



Linking the NC READY EOG Reading/EOC English II with the Lexile® Framework

*A Study to Link the North Carolina READY EOG
Reading/EOC English II with The Lexile®
Framework for Reading*

November 2013
Updated April 2015

Prepared by MetaMetrics for:

**North Carolina Department of Public Instruction
Division of Accountability Services
301 N. Wilmington Street
Raleigh, NC 27601**



MetaMetrics.

1000 Park Forty Plaza Drive, Suite 120
Durham, North Carolina 27713
www.MetaMetricsInc.com
www.Lexile.com

Preface

Lexile Scale Enhancements

The Lexile® Framework for Reading is a scientific approach to measuring reading ability and the complexity of reading materials. The Lexile Framework includes a Lexile measure and the Lexile scale. A Lexile measure represents both the complexity of a text, such as a book or article, and an individual’s reading ability. Lexile measures are expressed as numeric measures followed by an “L” (e.g., 850L), and are placed on the Lexile scale. (There is no space between the measure and the “L.”) The Lexile scale is a developmental scale for reporting reader ability and text complexity, ranging from below 200L for emergent readers and emergent-reader texts to above 1600L for advanced readers and texts. Lexile measures of one thousand or greater are reported without a comma (e.g., 1050L). All Lexile reader measures should be rounded to the nearest 5L to avoid over-interpretation of the measures. As with any test score, uncertainty in the form of measurement error is present. If the Lexile reader measure is xxx2.5 or higher or xxx7.5 or higher, it is rounded up to the next highest 5L; below those points, the measure is rounded down to the next lowest 5L. For example, if a computed Lexile reader measure is 772.51, it should be reported as 775L. If the computed Lexile reader measure is 777.42, it should be reported as 775L.

Prior to May 1, 2014, all Lexile reader measures at or below 0L were reported as BR (Beginning Reader). Starting in spring 2014, Lexile reader measures below 0L may be reported with a more specific measure. These BR measures are shown as “BRxxxL.” For example, a Lexile reader measure of -150 is reported as BR150L where “BR” stands for “Beginning Reader” and replaces the negative sign in the number. The Lexile scale is like a thermometer, with numbers below zero indicating decreasing reading ability as the number moves away from zero. The smaller the number following the BR code, the more advanced the reader is. For example, a BR150L reader is more advanced than a BR200L reader. Above 0L, measures indicate increasing reading ability as the numbers increase. For example, a 200L reader is more advanced than a 150L reader.

Lexile measures that are reported for an individual student should reflect the purpose for which they will be used. If the purpose is research (e.g., to measure growth at the student, grade, school, district, or state level), then actual measures should be used at all score points, rounded to the nearest integer. A computed Lexile measure of 772.51 would be represented as 773L. If the purpose is instructional, then the Lexile measures should be capped at the upper bound of measurement error (e.g., at the 95th percentile point of the national Lexile norms) to ensure developmental appropriateness of the material. MetaMetrics expresses these measures used for instructional purposes as “Reported Lexile Measures” and recommends that they be used on individual score reports. In instructional environments where the purpose of the Lexile measure is to

appropriately match readers with text, all scores below 0L should be reported as “BRxxxL.” No student should receive a negative Lexile measure on a score report. The lowest reported value below 0L is BR400L.

Table i. Maximum reported Lexile measures by grade.

Grade	Lexile Caps
K	850L
1	900L
2	1100L
3	1200L
4	1300L
5	1400L
6	1500L
7	1600L
8	1700L
9	1725L
10	1750L
11	1800L
12	1825L

Some assessments report a Lexile range for each student rather than a specific Lexile reader measure. The Lexile range is 50L above to 100L below the student’s actual Lexile measure. For example, the Lexile range for a specific reader measure of 700L is 600L to 750L. This range represents the boundaries between relatively easy reading material for the student and the level at which the student will be more challenged, yet can still read successfully.

Text within the Technical Report has been updated to correspond with the language of the enhanced Lexile scale.

Table of Contents

Introduction	1
The Lexile Framework for Reading	3
The Semantic Component	3
The Syntactic Component	4
Calibration of Text Difficulty	5
The Lexile Scale.....	5
Validity of The Lexile Framework for Reading	7
Text Measure Error Associated with the Lexile Framework.....	11
Lexile Item Bank	13
The NC READY EOG Reading/EOC English II – Lexile Framework Linking Process	17
Description of the Assessments.....	17
Study Design.....	20
Description of the Sample	21
Linking the NC READY EOG Reading/EOC English II Scale Scores with the Lexile Scale	32
Validity of the NC READY EOG Reading/EOC English II – Lexile Link.....	34
The Lexile Framework and Forecasted Comprehension Rates	43
Conclusions, Caveats, and Recommendations	47
References.....	59

Introduction

Often it is desirable to convey more information about test performance than can be incorporated into a single primary score scale. Two examples arise in large-scale assessment. In one situation, one test can provide a unique type of information (such as national comparisons available from NAEP) but is not administered very often. At the same time another test is administered more often, but is not able to provide the breadth of information (such as a state assessment). An auxiliary score scale for a test can be established to provide this additional information through assessment scale linkages. Once linkages are established between the two assessments, then the results of the more-frequently-administered assessment can be translated in terms of the scale for the other assessment.

In another situation, the linkage between two score scales can be used to provide a context for understanding the results of one of the assessments. For example, sometimes it is hard to explain what a student can read based on the results of a reading comprehension test. Parents typically ask the questions “If my child is in the fourth grade and scores 450 on the NC READY EOG Reading assessment, what does this mean?” or “Based on my child’s test results, what can he or she read and how well?” or “Is my child well prepared to meet the reading demands of grade level materials?” Once a linkage is established with an assessment that is related to specific book or text titles, then the results of the assessment can be explained and interpreted in the context of the specific titles that a student should be able to read.

Auxiliary score scales can be used to “convey additional normative information, test-content information, and information that is jointly normative and content based. For many test uses, an auxiliary scale conveys information that is more crucial than the information conveyed by the primary score scale. In such instances, the auxiliary score is the one that is focused on, and the primary scale can be viewed more as a vehicle for maintaining interpretability over time” (Petersen, Kolen, and Hoover, 1989, p. 222). One such auxiliary scale is The Lexile[®] Framework for Reading, which was developed to appropriately match readers with text at a level that provides challenge but not frustration.

Linking assessment results with the Lexile Framework provides a mechanism for matching each student’s reading ability with text on a common scale. It serves as an anchor to which texts and assessments can be connected allowing parents, teachers, and administrators to speak the same language. In addition, the Lexile Framework provides a common way to monitor if students are “on track” for the reading demands of various postsecondary endeavors. By using the Lexile Framework, the same metric is applied to

the books students read, the tests they take, and the results that are reported. Parents often ask questions like the following:

- How can I help my child become a better reader?
- How do I challenge my child to read so that she is ready for various college and career options?

Questions like these can be challenging for parents and educators. By linking the NC READY EOG Reading/EOC English II assessment with The Lexile Framework for Reading, educators and parents will be able to answer these questions and will be better able to use the results from the test to improve instruction and to develop each student's level of reading comprehension.

This research study was designed to determine a mechanism to provide reading levels that can be matched to text based on the NC READY EOG Reading/EOC English II test scores. The study was conducted by MetaMetrics, Inc. in collaboration with the North Carolina Department of Public Instruction (NCDPI) (Contract No. NC10025818 dated December 17, 2012). The primary purposes of this study were to:

- present a solution for matching readers with text;
- provide North Carolina with Lexile measures on the NC READY EOG Reading/EOC English II assessment;
- develop tables for converting NC READY EOG Reading/EOC English II scale scores to Lexile measures; and
- produce a report that describes the linking analysis procedures.

The Lexile Framework for Reading

All symbol systems share two features: a semantic component and a syntactic component. In language, the semantic units are words. Words are organized according to rules of syntax into thought units and sentences (Carver, 1974). In all cases, the semantic units vary in familiarity and the syntactic structures vary in complexity. The comprehensibility or difficulty of a message is dominated by the familiarity of the semantic units and by the complexity of the syntactic structures used in constructing the message.

The Semantic Component

As far as the semantic component is concerned, it is clear that most operationalizations are proxies for the probability that an individual will encounter a word in a familiar context and thus be able to infer its meaning (Bormuth, 1966). This is the basis of exposure theory, which explains the way receptive or hearing vocabulary develops (Miller and Gildea, 1987; Stenner, Smith, and Burdick, 1983). Klare (1963) hypothesized that the semantic component varied along a familiarity-to-rarity continuum. This concept was further developed by Carroll, Davies, and Richman (1971), whose word-frequency study examined the reoccurrence of words in a five-million-word corpus of running text. Knowing the frequency of words as they are used in written and oral communication provided the best means of inferring the likelihood that a word would be encountered by a reader and thus become a part of that individual's receptive vocabulary.

Variables such as the average number of letters or syllables per word have been observed to be proxies for word frequency. There is a strong negative correlation between the length of words and the frequency of word usage. Polysyllabic words are used less frequently than monosyllabic words, making word length a good proxy for the likelihood that an individual will be exposed to a word.

In a study examining receptive vocabulary, Stenner, Smith, and Burdick (1983) analyzed more than 50 semantic variables in order to identify those elements that contributed to the difficulty of the 350 vocabulary items on Forms L and M of the *Peabody Picture Vocabulary Test – Revised* (Dunn and Dunn, 1981). Variables included part of speech, number of letters, number of syllables, the modal grade at which the word appeared in school materials, content classification of the word, the frequency of the word from two different word counts, and various algebraic transformations of these measures.

The first word frequency measure used was the raw count of how often a given word appeared in a corpus of 5,088,721 words sampled from a broad range of school materials (Carroll, Davies, and Richman, 1971). For example, the word “accident”

appears 176 times in the 5,088,721-word corpus. The second word frequency measure used was the frequency of the “word family.” A word family included: (1) the stimulus word; (2) all plurals (adding “-s” or “-es” or changing “-y” to “-ies”); (3) adverbial forms; (4) comparatives and superlatives; (5) verb forms (“-s,” “-d,” “-ed,” and “-ing”); (6) past participles; and (7) adjective forms. For example, the word family for “accident” would include “accidental,” “accidentally,” “accidentals,” and “accidents,” and they would all have the same word frequency of 334. The frequency of a word family was based on the sum of the individual word frequencies from each of the types listed.

Correlations were computed between algebraic transformations of these means (mean frequency of the words in the test item and mean frequency of the word families in the test item) and the rank order of the test items. Since the items were ordered according to increasing difficulty, the rank order was used as the observed item difficulty. The log of the mean word frequency provided the strongest correlation with item rank order ($r = -0.779$) for the items on the combined form.

The Lexile Framework currently employs a 600-million-word corpus when examining the semantic component of text. This corpus was assembled from the more than 15,000 texts that were measured by MetaMetrics for publishers from 1998 through 2002. When text is analyzed by MetaMetrics, all electronic files are initially edited according to established guidelines used with the Lexile Analyzer software. These guidelines include the removal of all incomplete sentences, chapter titles, and paragraph headings; running of a spell check; and re-punctuating where necessary to correspond to how the book would be read by a child (for example, at the end of a page). The text is then submitted to the Lexile Analyzer that examines the lengths of the sentences and the frequencies of the words and reports a Lexile measure for the book. When enough additional texts have been analyzed to make an adjustment to the corpus necessary and desirable, a linking study will be conducted to adjust the calibration equation such that the Lexile measure of a text based on the current corpus will be equivalent to the Lexile measure based on the new corpus.

The Syntactic Component

Klare (1963) provides a possible interpretation for how sentence length works in predicting passage difficulty. He speculated that the syntactic component varied with the load placed on short-term memory. Crain and Shankweiler (1988), Shankweiler and Crain (1986), and Liberman, Mann, Shankweiler, and Westelman (1982) have also supported this explanation. The work of these individuals has provided evidence that sentence length is a good proxy for the demand that structural complexity places upon verbal short-term memory.

While sentence length has been shown to be a powerful proxy for the syntactic complexity of a passage, an important caveat is that sentence length is not the underlying causal influence (Chall, 1988). Researchers sometimes incorrectly assume that manipulation of sentence length will have a predictable effect on passage difficulty. Davidson and Kantor (1982), for example, illustrated rather clearly that sentence length can be reduced and difficulty increased and vice versa.

Based on previous research, it was decided to use sentence length as a proxy for the syntactic component of reading difficulty in the Lexile Framework.

Calibration of Text Difficulty

The research study on semantic units (Stenner, Smith, and Burdick, 1983) was extended to examine the relationship of word frequency and sentence length to reading comprehension. In 1987(a), Stenner, Smith, Horabin, and Smith performed exploratory regression analyses to test the explanatory power of these variables. This analysis involved calculating the mean word frequency and the log of the mean sentence length for each of the 66 reading comprehension passages on the *Peabody Individual Achievement Test* (Dunn and Markwardt, 1970). The observed difficulty of each passage was the mean difficulty of the items associated with the passage (provided by the publisher) converted to the logit scale. A regression analysis based on the word-frequency and sentence-length measures produced a regression equation that explained most of the variance found in the set of reading comprehension tasks. The resulting correlation between the observed logit difficulties and the theoretical calibrations was 0.97 after correction for range restriction and measurement error. The regression equation was further refined based on its use in predicting the observed difficulty of the reading comprehension passages on 8 other standardized tests. The resulting correlation between the observed logit difficulties and the theoretical calibrations across the 9 tests was 0.93 after correction for range restriction and measurement error.

Once a regression equation is established linking the syntactic and semantic features of text to the difficulty of text, the equation can be used to calibrate test items and text.

The Lexile Scale

In developing the Lexile Scale, the Rasch model (Wright and Stone, 1979) was used to estimate the difficulties of the items and the abilities of the persons on the logit scale.

