Assessments Reporting the Measure

<table>
<thead>
<tr>
<th></th>
<th>REAP</th>
<th>ATOS</th>
<th>DRP</th>
<th>Lexile Measure</th>
<th>Reading Maturity</th>
<th>SourceRater</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>State Assessments</strong></td>
<td>--</td>
<td>0</td>
<td>1*</td>
<td>21**</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Norm-Referenced Assessments</strong></td>
<td>--</td>
<td>0</td>
<td>--</td>
<td>10</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>Benchmark Assessments</strong></td>
<td>--</td>
<td>0</td>
<td>--</td>
<td>9</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>Student Measures Each Year</strong></td>
<td>--</td>
<td>30,000+***</td>
<td>--</td>
<td>35,000,000</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>


*As of November 2011, this information was not publically available.
Tables 3, 4 and 5 on the previous page highlight some of the key features and dimensions of the Standards’ text complexity tools. Table 3 focuses on the foundational characteristics; Table 4 focuses on the type and quantity of text that has been measured by each tool; and Table 5 focuses on the measurement of students. As these tables illustrate, there are a number of distinct advantages for using Lexile measures.

First, the Lexile Framework is the only text complexity measurement system that is originally psychometrically derived from individuals reading texts—the reader-text-task transaction. MetaMetrics theorized this transaction and discovered that select semantic and syntactic text features could serve as proxies for the mental processes individuals use when reading texts. These text proxies account for an extremely high amount of variance (94 percent) in readers' comprehension.

MetaMetrics has conducted studies to explore additional proxies that could explain the remaining six percent of variance (Hanlon, Swartz, Stenner, Burdick, & Burdick, 2010; Kershaw, Barth, Francis, Swartz, & Stenner, 2011; Stenner, Burdick, Burdick, Hanlon, & Swartz, 2010; Stenner, Swartz, Burdick, Burdick, & Hanlon, 2010; Swartz, Burdick, Hanlon, Stenner, Kyngdon, Burdick, & Smith, 2011). The organization has also provided its data sets to third-party researchers and organizations to see if they could improve the Lexile measure. While this research continues, it is important to note that, to date, none of the studies has led to the discovery of additional proxies (in a statistically significant way). While additional proxies could be added, MetaMetrics has found that they do not provide any more explanatory power to the Lexile equation. And as any researcher will attest, simply adding more variables to an equation does not necessarily produce a more accurate result.

Second, the Lexile Framework was built using a vertical, developmental and equal-interval scale, permitting finer-grained analyses than grade-related scales and avoiding the pitfalls associated with grade-related scales. More than three decades ago, the International Reading Association resolved that grade-equivalent reports should be abandoned because they were being misconstrued and misused (IRA, 1981).

MetaMetrics’ and Questar’s tools are built upon vertical, equal-interval scales. Renaissance Learning, ETS and Pearson Educational Measurement have opted to create grade-based or grade-equivalent scales. On the surface, grade equivalents appear to be an intuitive way to report students’ test scores. However, this seemingly simplistic method glosses over some significant limitations that often promote misleading and inaccurate interpretations of the data. These misconceptions include: 1) norms are standards of what should be; 2) grade equivalents indicate the appropriate grade placement for a student; 3) all students should be expected to grow one grade-equivalent unit per year; 4) the units are equal throughout the score range; 5) grade equivalents for different tests are comparable; and 6) the scores that are based on extrapolations to grades well above or below the test level are meaningful (Miller, Linn & Gronlund, 2009).

Third, the Lexile Framework places reader and text on the same developmental (Lexile) scale, which enables for more accurate reader-with-text matches. As numerous researchers and educational practitioners attest, this feature provides classroom teachers with an actionable resource to support differentiated instruction (Chall & Dale, 1995; Hall & Moats, 1999; Hiebert, 2009; Hiebert & Mesmer, 2011; Mesmer, 2007).