The calibrations of the items from the Rasch model are objective in the sense that the relative difficulties of the items will remain the same across different samples of persons (specific objectivity). When two items are administered to the same group it can be

determined which item is harder and which one is easier. This ordering should hold when the same two items are administered to a second group. If two different items are administered to the second group, there is no way to know which set of items is harder and which set is easier. The problem is that the location of the scale is not known. General objectivity requires that scores obtained from different test administrations be tied to a common zero – absolute location must be sample independent (Stenner, 1990). To achieve general objectivity, the theoretical logit difficulties must be transformed to a scale where the ambiguity regarding the location of zero is resolved.

The first step in developing a scale with a fixed zero was to identify two anchor points for the scale. The following criteria were used to select the two anchor points: they should be intuitive, easily reproduced, and widely recognized. For example, with most thermometers the anchor points are the freezing and boiling points of water. For the Lexile Scale, the anchor points are text from seven basal primers for the low end and text from *The Electronic Encyclopedia* (Grolier, Inc., 1986) for the high end. These points correspond to the middle of first grade text and the midpoint of workplace text.

The next step was to determine the unit size for the scale. For the Celsius thermometer, the unit size (a degree) is $1/100^{\text{th}}$ of the difference between freezing (0 degrees) and boiling (100 degrees) water. For the Lexile Scale the unit size (a Lexile) was defined as $1/1000^{\text{th}}$ of the difference between the mean difficulty of the primer material and the mean difficulty of the encyclopedia samples. Therefore, a Lexile by definition equals $1/1000^{\text{th}}$ of the difference between the difficulty of the primers and the difficulty of the encyclopedia.

The third step was to assign a value to the lower anchor point. The low-end anchor on the Lexile Scale was assigned a value of 200.

Finally, a linear equation of the form

$$[(\text{Logit} + \text{constant}) \times \text{CF}] + 200 = \text{Lexile text measure} \quad \text{Equation (1)}$$

was developed to convert logit difficulties to Lexile calibrations. The values of the conversion factor (CF) and the constant were determined by substituting in the low-end anchor point and then solving the system of equations.

The Lexile Scale ranges from below 200L to above 1600L. There is not an explicit bottom or top to the scale, but rather two anchor points on the scale (described above) that describe different levels of reading comprehension. The Lexile Map, a graphic representation of the Lexile Scale from 200L to 1500L+, provides a context for understanding reading comprehension.

Validity of The Lexile Framework for Reading

Validity refers to the “degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests” (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999). In other words, does the test measure what it is supposed to measure? For the Lexile Framework, which measures a skill, the most important aspect of validity that should be examined is construct validity. The validity of the Lexile Framework can be evaluated by examining how well Lexile measures relate to other measures of reading comprehension and text difficulty.

Lexile Framework and other Measures of Reading Comprehension. Table 1 presents the results from studies where students were administered a Lexile assessment and another assessment of reading comprehension. There is a strong relationship between reading comprehension ability as measured by the Lexile Framework and reading comprehension ability as measured by other assessments.

Table 1. Results from linking studies conducted with The Lexile Framework for Reading.

Standardized Test	Grades in Study	N	Correlation Between Test Score and Lexile Measure
Gates-MacGinitie Reading Test	2, 4, 6, 8, 10	4,644	0.90
Metropolitan Achievement Test (8 th ed.)	2, 4, 6, 8, 10	2,382	0.93
Texas Assessment of Knowledge and Skills (TAKS)	3, 5, 8	1,960	0.60 to 0.73*
The Iowa Tests (Iowa Tests of Basic Skills and Iowa Tests of Educational Development)	3, 5, 7, 9, and 11	4,666	0.88
Stanford Achievement Test (Tenth Edition)	2, 4, 6, 8, and 10	3,064	0.93
Oregon Reading/Literature Knowledge and Skills Test	3, 5, 8, and 10	3,180	0.89
Mississippi Curriculum Test	2, 4, 6, and 8	7,045	0.90
Georgia Criterion Referenced Competency Test (CRCT and GHSGT)	1 – 8, and 11	16,363	0.72 to 0.88*
Wyoming Performance Assessment for Wyoming Students (PAWS)	3, 5, 7, and 11	3,871	0.91
Arizona Instrument to Measure Progress (AIMS)	3, 5, 7, and 10	7,735	0.89
South Carolina Palmetto Achievement Challenge Tests (PACT)	3 – 8	15,559	0.87 to 0.88*
Comprehensive Testing Program (CPT 4 – ERB)	2, 4, 6, and 8	924	0.83 to 0.88
Oklahoma Core Competency Tests (OCCT)	3 – 8	10,691	0.71 to 0.75*
TOEFL iBT	NA	2,906	0.63 to 0.67
TOEIC	NA	2,799	0.73 to 0.74
Kentucky Performance Rating for Educational Progress (K-PREP)	3 – 8	6,480	0.71 to 0.79*
North Carolina ACT	11	3,472	0.84
North Carolina READY End-of-Grades/End-of-Course Tests (NC READY EOG/EOC)	3, 5, 7, 8, and E2	12,356	0.88 to 0.89

Notes: Results are based on final samples used with each linking study.

*Not vertically equated; separate linking equations were derived for each grade.

Lexile Framework and the Difficulty of Basal Readers. In a study conducted by Stenner, Smith, Horabin, and Smith (1987b) Lexile calibrations were obtained for units in 11 basal series. It was presumed that each basal series was sequenced by difficulty. So, for example, the latter portion of a third-grade reader is presumably more difficult than the first portion of the same book. Likewise, a fourth-grade reader is presumed to be more difficult than a third-grade reader. Observed difficulties for each unit in a basal series were estimated by the rank order of the unit in the series. Thus, the first unit in the first book of the first grade was assigned a rank order of one and the last unit of the eighth-grade reader was assigned the highest rank order number.

Correlations were computed between the rank order and the Lexile calibration of each unit in each series. After correction for range restriction and measurement error, the average disattenuated correlation between the Lexile calibration of text comprehensibility and the rank order of the basal units was 0.995 (see *Table 2*).

Table 2. Correlations between theory-based calibrations produced by the Lexile equation and rank order of unit in basal readers.

Basal Series	Number of Units	r_{OT}	R_{OT}	R'_{OT}
Ginn Rainbow Series (1985)	53	.93	.98	1.00
HBJ Eagle Series (1983)	70	.93	.98	1.00
Scott Foresman Focus Series (1985)	92	.84	.99	1.00
Riverside Reading Series (1986)	67	.87	.97	1.00
Houghton-Mifflin Reading Series (1983)	33	.88	.96	.99
Economy Reading Series (1986)	67	.86	.96	.99
Scott Foresman American Tradition (1987)	88	.85	.97	.99
HBJ Odyssey Series (1986)	38	.79	.97	.99
Holt Basic Reading Series (1986)	54	.87	.96	.98
Houghton-Mifflin Reading Series (1986)	46	.81	.95	.98
Open Court Headway Program (1985)	52	.54	.94	.97
Total/Means*	660	.839	.965	.995

r_{OT} = raw correlation between observed difficulties (O) and theory-based calibrations (T).

R_{OT} = correlation between observed difficulties (O) and theory-based calibrations (T) corrected for range restriction.

R'_{OT} = correlation between observed difficulties (O) and theory-based calibrations (T) corrected for range restriction and measurement error.

*Mean correlations are the weighted averages of the respective correlations.

Based on the consistency of the results in *Table 2*, the Lexile theory was able to account for the unit rank ordering of the 11 basal series even with numerous differences in the series – prose selections, developmental range addressed, types of prose introduced (i.e., narrative versus expository), and purported skills and objectives emphasized.

Lexile Framework and the Difficulty of Reading Test Items. In a study conducted by Stenner, Smith, Horabin, and Smith (1987a), 1,780 reading comprehension test items appearing on nine nationally-normed tests were analyzed. The study correlated empirical item difficulties provided by the publishers with the Lexile calibrations specified by the computer analysis of the text of each item. The empirical difficulties were obtained in one of three ways. Three of the tests included observed logit difficulties from either a Rasch or three-parameter analysis (e.g., NAEP). For four of the tests, logit difficulties were estimated from item p -values and raw score means and standard deviations (Poznanski, 1990; Wright, and Linacre, 1994). Two of the tests provided no item parameters, but in each case items were ordered on the test in terms of difficulty (e.g., PIAT). For these two tests, the empirical difficulties were approximated by the difficulty rank order of the items. In those cases where multiple questions were asked about a single passage, empirical item difficulties were averaged to yield a single observed difficulty for the passage.

Once theory-specified calibrations and empirical item difficulties were computed, the two arrays were correlated and plotted separately for each test. The plots were checked for unusual residual distributions and curvature, and it was discovered that the Lexile equation did not fit poetry items or noncontinuous prose items (e.g., recipes, menus, or shopping lists). This indicated that the universe to which the Lexile equation could be generalized was limited to continuous prose. The poetry and noncontinuous prose items were removed and correlations were recalculated. *Table 3* contains the results of this analysis.

Table 3. Correlations between theory-based calibrations produced by the Lexile equation and empirical item difficulties.

Test	Number of Questions	Number of Passages	Mean	SD	Range	Min	Max	r_{OT}	R_{OT}	R'_{OT}
SRA	235	46	644	353	1303	33	1336	.95	.97	1.00
CAT-E	418	74	789	258	1339	212	1551	.91	.95	.98
Lexile	262	262	771	463	1910	-304	1606	.93	.95	.97
PIAT	66	66	939	451	1515	242	1757	.93	.94	.97
CAT-C	253	43	744	238	810	314	1124	.83	.93	.96
CTBS	246	50	703	271	1133	173	1306	.74	.92	.95
NAEP	189	70	833	263	1162	169	1331	.65	.92	.94
Battery	26	26	491	560	2186	-702	1484	.88	.84	.87
Mastery	85	85	593	488	2135	-586	1549	.74	.75	.77
Total/ Mean	1780	722	767	343	1441	50	1491	.84	.91	.93

r_{OT} = raw correlation between observed difficulties (O) and theory-based calibrations (T).

R_{OT} = correlation between observed difficulties (O) and theory-based calibrations (T) corrected for range restriction.

R'_{OT} = correlation between observed difficulties (O) and theory-based calibrations (T) corrected for range restriction and measurement error.

*Means are computed on Fisher Z transformed correlations.

The last three columns in *Table 3* show the raw correlation between observed (O) item difficulties and theoretical (T) item calibrations, with the correlations corrected for restriction in range and measurement error. The Fisher Z mean of the raw correlations (r_{OT}) is 0.84. When corrections are made for range restriction and measurement error, the Fisher Z mean disattenuated correlation between theory-based calibration and empirical difficulty in an unrestricted group of reading comprehension items (R'_{OT}) is 0.93. These results show that most attempts to measure reading comprehension, no matter what the item form, type of skill objectives assessed, or response requirement used, measure a common comprehension factor specified by the Lexile theory.

Text Measure Error Associated with the Lexile Framework

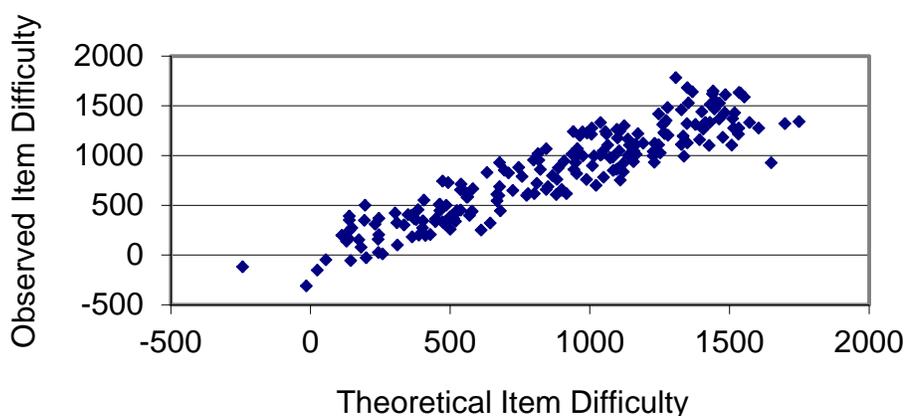
To determine a Lexile measure for a text, the standard procedure is to process the entire text. All pages in the work are concatenated into an electronic file that is processed by a software package called the Lexile Analyzer (developed by MetaMetrics, Inc.). The analyzer “slices” the text file into as many 125-word passages as possible, analyzes the set of slices, and then calibrates each slice in terms of the logit metric. That set of calibrations is then processed to determine the Lexile measure corresponding to a 75% comprehension rate. The analyzer uses the slice calibrations as test item calibrations and

then solves for the measure corresponding to a raw score of 75% (e.g., 30 out of 40 correct, as if the slices were test items). The Lexile Analyzer automates this process, but what “certainty” can be attached to each text measure?

Using the bootstrap procedure to examine error due to the text samples, the above analysis could be repeated (Efron, 1981; Sitter, 1992). The result would be an identical text measure to the first because there is no sampling error when a complete text is calibrated.

There is, however, another source of error that increases the uncertainty about where a text is located on the Lexile Map. The Lexile Theory is imperfect in its calibration of the difficulty of individual text slices. To examine this source of error, 200 items that had been previously calibrated and shown to fit the model were administered to 3,026 students in Grades 2 through 12 in a large urban school district. For each item the observed item difficulty calibrated from the Rasch model was compared with the theoretical item difficulty calibrated from the regression equation used to calibrate texts. A scatter plot of the data is presented in *Figure 1*.

Figure 1. Scatter plot between observed item difficulty and theoretical item difficulty.



The correlation between the observed and the theoretical calibrations for the 200 items was 0.92 and the root mean square error was 178L. Therefore, for an individual slice of text the measurement error is 178L.

The standard error of measurement associated with a text is a function of the error associated with one slice of text (178L) and the number of slices that are calibrated from a text. Very short books have larger uncertainties than longer books. A book with only four slices would have an uncertainty of 89L whereas a longer book such as *War and Peace* (4,082 slices of text) would only have an uncertainty of 3L (*Table 4*).

Table 4. Standard errors for selected values of the length of the text.

Title	Number of Slices	Text Measure	Standard Error of Text
<i>The Stories Julian Tells</i>	46	520	26
<i>Bunnicula</i>	102	710	18
<i>The Pizza Mystery</i>	137	620	15
<i>Meditations of First Philosophy</i>	206	1720	12
<i>Metaphysics of Morals</i>	209	1620	12
<i>Adventures of Pinocchio</i>	294	780	10
<i>Red Badge of Courage</i>	348	900	10
<i>Scarlet Letter</i>	597	1420	7
<i>Pride and Prejudice</i>	904	1100	6
<i>Decameron</i>	2431	1510	4
<i>War and Peace</i>	4082	1200	3

A typical Grade 3 reading test has approximately 2,000 words in the passages. To calibrate this text, it would be sliced into 16 125-word passages. The error associated with this text measure would be 45L. A typical Grade 7 reading test has approximately 3,000 words in the passages and the error associated with the text measure would be 36L. A typical Grade 10 reading test has approximately 4,000 words in the passages and the error associated with the text measure would be 30L.

The Find A Book (www.Lexile.com) contains information about each book analyzed: author, Lexile measure and Lexile Code, awards, ISBN, and developmental level as determined by the publisher. Information concerning the length of a book and the extent of illustrations – factors that affect a reader’s perception of the difficulty of a book – can be obtained from MetaMetrics.

Lexile Item Bank

The Lexile Item Bank contains over 10,000 items that have been developed since 1986 for research purposes with the Lexile Framework.

Passage Selection. Passages selected for use are selected from “real world” reading materials that students may encounter both in and out of the classroom. Sources include textbooks, literature, and periodicals from a variety of interest areas and material written by authors of different backgrounds. The following criteria are used to select passages:

- the passage must develop one main idea or contain one complete piece of information;
- understanding of the passage is independent of the information that comes before or after the passage in the source text; and

- understanding of the passage is independent of prior knowledge not contained in the passage.

With the aid of a computer program, item writers examine blocks of text (minimum of three sentences) that are calibrated to be within 100L of the source text. From these blocks of text item writers are asked to select four to five that could be developed as items. If it is necessary to shorten or lengthen the passage in order to meet the criteria for passage selection, the item writer can immediately recalibrate the text to ensure that it is still targeted within 100L of the complete text (source targeting).

Item Format. The native Lexile item format is embedded completion. The embedded completion format is similar to the fill-in-the-blank format. When properly written, this format directly assesses the reader’s ability to draw inferences and establish logical connections between the ideas in the passage (Haladyna, 1994). The reader is presented with a passage of approximately 30 to 150 words in length. The passages are shorter for beginning readers and longer for more advanced readers. The passage is then response illustrated (a statement is added at the end of the passage with a missing word or phrase followed by four options). From the four presented options, the reader is asked to select the “best” option that completes the statement. With this format, all options are semantically and syntactically appropriate completions of the sentence, but one option is unambiguously the “best” option when considered in the context of the passage.

The statement portion of the embedded completion item can assess a variety of skills related to reading comprehension: paraphrase information in the passage, draw a logical conclusion based on the information in the passage, make an inference, identify a supporting detail, or make a generalization based on the information in the passage. The statement is written to ensure that by reading and comprehending the passage the reader is able to select the correct option. When the embedded completion statement is read by itself, each of the four options is plausible.

Item Writer Training. Item writers are classroom teachers and other educators who have had experience with the everyday reading ability of students at various levels. The use of individuals with these types of experiences helps to ensure that the items are valid measures of reading comprehension. Item writers are provided with training materials concerning the embedded completion item format and guidelines for selecting passages, developing statements, and selecting options. The item writing materials also contain incorrect items that illustrate the criteria used to evaluate items and corrections based on those criteria. The final phase of item writer training is a short practice session with three items.

Item writers are provided vocabulary lists to use during statement and option development. The vocabulary lists were compiled from spelling books one grade level below the level the item would typically be used with. The rationale was that these

words should be part of a reader’s “working” vocabulary since they had been learned the previous year.

Item writers are also given extensive training related to “sensitivity” issues. Part of the item writing materials address these issues and identify areas to avoid when selecting passages and developing items. The following areas are covered: violence and crime, depressing situations/death, offensive language, drugs/alcohol/tobacco, sex/attraction, race/ethnicity, class, gender, religion, supernatural/magic, parent/family, politics, animals/environment, and brand names/junk food. These materials were developed based on material published by McGraw-Hill (*Guidelines for Bias-Free Publishing*, 1983). This publication discusses the equal treatment of the sexes, fair representation of minority groups, and the fair representation of disabled individuals.

Item Review. All items are subjected to a two-stage review process. First, items are reviewed and edited by an editor according to the 19 criteria identified in the item writing materials and for sensitivity issues. Approximately 25% of the items developed are deleted for various reasons. Where possible items are edited and maintained in the item bank.

Items are then reviewed and edited by a group of specialists that represent various perspectives – test developers, editors, and curriculum specialists. These individuals examine each item for sensitivity issues and for the quality of the response options. During the second stage of the item review process, items are either “approved as presented,” “approved with edits,” or “deleted.” Approximately 10% of the items written are “approved with edits” or “deleted” at this stage. When necessary, item writers receive additional on-going feedback and training.

Item Analyses. As part of the linking studies and research studies conducted by MetaMetrics, items in the Lexile Item Bank are evaluated in terms of difficulty (relationship between logit [observed Lexile measure] and theoretical Lexile measure), internal consistency (point-biserial correlation), and bias (ethnicity and gender where possible). Where necessary, items are deleted from the item bank or revised and recalibrated.

During the spring of 1999, 8 levels of a Lexile assessment were administered in a large urban school district to students in grades 1 through 12. The 8 test levels were administered in grades 1, 2, 3, 4, 5, 6, 7-8, and 9-12 and ranged from 40 to 70 items depending on the grade level. A total of 427 items were administered across the 8 test levels. Each item was answered by at least 9,000 students (the number of students per level ranged from 9,286 in grade 2 to 19,056 in grades 9-12). The item responses were submitted to a Winsteps IRT analysis. The resulting item difficulties (in logits) were

assigned Lexile measures by multiplying by 180 and anchoring each set of items to the mean theoretical difficulty of the items on the form.

The NC READY EOG Reading/EOC English II – Lexile Framework Linking Process

Description of the Assessments

North Carolina READY End-of-Grade Language Arts/Reading Assessments and End-of-Course English II Assessment. The 2013 North Carolina READY End-of-Grade Language Arts/Reading Assessments and End-of-Course English II Assessment are designed to measure students' proficiency on the Common Core State Standards (CCSS) for English Language Arts, adopted by the North Carolina State Board of Education in June 2010 (NCDPI, 2013d, 2013e). The Common Core State Standards are divided into strands which address a specific set of College and Career Readiness Anchor Standards. These strands are reading, writing, speaking, listening, and language.

The EOG assessments are administered annually to students in Grades 3 through 8 and the English II assessment is administered to students enrolled in English II (generally Grade 10) at the end of the course. Assessment results will be used both for school and district accountability under the NC READY Accountability Model and for Federal reporting purposes (NCDPI, 2013c).

The EOG English Language Arts/Reading assessments at Grades 3 through 8 are multiple-choice tests. These assessments are available only in paper-and pencil format for the 2012–13 school year. Students read authentic selections and then answer questions related to the selections. The reading selections are comprised of literary and informational text based on the *Common Core State Standards*. Knowledge of vocabulary is assessed indirectly through application and understanding of terms within the context of the selection and questions. The EOG assessments of English Language Arts/Reading at Grades 3 through 5 contain 52 total test items. The assessments at Grades 6 through 8 contain 56 total test items (NCDPI, 2013e).

The NC READY EOG Reading assessments were vertically scaled across grades. Each test has scale scores that range from 400 to 500. These scale scores can be compared directly from grade-to-grade.

The NC READY EOC English II assessment addresses a common set of standards for the second-year high school course of English language arts (NCDPI, 2013c). The English II assessment consists of reading passages and associated items addressing three strands of the CCSS: Reading, Language and Writing. The reading strand is further divided into two sub-strands of Reading Literature and Reading Information. The NC READY tests are approximately 30-35% Reading Literature, 35-40% Reading Information, 15-20% Language, and 15-20% Writing. The Speaking and Listening strands of the CCSS are not included in the assessment (NCDPI, 2013c).

The English II assessment is a criterion-referenced test (CRT) consisting of 50 operational four-response-option multiple-choice items and 3 operational constructed-response items. The constructed-response items appear throughout the test, integrated with multiple choice items related to text passages. The EOC English II scale scores range from 100 and 200, and these scale scores are on a separate scale.

The Lexile Framework for Reading. The Lexile Framework is a tool that can help teachers, parents, and students locate appropriate reading materials. Text complexity (difficulty) and reader ability are measured in the same unit—the Lexile. Text complexity is determined by examining such characteristics as word frequency and sentence length. Items and text are calibrated using the Rasch model. The typical range of the Lexile Scale is from 200L to 1600L, although actual Lexile measures can range from below zero (BR) to above 1600L (see the discussion on pages 5-6 for more information).

Using multiple-choice items, the Lexile Framework measures reading ability by focusing on skills readers use when studying written materials sampled from various content areas. Each test item consists of a passage that is response-illustrated (a statement is added at the end of the passage with a missing word or phrase followed by four options, or distractors). The skills measured by these items include referring to details in the passage, drawing conclusions, and making comparisons and generalizations. Lexile items do not require prior knowledge of ideas outside of the passage, vocabulary taken out of context, or formal logic.

The Lexile Linking Tests were developed for administration to students in Grades 3, 5, 7, 8, and English II. Characteristics of the Lexile Linking Tests were as similar as possible to the NC READY EOG Reading/EOC English II assessments, including the number of operational items per test and difficulty of the items. For each grade/course, two equivalent forms were developed and administered.

The Lexile Linking Tests contained 44 items on each test form for Grades 3 and 5, and 48 items on each test form for Grades 7 and 8. The number of items on the test for each grade was determined by the number of items on the NC READY EOG Reading/EOC English II assessments. Approximately 80% (35 for Grades 3 and 5, and 38 for Grades 7 and 8) of the items were common across the two grade-level test forms.

The English II Lexile Linking Test contained 56 items. The NC READY EOC English II assessment contains 50 operational multiple-choice items with 3 operational polytomous items and 15 experimental items. Because the Lexile Linking Test includes only dichotomous items, the total possible score for items on the NC READY EOC English II assessment was computed by summing the number of one-point multiple-choice items and the number of score points for the open-ended items. This process yielded a total of 56 score points.

The items for the Lexile Linking Tests were chosen to optimize the match to the target test. The IRT difficulty values associated with the NC READY EOG Reading/EOC English II items were converted to Lexile measures using a computer program developed by MetaMetrics, Inc. (no date). Each Lexile Linking Test had a mean Lexile measure established through analysis of the difficulties of the passages on the target test, normative grade-level means, and the item difficulties for the NC READY EOG Reading/EOC English II assessments for 2013. The following mean targets were set: Grade 3, 722L; Grade 5, 963L; Grade 7, 1129L; Grade 8, 1205L; and English II, 1273L.

Evaluation of T-parallel Lexile Linking Tests. After administration, the Lexile Linking Test items were reviewed. Based on the item examination, four items were removed from further analyses, one item from Grade 3 Form 1, one item from Grade 5 Form 1, one item from Grade 5 Form 2, and one item from English II Form 1. These items indicated an alternate answer choice was more attractive than the correct answer choice. While a few items retained on the tests had low point-biserial correlations, the items performed adequately (average ability measure for the correct answer was highest compared to the average ability measures of the three distractors from the Winsteps analyses). The raw score descriptive statistics for the Lexile Linking Tests are presented in *Table 5*.

Table 5. Descriptive statistics from the development of the Lexile Linking Tests raw scores.

Grade	Test Form	N	Raw Score Mean (SD)	Minimum Score		Maximum Score	
				Observed	Possible	Observed	Possible
3	1	1,197	27.72 (9.3)	4	0	43	43
3	2	1,144	28.97 (9.7)	5	0	44	44
5	1	1,151	31.18 (7.8)	1	0	43	43
5	2	1,134	31.18 (7.9)	8	0	43	43
7	1	1,142	33.15 (9.5)	2	0	48	48
7	2	1,110	32.79 (9.5)	0	0	48	48
8	1	1,485	31.27 (9.8)	5	0	48	48
8	2	1,473	31.11 (9.4)	2	0	48	48
Eng II	1	1,334	38.67 (11.9)	0	0	55	55
Eng II	2	1,320	38.92 (11.9)	4	0	56	56
Total		12,490					

Selected item statistics for the Lexile Linking Tests are presented in *Table 6*.

Table 6. Item statistics from the administration of the Lexile Linking Tests.

Grade		<i>N</i> (Persons)	<i>N</i> (Items)	Percent Correct Mean (Range)	Point- Biserial Range	Coefficient Alpha
3	1	1,197	43	64 (22 - 94)	0.24 - 0.60	0.920
3	2	1,144	44	66 (25 - 89)	0.29 - 0.61	0.926
5	1	1,151	43	73 (28 - 97)	0.08 - 0.57	0.902
5	2	1,134	43	73 (34 - 98)	0.23 - 0.57	0.903
7	1	1,142	48	69 (31 - 92)	0.13 - 0.59	0.918
7	2	1,110	48	68 (21 - 93)	0.12 - 0.61	0.918
8	1	1,485	48	65 (28 - 89)	0.11 - 0.56	0.919
8	2	1,473	48	65 (33 - 90)	0.11 - 0.54	0.910
Eng II	1	1,334	55	70 (31 - 91)	0.26 - 0.64	0.944
Eng II	2	1,320	56	70 (26 - 93)	0.20 - 0.64	0.941
Total		12,490				

The Coefficient Alpha correlations for each of the ten Lexile Linking Tests, two for each grade/course, ranged from 0.902 to 0.944. This indicates strong internal consistency reliability for each of the ten tests and high consistency across these ten tests.