Fourth, a topic that is often overlooked but has profound implications is the importance of “editing” a text for analysis, as different tools have different rules for how text should be prepared for measurement. MetaMetrics makes its rules transparent by publishing them on the Lexile website. Without precise rules for how text should be prepared for analysis there will be volatility in the measures. MetaMetrics also provides open access to its Lexile® Analyzer to generate Lexile measures for various length text files, from short samples to entire books. More than 228,000 individuals have registered for this free resource.

Fifth, pragmatic factors should be considered when selecting a text complexity tool. Factors such as ubiquity of use, access to complementary resources, and the number of books, assessments and products that are built on these text complexity tools are paramount. Lexile measures, for example, are available for more than 150,000 books from more than 200 publishers in MetaMetrics’ free “Find a Book” search, and in the search tools from domestic and international booksellers like Barnes & Noble and Korea’s Interpark Books. In addition, more than 5,000 textbooks from more than 20 publishers have been measured. More and more, publishers are submitting their materials for Lexile measurement to ensure alignment with the Common Core. Tens of millions of articles also have Lexile measures and are available from content aggregators such as EBSCO, ProQuest, Gale Cengage and Newsbank. And more than 200,000 web pages have been measured. All of these resources with Lexile measures—many of which are used in schools and libraries—allow students to read targeted narrative and informational text on classroom
relevant topics or personal interests.

A final consideration involves the assessment of the reader. As noted previously, it is imperative that readers and texts are placed on a common scale. Currently, there are only three systems (ATOS, DRP and the Lexile Framework) which measure the reader and the text. However, the availability of Renaissance Learning's ATOS and Questar Assessment's DRP metrics are limited to those organizations' proprietary tests. Schools and districts must purchase the STAR test from Renaissance Learning or the DRP test from Questar. MetaMetrics, in contrast, does not publish or sell its own test. The organization maintains open, nonproprietary relationships with major educational test publishers, having linked the Lexile Framework with more than 60 popular reading assessments and programs. ETS and Pearson alone offer Lexile measures from a total of eleven reading assessments. This approach provides districts and schools with more options and greater flexibility in determining which instrument to use. (Please see Appendix A for a complete list of assessments and reading programs that report Lexile measures.)

**Conclusion**

Only three of the text complexity tools included in the “Measures of Text Difficulty” report have a direct and empirical link to readers and texts: ATOS, DRP and the Lexile measure. The other three tools, SourceRater, REAP and the Reading Maturity measure, evaluate text only and, to date, are not reported on any standardized tests.

When it comes to selecting a tool to measure reading ability and text complexity, there are two important criteria to consider. The first is ubiquity, or how many resources like books, articles and websites have been measured. As mentioned earlier, Lexile measures are available for more than 150,000 books, tens of millions of articles, and hundreds of thousands of web pages.

The second criteria is the “accessibility” of the tools for measuring text complexity. The success of the Common Core will be determined by how the Standards are translated into classroom practice. MetaMetrics strives to provide the tools and complementary resources to make this a reality. For example, the Lexile Analyzer is freely available on the Lexile website for educators and others to determine the Lexile measure of a text. In addition, publishers can license the software to build commercial products or to report certified Lexile measures in public domains like websites, databases and catalogs. To date, MetaMetrics has provided more than 228,000 users with access to the free Lexile Analyzer. Over the past two years alone, users have analyzed nearly one million text files. And the number of words analyzed for commercial use exceeds 1.5 billion.