Study Design

A single-group/common-person design was chosen for this study (Kolen and Brennan, 2004). This design is most useful “when (1) administering two sets of items to examinees is operationally possible, and (2) differential order effects are not expected to occur” (pp. 16–17). The NC READY EOG Reading assessments were administered between April 8, 2013 and April 26, 2013. The Lexile Linking Tests were administered within two weeks of the administration of the NC READY EOG Reading assessments. The NC READY EOC English II assessment was administered between April 29, 2013 and May 15, 2013. The Lexile Linking Test was administered within two weeks of the administration of the NC READY EOC English II assessment.

Description of the Sample

The sample of students for the study was selected by the North Carolina Department of Public Instruction. The participating schools were located from across North Carolina with a total of 121 schools from 75 districts participating in the linking study.

Table 7 presents the number of students tested in the linking study and the percentage of students with complete data (both a NC READY EOG Reading/EOC English II score and a Lexile Linking Test Lexile measure). A total of 12,356 students (Grades 3, 5, 7, 8, and English II), or 98.9%, had both test scores. This sample will be referred to as the matched sample.

Table 7. Number of student tests received and number of students in the matched sample.

Grade	NC READY EOG Reading/EOC English II Received <i>N</i>	Lexile Linking Test <i>N</i>	Matched <i>N</i>	Matched Percent
3	103,173	2,341	2,318	99.0
5	109,836	2,285	2,260	98.9
7	110,944	2,252	2,224	98.8
8	108,983	2,958	2,939	99.4
Eng II	108,188	2,654	2,615	98.5
Total	541,124	12,490	12,356	98.9

All students and items were submitted to a Winsteps (Linacre, 2011) analysis using a logit convergence criterion of 0.0001 and a residual convergence criterion of 0.003.

To account for individual differences in motivation when responding to the two assessments, the sample set was trimmed. Test scores from each of the assessments were rank ordered and then converted to percentiles. For each student, the difference in percentiles between the two assessments was examined. A screen of a 25-percentile-point difference was selected for all tests. This helped to minimize the number of students removed from the sample and maintain the characteristics of the distribution, while at the same time removing students that were obvious outliers on one or both of the assessments.

For the final sample of students used in the study, students in the matched sample with the following score patterns were removed:

- Accommodations that effect the construct being measured,
- 100% correct on the Lexile Linking Test,
- Missing total score on the NC READY EOG Reading/EOC English II assessment,
- Misfit to the Rasch model, or
- Showed greater than a 25-percentile-rank difference between the NC READY EOG Reading/EOC English II assessment scale scores and Lexile Linking Test Lexile measures within grade.

Table 8 shows, for each grade, the number of students (*N*) in the final sample and the percent each grade *N*-count represents of the original matched sample. Of the 12,356 students in the matched sample, 9,777 (79.1%) remained in the final sample. The table also summarizes the number of student test scores (by grade) removed from analysis, and the reason for their removal.

Table 8. Comparison of matched sample and final sample and the reason for student removal.

Matched Sample		<i>N</i> Removed by Reason				Final Sample	
Grade	<i>N</i>	Accommodated Students	Misfit to Rasch	Scores Removed*	Percentile Rank Difference	<i>N</i>	Percent of Matched Sample
3	2,318	3	91	40	281	1,903	82.1
5	2,260	2	130	24	377	1,727	76.4
7	2,224	1	59	15	379	1,770	79.6
8	2,939	9	74	23	524	2,309	78.6
Eng II	2,615	0	47	49	451	2,068	79.1
Total	12,356	15	401	151	2,012	9,777	79.1

* Note: Students with a 100% correct on the linking test or with an invalid NC READY EOG Reading/EOC English II assessment score.

Table 9 presents the demographic characteristics of all students in the NC READY EOG Reading/EOC English II sample, the matched sample, and the final sample of students included in this study. Across the samples, the final sample is similar to the other two samples.

Table 9. Percentage of students in the NC READY EOG Reading/EOC English II sample, matched sample, and final sample for selected demographic characteristics.

Student Characteristic	Category	State Sample N=541,124	Matched Sample N= 12,356	Final Sample N=9,777
Grade or Course	3	19.1	18.8	19.5
	5	20.3	18.3	17.7
	7	20.5	18.0	18.1
	8	20.1	23.8	23.6
	English II	20.0	21.2	21.2
Gender	Female	49.6	49.6	50.4
	Male	50.4	50.4	49.6
	Unknown/not avail	0.1	0.0	0.0
Race/Ethnicity	American Indian	1.4	0.9	1.0
	Asian	2.6	2.4	2.4
	Black	25.7	24.7	24.5
	Hispanic	13.4	12.8	13.2
	Pacific Islander	0.1	0.2	0.2
	White	53.1	55.6	55.3
	Two or more	3.7	3.4	3.5
	N/A	0.1	0.0	0.0
LEP Status	Currently identified	5.4	5.1	5.4
	Exit by committee	0.0	0.0	0.0
	Exits LEP	5.6	5.7	5.7
	Never identified	88.8	89.1	88.7
	No Status	0.1	0.0	0.0
	Parental refusal of IPT testing	0.1	0.1	0.1
Student/Disability	Exited within 2 years	1.7	1.6	1.5
	Yes	8.9	8.5	8.8
	No	89.4	90.0	89.7

Student Characteristic	Category	State Sample N=541,124	Matched Sample N=12,356	Final Sample N=9,777
EC Code	Autism	0.5	0.6	0.6
	Deaf-Blindness	0.0	0.0	0.0
	Deafness	0.0	0.0	0.0
	Developmental Delay	0.1	0.0	0.0
	Hearing Impairment	0.1	0.1	0.1
	Intell. Disability - Mild	0.2	0.2	0.2
	Intell. Disability - Moderate	0.0	0.0	0.0
	Multiple Disabilities	0.0	0.0	0.0
	Not Provided	89.4	90.0	89.7
	Orthopedic Impairment	0.0	0.1	0.1
	Other Health Impairment	2.3	2.1	2.1
	Serious Emotional Disability	0.4	0.2	0.2
	Specific Learning Disability	5.2	4.7	4.9
	Speech or Language Impairment	1.9	2.1	2.1
	Traumatic Brain Injury	0.0	0.0	0.0
	VI	0.0	0.0	0.0
Plan-504	Yes	1.1	1.4	1.4
	No	98.9	98.6	98.6
Word To Word Bilingual	Yes	0.2	0.1	0.0
	No	99.8	99.9	100.0
Acad/Intell Gifted - Reading	Yes	10.8	10.1	10.0
	No	89.2	89.9	90.0

Table 10 presents the descriptive statistics for the NC READY EOG Reading/EOC English II scale scores and the Lexile Linking Test Lexile measures for the matched sample. The correlations between the NC READY EOG Reading/EOC English II scale scores and the Lexile Linking Test measures range from 0.769 to 0.824. Based upon the correlations between the NC READY EOG Reading/EOC English II scale scores and the

Lexile Linking Test Lexile measures presented in *Table 10*, it can be concluded that the two tests are measuring similar reading comprehension constructs.

Table 10. Descriptive statistics for the NC READY EOG Reading/EOC English II scale scores and Lexile measures and the Lexile Linking Test Lexile measures, matched sample ($N = 12,356$).

Grade	N	Matched Sample NC READY EOG Reading/EOC English II Scale Score Mean (SD)	Matched Sample Lexile Linking Test Lexile Measure Mean (SD)	r
3	2,318	440.18 (10.4)	697.98 (253.4)	0.824
5	2,260	449.18 (9.5)	1019.58 (226.5)	0.795
7	2,224	455.81 (10.2)	1138.34 (237.4)	0.769
8	2,939	458.55 (10.7)	1168.69 (226.8)	0.770
Eng II	2,615	150.68 (9.0)	1295.86 (259.2)	0.769
Total	12,356			

Table 11 presents the descriptive statistics of the NC READY EOG Reading/EOC English II test scale scores as well as the Lexile Linking Test Lexile measures for the final sample. The correlations between the final sample NC READY EOG Reading/EOC English II scale scores and the final sample Lexile Linking Test measures range from 0.877 to 0.893. These correlations between the two scores are strong and higher than the matched sample.

Table 11. Descriptive statistics for the NC READY EOG Reading/EOC English II scale scores and the Lexile Linking Test Lexile measures, final sample ($N = 9,777$).

Grade	N	Final Sample NC READY EOG Reading/EOC English II Scale Score Mean (SD)	Final Sample Lexile Linking Test Lexile Measure Mean (SD)	r
3	1,903	439.69 (10.1)	686.13 (233.3)	0.893
5	1,727	449.12 (9.3)	1016.02 (209.8)	0.883
7	1,770	455.65 (10.3)	1135.65 (229.9)	0.877
8	2,309	458.41 (10.7)	1169.21 (217.5)	0.888
Eng II	2,068	150.30 (9.1)	1285.82 (239.1)	0.887
Total	9,777			

Figures 2 through 11 shows the relationship between the NC READY EOG Reading/EOC English II scale scores and the Lexile Linking Test Lexile measures for the matched and final samples for each grade/course. In each grade/course, it can be seen that there is a linear relationship between the NC READY EOG Reading/EOC English II scale score and the final sample Lexile measure reinforcing the use of linear equating.

Figure 2. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 3 matched sample (N = 2,318).

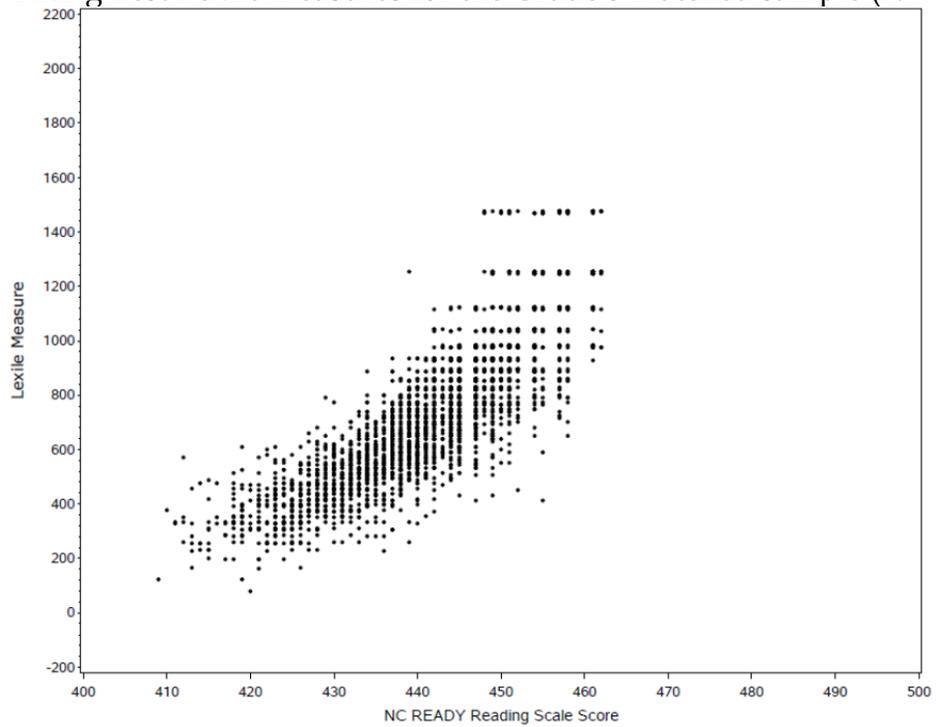


Figure 3. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 3 final sample (N = 1,903).

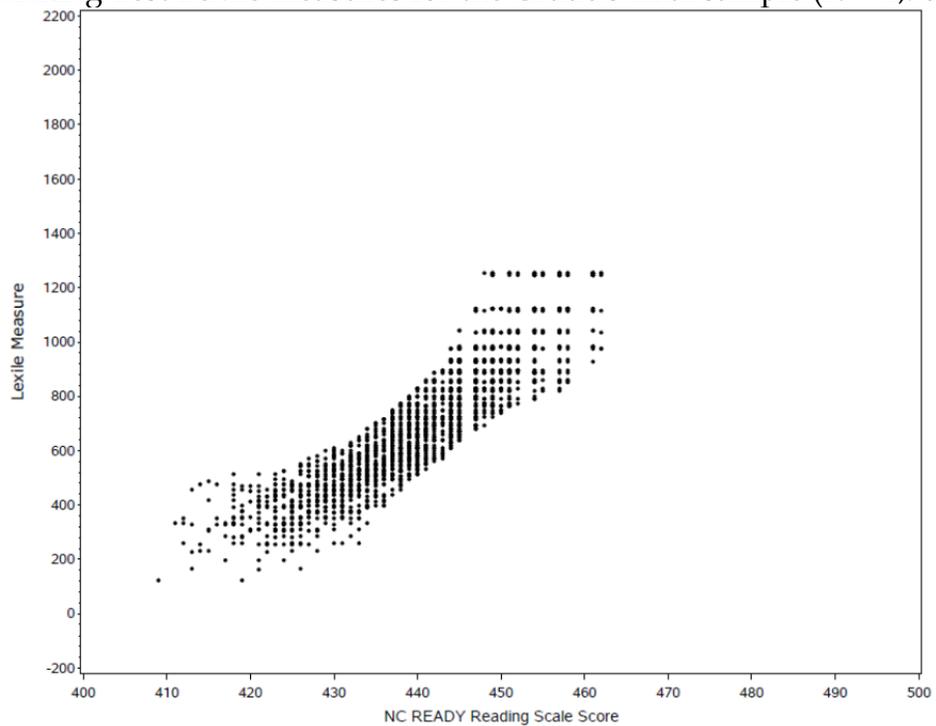


Figure 4. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 5 matched sample ($N = 2,260$).

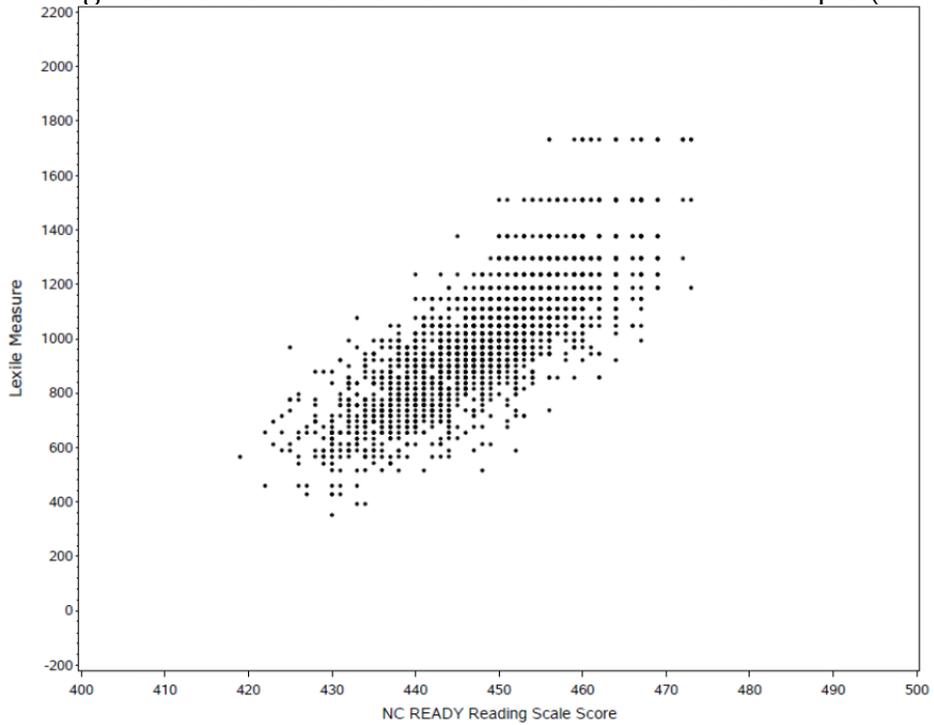


Figure 5. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 5 final sample ($N = 1,727$).

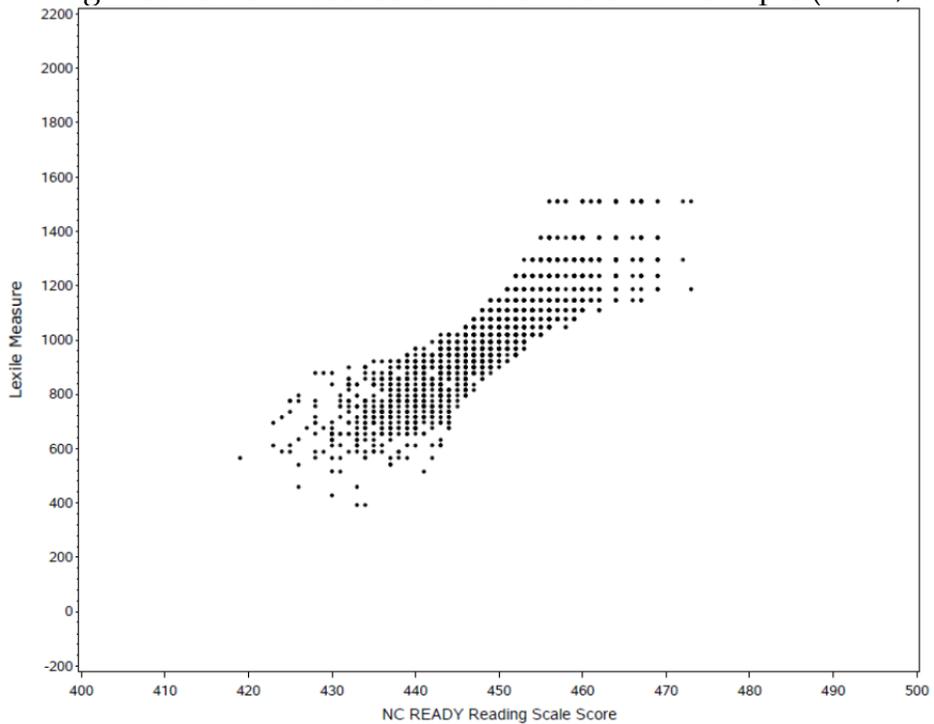


Figure 6. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 7 matched sample (N = 2,224).

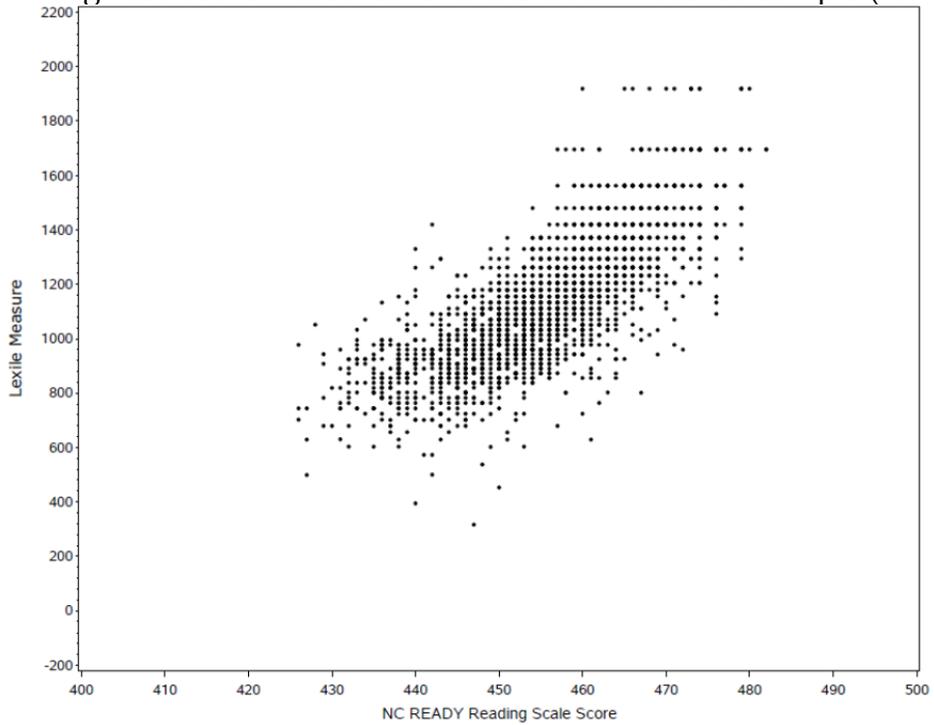


Figure 7. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 7 final sample (N = 1,770).

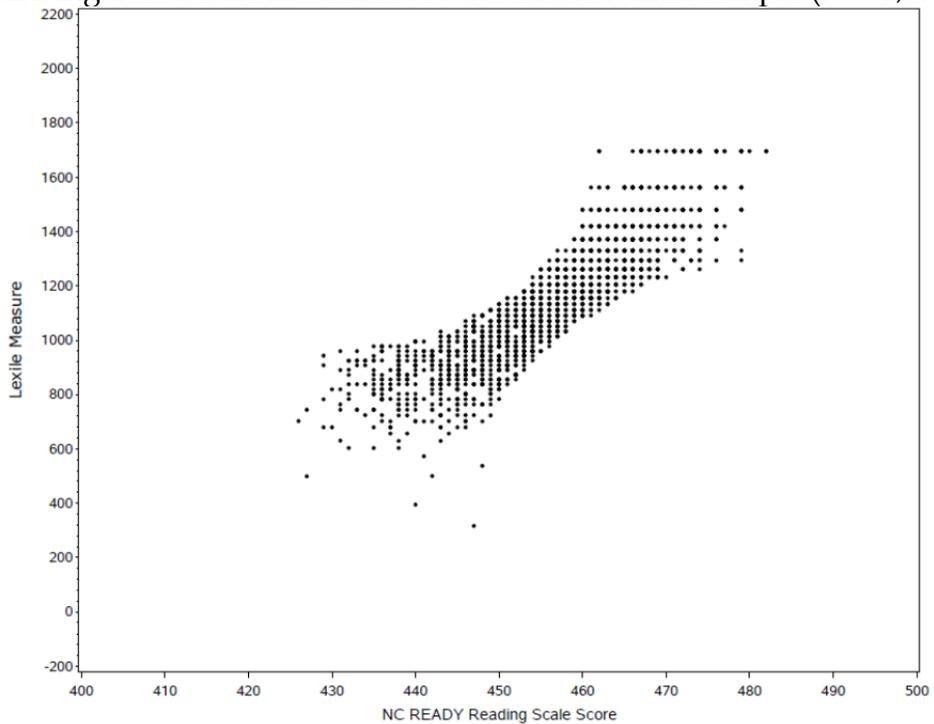


Figure 8. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 8 matched sample ($N = 2,939$).

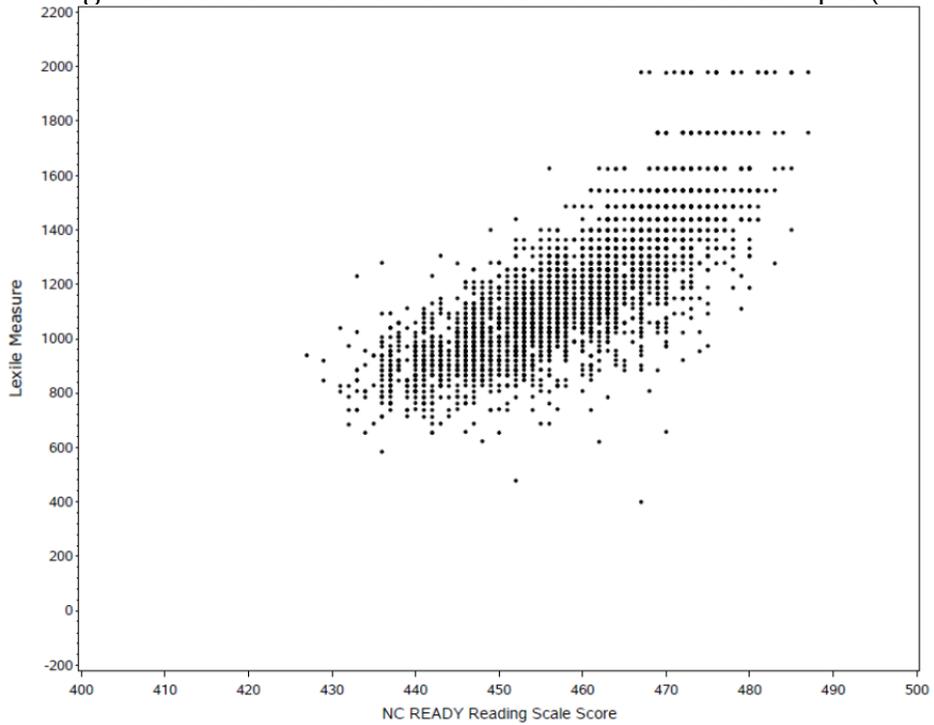


Figure 9. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 8 final sample ($N = 2,309$).

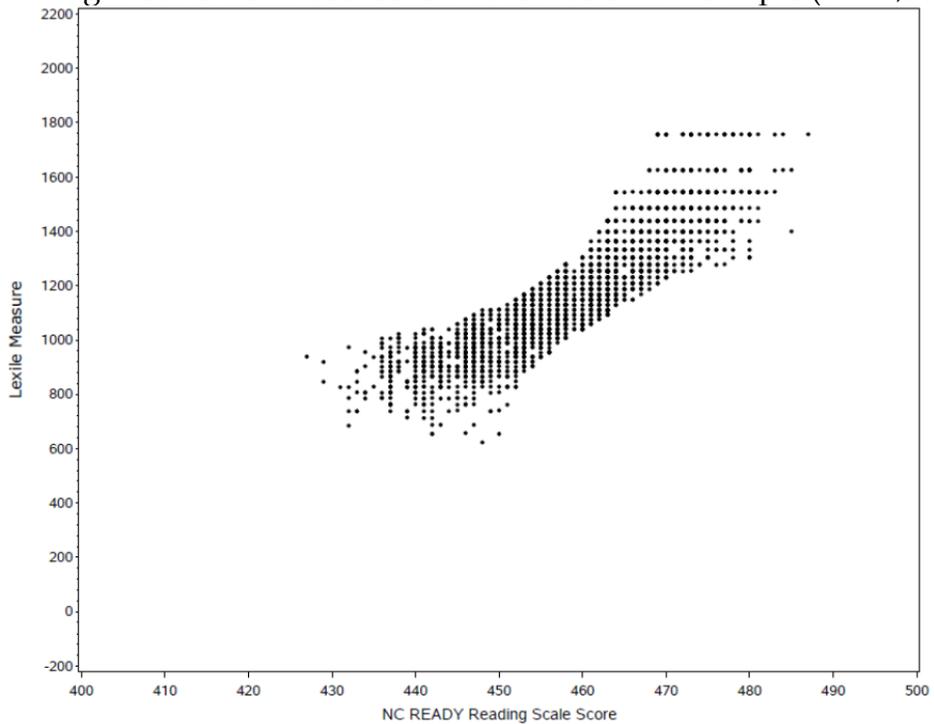


Figure 10. Scatter plot of the NC READY EOC English II scale scores and the Lexile Linking Test Lexile measures for the English II matched sample ($N = 2,615$).