As Stephen Covey points out in his book, “The Seven Habits of Highly Successful People,” it is best to begin with the end in mind. The authors of the “Measures of Text Difficulty” report offer a text complexity staircase for college- and career-readiness in Appendix C. By mapping required Lexile growth backwards across grade levels and providing access to resources to measure text complexity and monitor student growth, policy makers and educators will be better able to implement and sustain the Common Core as they prepare all students for the reading demands of the postsecondary world.
References


## Appendix: Assessments and Reading Programs that Report Lexile Measure

### State Assessments
- Arizona's Instrument to Measure Standards
- California English-Language Arts Standards Test
- Delaware Comprehensive Assessment System*
- Georgia Criterion-Referenced Competency Tests and End-of-Course Tests (Ninth Grade Literature and Composition, and American Literature and Composition)
- Hawaii State Assessment
- Illinois Standards Achievement Test
- Kansas State Assessments of Reading
- Kentucky Performance Rating for Educational Progress
- Minnesota Comprehensive Assessments
- New Mexico Standards-Based Assessment
- North Carolina End-of-Grade Tests and English I End-of-Course Test
- Oklahoma Core Curriculum Tests
- Oregon Assessment of Knowledge and Skills
- Proficiency Assessments for Wyoming Students
- South Carolina Palmetto Assessment of State Standards
- South Dakota State Test of Educational Progress
- Tennessee Comprehensive Assessment Program Achievement Test
- Texas Assessment of Knowledge and Skills**
- Virginia Standards of Learning Tests
- West Virginia WESTEST 2

### Norm-Referenced Assessments
- CTB/McGraw-Hill: TerraNova (CAT/6 and CTBS/5) and Tests of Adult Basic Education (TABE)
- Pearson: Stanford 9 & 10, MAT 8 and Aprenda 3
- Riverside Publishing: The Iowa Tests (ITBS and ITED) and Gates-MacGinitie Reading Tests, Fourth Edition

### Interim/Benchmark Assessments
- American Education Corporation: A+ Learning Link
- Curriculum Associates: i-Ready* Diagnostic & Instruction
- Dynamic Measurement Group: DIBELS
- Measured Progress: Progress Toward Standards (PTS3)
- Measurement, Inc.: Item Bank
- NWEA: Measures of Academic Progress* (MAP*)
- Scantron: Performance Series
- Scholastic: Scholastic Reading Inventory

### Reading Programs and Interventions
- Achieve3000: KidBiz3000 and TeenBiz3000
- Cambium Learning Group (Sopris/Voyager): Language! and Passport Reading Journeys
- Capstone Digital: myON reader
- EDmin: Total Reader
- EPS: MCI (Making Connections® Intervention)
- Hampton-Brown: Edge and Inside
- Houghton Mifflin: Earobics Reach
- McGraw-Hill: SRA FLEX Literacy™
- Pearson/Longman/Prentice Hall: MyReadingLab
- Scholastic: Read 180 and Reading Counts!

### International Assessments
- ETS: TOEFL iBT® test, TOEFL® Junior™ test and TOEIC® test
- GL Assessment: Progress in English (PIE) Assessment
- Natmal: E-LQ Assessment

### Spanish Assessments
- Achieve3000: KidBiz3000 and TeenBiz3000
- New Mexico Standards-Based Assessment
- Pearson: Aprenda 3
- Scholastic: Scholastic Reading Inventory

*Linking study in 2012
**Grades 10 and 11 only
MetaMetrics is focused on improving education for learners of all ages and ability levels. The organization develops scientific measures of academic achievement and complementary technologies that link assessment results with real-world instruction. MetaMetrics’ products and services for reading (The Lexile® Framework for Reading, El Sistema Lexile para Leer), mathematics (The Quantile® Framework for Mathematics), and writing (The Lexile Framework for Writing) provide unique insights about academic ability and the potential for growth, enabling individuals to achieve their goals at every stage of development. For more information, visit www.MetaMetricsInc.com.

METAMETRICS®, the METAMETRICS® logo and tagline, LEXILE®, LEXILE FRAMEWORK®, LEXILE ANALYZER® and the LEXILE® logo are trademarks of MetaMetrics, Inc., and are registered in the United States and abroad. The trademarks and names of other companies and products mentioned herein are the property of their respective owners. Copyright © MetaMetrics, Inc. All rights reserved.