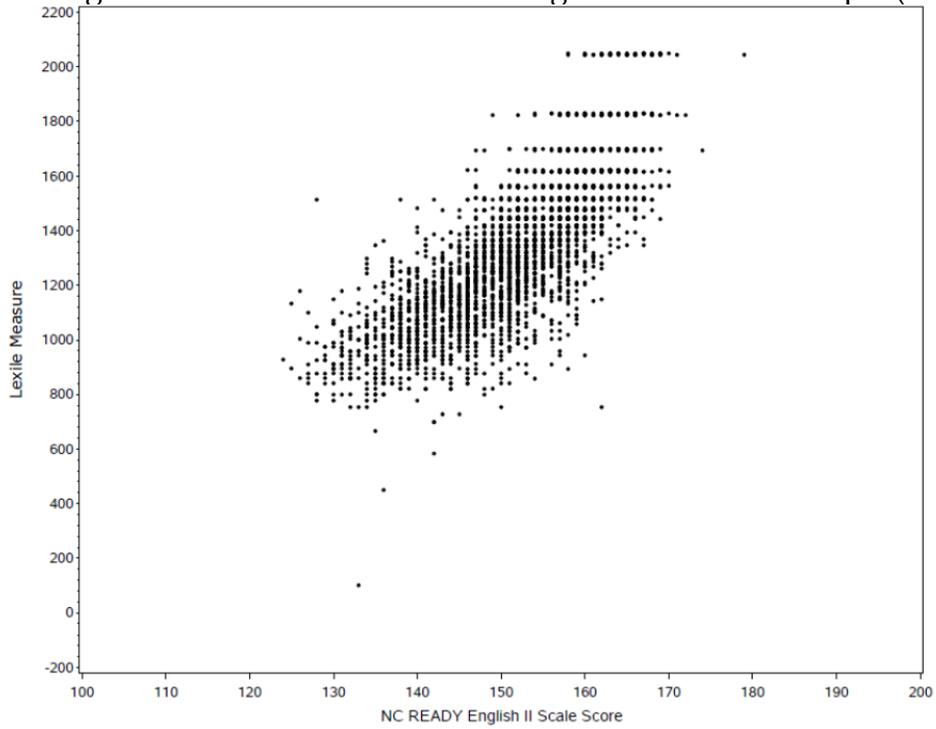
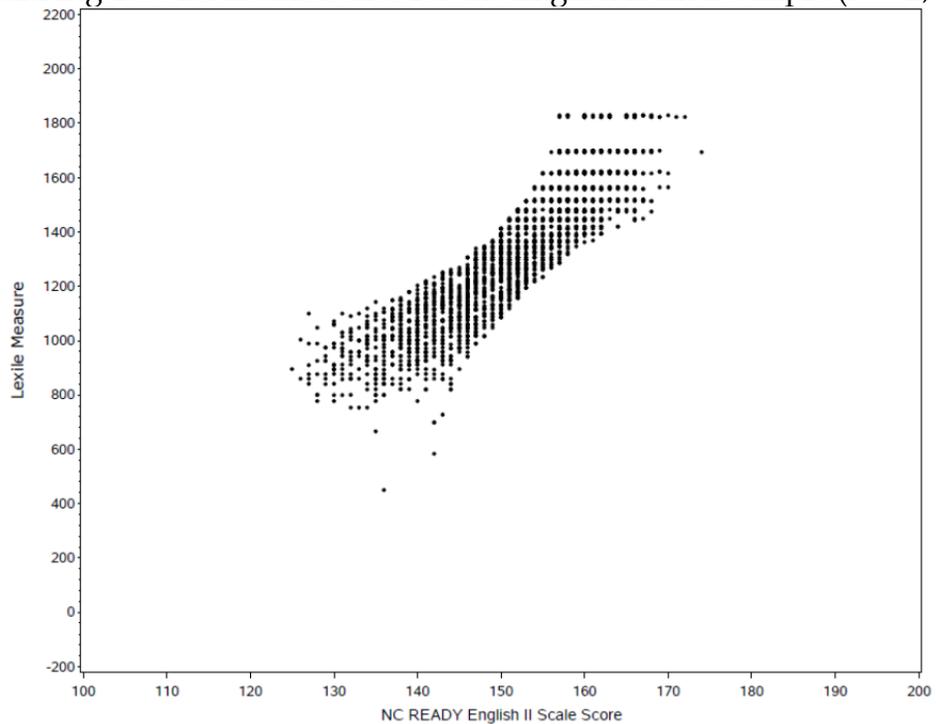


Figure 11. Scatter plot of the NC READY EOC English II scale scores and the Lexile Linking Test Lexile measures for the English II final sample ($N = 2,068$).



Linking the NC READY EOG Reading/EOC English II Scale Scores with the Lexile Scale

Linking in general means “putting the scores from two or more tests on the same scale” (National Research Council, 1999, p.15). MetaMetrics and the North Carolina Department of Public Instruction conducted this linking study for the purpose of matching students with books and texts – to predict the books and texts a student should be matched with for successful reading experiences, given their performance on the NC READY EOG Reading/EOC English II assessment.

Evaluation of linkage assumptions. Factors that affect the linkage between two assessments include the domain to be assessed, the definition of the framework for assessment, the test specifications, and the items sampled.

Based upon the correlations between the NC READY EOG Reading/EOC English II scale scores and the Lexile Linking Tests Lexile measures presented in *Table 11*, it can be concluded that the two assessments measure similar constructs. The correlations between the two assessments are above or within the typical range of alternate-form reliability coefficients; therefore, the Lexile Linking Tests can be considered a T-parallel form of the NC READY EOG Reading/EOC English II test (see Note 1). By using alternate-form reliability coefficients as a comparison, similar sources of variation are accounted for (differences in testing occasions and items). In addition, the linking tests were constructed to have a similar number of items and the same level of difficulty as the NC READY EOG Reading/EOC English II assessments.

Linking Analyses. Two score scales (e.g., the NC READY EOG Reading/EOC English II scale and the Lexile Scale) can be linked using linear equating when (1) test forms have similar difficulties; and (2) simplicity in conversion tables or equations, in conducting analyses, and in describing procedures are desired (Kolen and Brennan, 2004).

In linear equating, a transformation is chosen such that scores on two sets of items are considered to be equated if they correspond to the same number of standard deviations above (or below) the mean in some group of examinees (Angoff, 1984, cited in Petersen, Kolen, and Hoover, 1989; Kolen and Brennan, 2004). Given scores x and y on tests X and Y , the linear relationship is

$$\frac{(x - \mu_x)}{\sigma_x} = \frac{(y - \mu_y)}{\sigma_y} \quad (\text{Equation 2})$$

and the linear transformation l_x (called the SD line in this report) used to transform scores on test Y to scores on text X is

$$x = I_x(y) = \left(\frac{\sigma_x}{\sigma_y} \right) y + \left(\mu_x - \frac{\mu_y \sigma_x}{\sigma_y} \right) \quad (\text{Equation 3})$$

Linear equating by definition has the same mean and standard deviation for the overall equation when the scale is vertically aligned. The means and standard deviations are the same for the Linking test and the Target test when calculated across grades. The values are somewhat different when the formula is developed by grade. Linear equating using an SD-line approach is preferable to linear regression because the tests are not perfectly correlated. With less than perfectly reliable tests, linear regression is dependent on which way the regression is conducted: predicting scores on test X from scores on test Y or predicting scores on test Y from scores on test X. The SD line provides the symmetric linking function that is desired.

The final linking equation between NC READY EOG Reading/EOC English II scale scores and Lexile measures can be written as:

$$\text{Lexile measure} = \text{Slope}_g(\text{NC READY EOG Reading/EOC English II scale score}) + \text{constant}_g \quad (\text{Equation 4})$$

where the slope is the ratio of the standard deviations of the NC READY EOG Reading/EOC English II scale scores and Lexile Linking Test Lexile measures. These values for each grade range/course can be found in *Table 11*.

Using the final sample data described in *Table 11*, the linear linking functions relating the NC READY EOG Reading/EOC English II scale scores and Lexile measures for students in the final sample are presented in *Table 12*. One linking function was developed for each of the following groups (*g*): (1) Grades 3 through 8 of the NC READY EOG Reading assessment and (2) EOC English II assessment.

Table 12. Linear linking equation coefficients used to predict Lexile measures from the NC READY EOG Reading and the EOC English II scale scores.

Group (<i>g</i>)	Slope	Intercept
3 - 8	23.488825	-9587.222
English II	26.264583	-2661.751

Conversion tables were developed for all grade levels in order to express the NC READY EOG Reading/EOC English II scale scores in the Lexile metric and were delivered to the North Carolina Department of Public Instruction in electronic format.

Table 13 contains the maximum reported Lexile measures by grade. The measures that are reported for an individual student should reflect the purpose for which they will be used. If the purpose of the test is accountability (at the student, school, or district level), then uncapped Lexile measures should be reported. If the purpose is instructional, then the scores should be capped at the upper bound of measurement error (e.g., at the 95th percentile point of the national Lexile norms). In an instructional environment where the purpose of the Lexile measure is to appropriately match readers with texts, all scores below 0L should be reported as “BRxxxL.” No student should receive a negative Lexile measure on a score report. The lowest reported value below 0L is BR400L.

Table 13. Capped values of the Lexile measure by grade/course.

Grade/Course	Capped Lexile Measure
3	1200L
4	1300L
5	1400L
6	1500L
7	1600L
8	1700L
Eng II	1750L

Validity of the NC READY EOG Reading/EOC English II – Lexile Link

Table 14 presents the descriptive statistics and effect size statistics of the NC READY EOG Reading/EOC English II Lexile measures as well as the Lexile Linking Test Lexile measures for the final sample.

Table 14. Descriptive statistics and effect size statistics for the final sample NC READY EOG Reading/EOC English II Lexile measures and the Lexile Linking Test Lexile measures.

Grade	<i>N</i>	Final Sample NC READY EOG Reading/EOC English II Lexile Measure Mean (SD)	Final Sample Lexile Linking Test Lexile Measure Mean (SD)	Effect Size
3	1,903	740.42 (237.1)	686.13 (233.3)	0.230793
5	1,727	961.98 (218.7)	1016.02 (209.8)	-0.252219
7	1,770	1115.5 (240.9)	1135.66 (229.9)	-0.085595
8	2,309	1180.38 (252.7)	1169.21 (217.5)	0.047384
Eng II	2,068	1285.82 (239.2)	1285.82 (239.1)	0.000003
Total	9,777			

The Hedges' *g* effect size shows the relationship between two variables or, in this case, between the NC READY EOG Reading/EOC English II Lexile measure and the Lexile Linking Test Lexile measure. A guideline to use for interpretation of the effect size is:

Table 15. Interpretation chart for effect size.

Small	0.20
Medium	0.50
Large	0.80

In Table 14, for the 5 comparisons, effect sizes were minimal for three comparisons indicating no significant difference between the NC READY EOG Reading/EOC English II Lexile measures and the Lexile Linking Test Lexile measures. Two comparisons, Grades 3 and 5, were slightly larger by at most only .05 within the medium range which was not a concern.

Table 16 contains the percentile ranks of the Lexile Linking Test Lexile measures and the NC READY EOG Reading/EOC English II assessment Lexile measures based on the final sample. The criterion of a half standard deviation (100L) on the Lexile scale was used to determine the size of the difference. In examining the values, the measures are very similar across the distributions. This supports the use of Lexile measures on the NC READY EOG Reading/EOC English II assessments.

Table 16. Comparison of the Lexile measures for selected percentile ranks for the final sample Lexile Linking Test and the NC READY EOG Reading/EOC English II assessment.

Grade 3		
Percentile Rank	Linking Test Lexile Measure	NC READY EOG Reading Sample Lexile Measure
1	255	184
5	333	349
10	398	419
25	507	583
50	659	748
75	852	912
90	983	1030
95	1115	1100
99	1254	1241

Grade 5		
Percentile Rank	Linking Test Lexile Measure	NC READY EOG Reading Sample Lexile Measure
1	567	466
5	675	583
10	736	677
25	878	818
50	1019	959
75	1187	1124
90	1296	1241
95	1377	1312
99	1510	1429

Grade 7		
Percentile Rank	Linking Test Lexile Measure	NC READY EOG Reading Sample Lexile Measure
1	679	560
5	783	701
10	855	795
25	960	959
50	1133	1124
75	1294	1288
90	1420	1429
95	1562	1500
99	1696	1617

Grade 8		
Percentile Rank	Linking Test Lexile Measure	NC READY EOG Reading Sample Lexile Measure
1	741	654
5	848	748
10	902	818
25	1007	1006
50	1149	1171
75	1305	1359
90	1485	1500
95	1546	1570
99	1756	1687

Table 16. (continued). Comparison of the Lexile measures for selected percentile ranks for the final sample Lexile Linking Test and the NC READY EOG Reading/EOC English II assessment.

English II		
Percentile Rank	Linking Test Lexile Measure	NC READY EOC English II Sample Lexile Measure
1	800	726
5	912	858
10	974	963
25	1104	1120
50	1279	1304
75	1449	1462
90	1616	1593
95	1694	1646
99	1829	1751

Performance standards provide a common meaning of test scores throughout a state or nation concerning what is expected at various levels of competence. The North Carolina Department of Instruction established four achievement levels: Level 1, Level 2, Level 3, and Level 4 (NCDPI, 2013b). As an example, the four achievement levels for the Grade 3 NC READY EOG Reading Assessment are:

Level 1: Students performing at this level have **limited command** of the knowledge and skills contained in the *Common Core State Standards (CCSS) Reading Standards for Literature* as assessed by referring to the text when asking and answering questions; recounting stories and determining a central message, explaining how the message is conveyed through key details in the text; describing characters and explaining how their actions contribute to the plot; and determining the meaning of words and phrases as they are used in a text, especially literal and nonliteral language. They will need academic support to engage successfully in this content area.

Level 2: Students performing at this level have **partial command** of the knowledge and skills contained in the *Common Core State Standards (CCSS) Reading Standards for Literature* as assessed by referring to the text when asking and answering questions; recounting stories and determining a central message, explaining how the message is conveyed through key details in the text; describing

characters and explaining how their actions contribute to the plot; and determining the meaning of words and phrases as they are used in a text, especially literal and nonliteral language. They will likely need academic support to engage successfully in this content area.

Level 3: Students performing at this level have **solid command** of the knowledge and skills contained in the *Common Core State Standards (CCSS) Reading Standards for Literature* as assessed by referring to the text when asking and answering questions; recounting stories and determining a central message, explaining how the message is conveyed through key details in the text; describing characters and explaining how their actions contribute to the plot; and determining the meaning of words and phrases as they are used in a text, especially literal and nonliteral language. They are academically prepared to engage successfully in this content area.

Level 4: Students performing at this level have **superior command** of the knowledge and skills contained in the *Common Core State Standards (CCSS) Reading Standards for Literature* as assessed by referring to the text when asking and answering questions; recounting stories and determining a central message, explaining how the message is conveyed through key details in the text; describing characters and explaining how their actions contribute to the plot; and determining the meaning of words and phrases as they are used in a text, especially literal and nonliteral language. They are academically well-prepared to engage successfully in this content area.

The four achievement levels for NC READY EOC English II Assessment (NCDPI, 2013a) are:

Level 1: Students performing at this level have **limited command** of the knowledge and skills contained in the *Common Core State Standards (CCSS) Reading Standards for Literature* as assessed by supporting analysis of the text with textual evidence; determining and analyzing the development and refinement of a theme or idea throughout a text; summarizing a text objectively; analyzing the development, interaction, and contribution of characters in a text; determining meanings of words or phrases in a text; analyzing the impact of word choice on meaning and tone; analyzing how authors' choices create literary effects, such as tension; analyzing point of view and cultural experiences in literature from outside the U.S., drawing on world literature. They will need academic support to engage successfully in this content area.

Level 2: Students performing at this level have **partial command** of the knowledge and skills contained in the *Common Core State Standards (CCSS) Reading Standards for Literature* as assessed by supporting analysis of the text with textual evidence; determining and analyzing the development and refinement of a theme or idea throughout a text; summarizing a text objectively; analyzing the development, interaction, and contribution of characters in a text; determining

meanings of words or phrases in a text; analyzing the impact of word choice on meaning and tone; analyzing how authors' choices create literary effects, such as tension; analyzing point of view and cultural experiences in literature from outside the U.S., drawing on world literature. They will likely need academic support to engage successfully in this content area.

Level 3: Students performing at this level have **solid command** of the knowledge and skills contained in the *Common Core State Standards (CCSS) Reading Standards for Literature* as assessed by supporting analysis of the text with textual evidence; determining and analyzing the development and refinement of a theme or idea throughout a text; summarizing a text objectively; analyzing the development, interaction, and contribution of characters in a text; determining meanings of words or phrases in a text; analyzing the impact of word choice on meaning and tone; analyzing how authors' choices create literary effects, such as tension; analyzing point of view and cultural experiences in literature from outside the U.S., drawing on world literature. They are academically prepared to engage successfully in this content area.

Level 4: Students performing at this level have **superior command** of the knowledge and skills contained in the *Common Core State Standards (CCSS) Reading Standards for Literature* as assessed by supporting analysis of the text with textual evidence; determining and analyzing the development and refinement of a theme or idea throughout a text; summarizing a text objectively; analyzing the development, interaction, and contribution of characters in a text; determining meanings of words or phrases in a text; analyzing the impact of word choice on meaning and tone; analyzing how authors' choices create literary effects, such as tension; analyzing point of view and cultural experiences in literature from outside the U.S., drawing on world literature. They are academically well-prepared to engage successfully in this content area.

Table 17 presents the achievement level cut scores on the NC READY EOG Reading/EOC English II assessments and the associated Lexile measures. There are four achievement levels: Level 1, Level 2, Level 3, and Level 4 (NCDPI, 2013a, 2013b). The values in the table are the cut scores associated with the bottom score for each category.

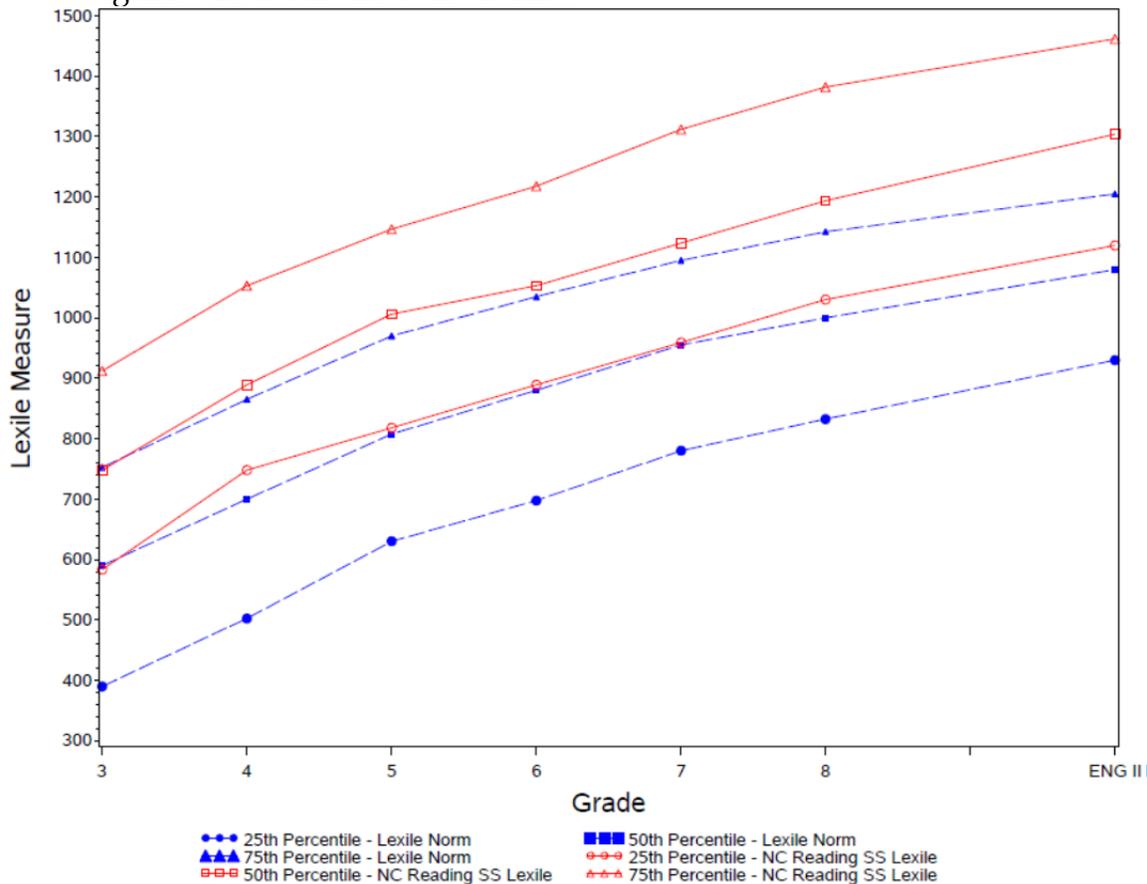
Table 17. NC READY EOG Reading/EOC English II performance level cut scores and the associated Lexile measures.

Grade	Level 2		Level 3		Level 4	
	NC READY EOG Reading/EOC English II Scale Score	Lexile Measure	NC READY EOG Reading/EOC English II Scale Score	Lexile Measure	NC READY EOG Reading/EOC English II Scale Score	Lexile Measure
3	432	560L	442	795L	452	1030L
4	439	725L	448	935L	460	1220L
5	443	820L	453	1055L	464	1310L
6	442	795L	454	1075L	465	1335L
7	445	865L	457	1145L	469	1430L
8	449	960L	462	1265L	473	1525L
E II	141	1040L	151	1305L	165	1670L

Figure 12 shows the Lexile measures for the NC READY EOG Reading/EOC English II assessment as compared to the norms that have been developed for use with The Lexile Framework for Reading. These norms were created based on linking studies conducted with the Lexile Framework.

Overall, it can be seen that the NC READY EOG Reading/EOC English II Lexile measures are higher across the grades at each percentile. The 25th percentile for the NC READY EOG Reading/EOC English II Lexile measures is closer to the 50th percentile Lexile measures. The 50th percentile for the NC READY EOG Reading/EOC English II Lexile measures is closer to the 75th percentile Lexile measures. Therefore, the NC READY EOG Reading/EOC English II scores were higher than the Lexile norms. This translates to the statement that the students in North Carolina were more able than the Lexile norms for a national population.

Figure 12. Selected Percentiles (25th, 50th, and 75th) plotted for the NC READY EOG Reading/EOC English II Lexile measure for the final sample ($N = 9,777$) against the Lexile measure norms.



The following box and whisker plots (*Figures 13, 14, and 15*) show the progression of scores (the y -axis) from grade to grade (the x -axis) (note, that English II is placed as Grade 10 which is the typical grade for students taking the course). For each grade, the box refers to the interquartile range. The line within the box indicates the median and the • represents the mean. The end of each whisker represents the minimum and maximum values of the scores (the y -axis).

The Lexile measures are on a vertical scale and *Figures 13, 14, and 15* demonstrate this by showing that as the grade increases so do the NC READY EOG Reading/EOC English II Lexile measures. All three plots show a similar profile.

Figure 13. Box and whisker plot of the Lexile Linking Tests Lexile measures by grade, final sample (N =9,777).

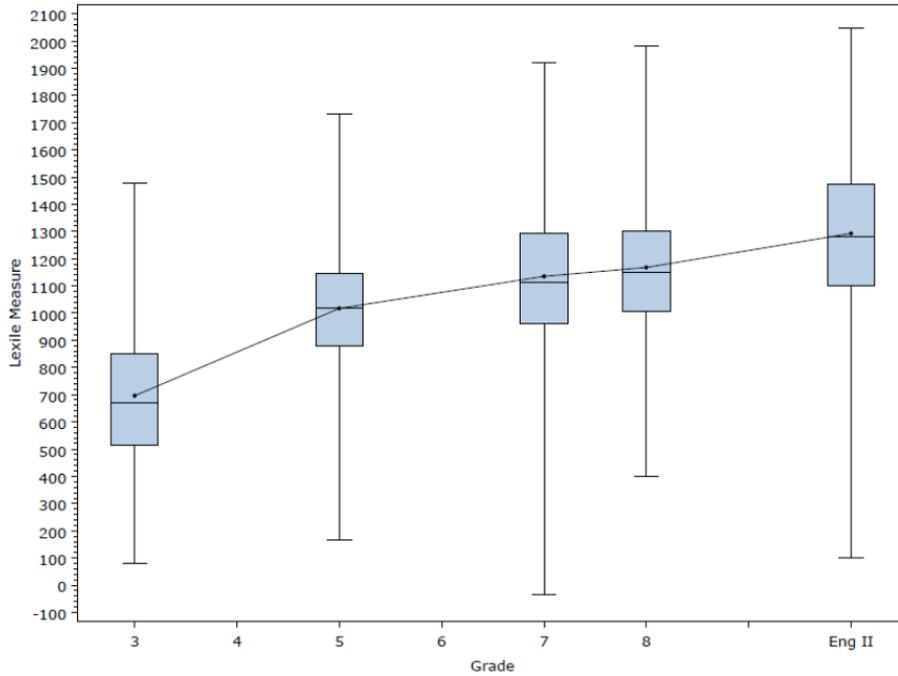


Figure 14. Box and whisker plot of the NC READY EOG Reading/EOC English II Lexile measures by grade, matched sample (N = 12,356).

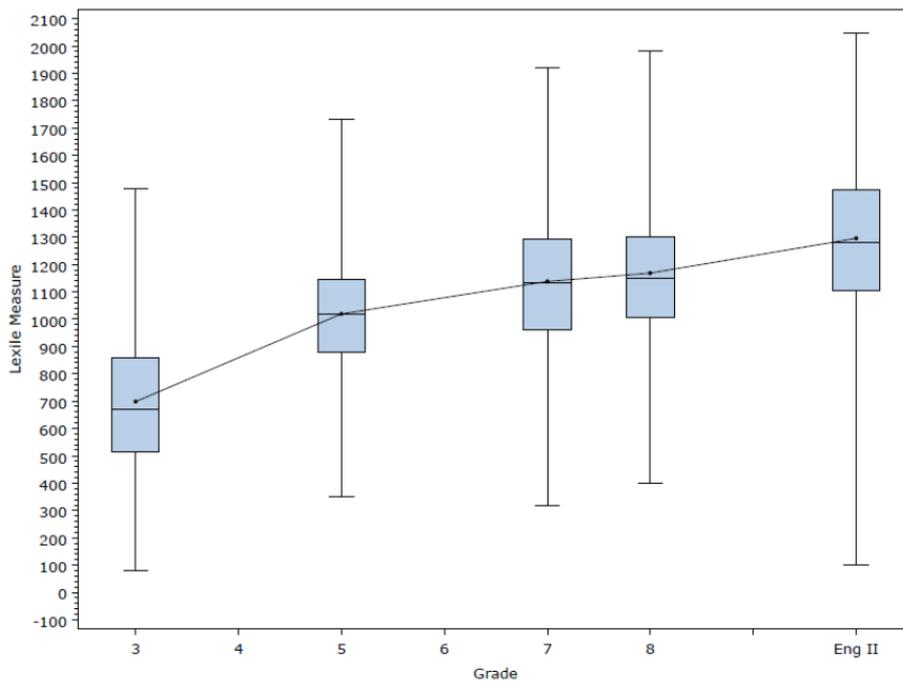
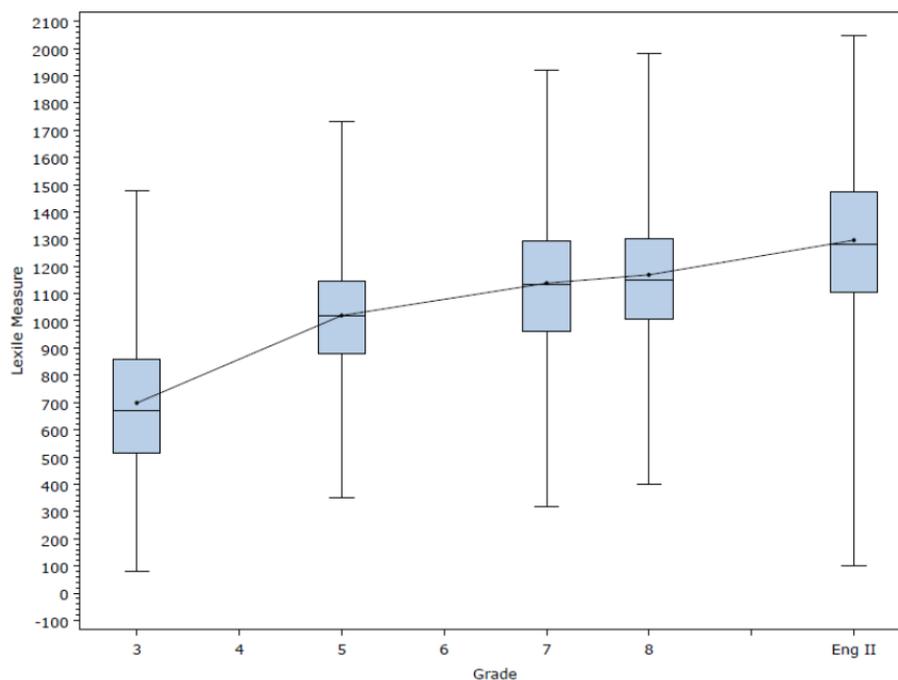


Figure 15. Box and whisker plot of the NC READY EOG Reading/EOC English II Lexile measures by grade, final sample (N = 9,777).



The Lexile Framework and Forecasted Comprehension Rates

A reader with a measure of 600L who is given a text measured at 600L is expected to have a 75-percent comprehension rate. This 75-percent comprehension rate is the basis for selecting text that is targeted to a reader's reading ability, but what exactly does it mean? And what would the comprehension rate be if this same reader were given a text measured at 350L or one at 850L?

The 75-percent comprehension rate for a reader-text pairing can be given an operational meaning by imagining the text is carved into item-sized slices of approximately 125-140 words with a question embedded in each slice. A reader who answers three-fourths of the questions correctly has a 75-percent comprehension rate.

Suppose instead that the text and reader measures are not the same. It is the difference in Lexile measures between reader and text that governs comprehension. If the text measure is less than the reader measure, the comprehension rate will exceed 75 percent. If not, it will be less. The question is "By how much?" What is the expected comprehension rate when a 600L reader reads a 350L text?

If all the item-sized slices in the 350L text had the same calibration, the 250L difference between the 600L reader and the 350L text could be determined using the Rasch model equation. This equation describes the relationship between the measure of a student's

level of reading comprehension and the calibration of the items. Unfortunately, comprehension rates calculated by this procedure would be biased because the calibrations of the slices in ordinary prose are not all the same. The average difficulty level of the slices *and* their variability both affect the comprehension rate.

Although the exact relationship between comprehension rate and the pattern of slice calibrations is complicated, Equation 5 is an unbiased approximation:

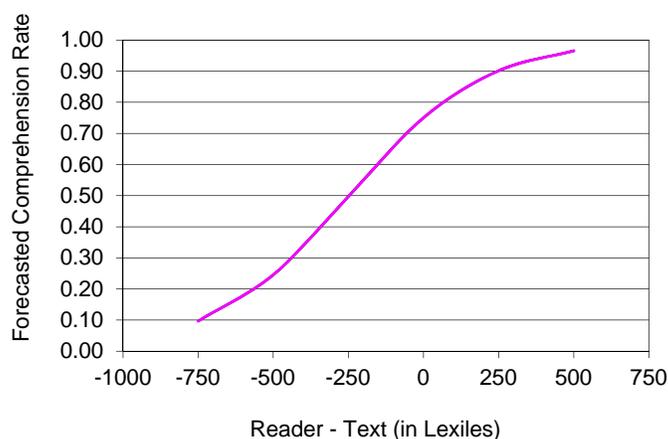
$$\text{Rate} = \frac{e^{\text{ELD}+1.1}}{1 + e^{\text{ELD}+1.1}} \quad (\text{Equation 5})$$

where ELD is the “effective logit difference” given by

$$\text{ELD} = (\text{Reader Lexile measure} - \text{Text Lexile measure}) \div 225. \quad (\text{Equation 6})$$

Figure 16 shows the general relationship between reader-text discrepancy and forecasted comprehension rate. When the reader measure and the text calibration are the same (difference of 0L) then the forecasted comprehension rate is 75 percent. In the example in the preceding paragraph, the difference between the reader measure of 600L and the text calibration of 350L is 250L. Referring to Figure 16 and using +250L (reader minus text), the forecasted comprehension rate for this reader-text combination would be 90 percent.

Figure 16. Relationship between reader-text discrepancy and forecasted comprehension rate.



Tables 18 and 19 show comprehension rates calculated for various combinations of reader measures and text calibrations.

Table 18. Comprehension rates for the same individual with materials of varying comprehension difficulty.

Person Measure	Text Calibration	Sample Titles	Forecast Comprehension
1000	500	<i>Tornado</i> (Byars)	96%
1000	750	<i>The Martian Chronicles</i> (Bradbury)	90%
1000	1000	<i>Reader's Digest</i>	75%
1000	1250	<i>The Call of the Wild</i> (London)	50%
1000	1500	<i>On the Equality Among Mankind</i> (Rousseau)	25%

Table 19. Comprehension rates of different person abilities with the same material.

Person Measure	Calibration for a Grade 10 Biology Textbook	Forecast Comprehension Rate
500	1000	25%
750	1000	50%
1000	1000	75%
1250	1000	90%
1500	1000	96%

The subjective experience of 50-percent, 75-percent, and 90-percent comprehension as reported by readers varies greatly. A 1000L reader reading 1000L text (75-percent comprehension) reports confidence and competence. Teachers listening to such a reader report that the reader can sustain the meaning thread of the text and can read with motivation and appropriate emotion and emphasis. In short, such readers appear to comprehend what they are reading. A 1000L reader reading 1250L text (50-percent comprehension) encounters so much unfamiliar vocabulary and difficult syntactic structures that the meaning thread is frequently lost. Such readers report frustration and seldom choose to read independently at this level of comprehension. Finally, a 1000L reader reading 750L text (90-percent comprehension) reports total control of the text, reads with speed, and experiences automaticity during the reading process.

The primary utility of the Lexile Framework is its ability to forecast what happens when readers confront text. With every application by teacher, student, librarian, or parent there is a test of the Framework’s accuracy. The Framework makes a point prediction every time a text is chosen for a reader. Anecdotal evidence suggests that the Lexile Framework predicts as intended. That is not to say that there is an absence of error in forecasted comprehension. There is error in text measures, reader measures, and their difference modeled as forecasted comprehension. However, the error is sufficiently small that the judgments about readers, texts, and comprehension rates are useful.

Relationship between Linking Error and Forecasted Comprehension Rate. Using Equation 5 with different combinations of reader measure and text difficulty, the effect of linking error on forecasted comprehension rate can be examined. *Table 20* shows the changes in the forecasted comprehension rate for different combinations of reader and text interactions. When the linking error is small, 5–10L, then the effect on forecasted comprehension rate is a minimal difference (1 to 2 percent) increase or decrease in comprehension.

Table 20. Effect of reader-text discrepancy on forecasted comprehension rate.

Reader Lexile Measure	Text Lexile Measure	Difference	Forecasted Comprehension Rate
1000L	970L	30L	77.4%
1000L	975L	25L	77.0%
1000L	980L	20L	76.7%
1000L	985L	15L	76.3%
1000L	990L	10L	75.8%
1000L	995L	5L	75.4%
1000L	1000L	0L	75.0%
1000L	1005L	-5L	74.6%
1000L	1010L	-10L	74.2%
1000L	1015L	-15L	73.8%
1000L	1020L	-20L	73.3%
1000L	1025L	-25L	72.9%
1000L	1030L	-30L	72.4%

Conclusions, Caveats, and Recommendations

Forging a link between scales is a way to add value to one scale without having to administer an additional test. Value can be in the form of any or all of the following:

- increased *interpretability* (e.g., “Based on this test score, what can my child actually read?”),
- increased *diagnostic capability* (e.g., “Based on this test score, what are the student’s weaknesses?”), or
- increased *instructional use* (e.g., “Based on these test scores, I need to modify my instruction to include these skills.”).

The link that has been established between the NC READY EOG Reading/EOC English II scale scores and the Lexile measures permits readers to be matched with books and texts that provide an appropriate level of challenge while avoiding frustration. The result of this purposeful match may be that students will read more, and, thereby read better. The real power of the Lexile Framework is in examining the growth of readers—wherever the reader may be in the development of his or her reading skills. Readers can be matched with texts that they are forecasted to read with 75-percent comprehension. As a reader grows, he or she can be matched with more demanding texts. And, as the texts become more demanding, then the reader grows.

Recommendations about reporting Lexile measures for readers. Lexile measures are reported as a number followed by a capital “L” for “Lexile.” There is no space between the measure and the “L,” and measures of 1,000 or greater are reported without a comma (e.g., 1050L). All Lexile measures should be rounded to the nearest 5L to avoid over interpretation of the measures. As with any test score, uncertainty in the form of measurement error is present.

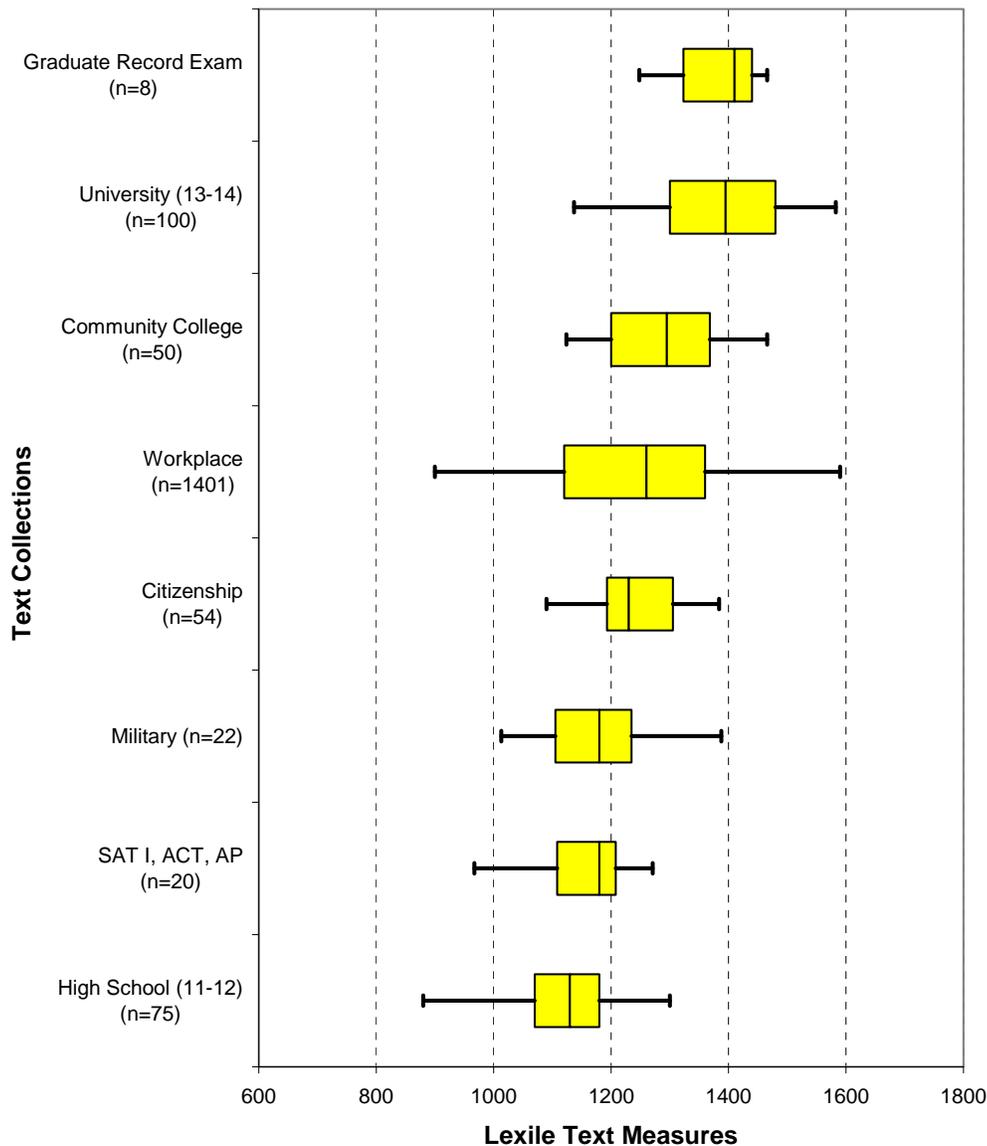
Lexile measures that are reported for an individual student should reflect the purpose for which they will be used. If the purpose is research (e.g., to measure growth at the student, grade, school, district, or state level), then actual measures should be used at all score points, rounded to the nearest integer. A computed Lexile measure of 772.51 would be reported as 773L. If the purpose is instructional, then the Lexile measures should be capped at the upper bound of measurement error (e.g., at the 95th percentile of the national Lexile norms) to ensure developmental appropriateness of the material. MetaMetrics expresses these as “Reported Lexile Measures” and recommends that these measures be reported on individual score reports. In instructional environments where the purpose of the Lexile measure is to appropriately match readers with texts, all scores below 0L should be reported as “BRxxxL.” No student should receive a negative Lexile measure on a score report. The lowest reported value below 0L is BR400L.

Some assessments report a Lexile range for each student, which is 50L above and 100L below the student's actual Lexile measure. This range represents the boundaries between the easiest kind of reading material for the student and the level at which the student will be more challenged, yet can still read successfully.

Text Complexity. There is increasing recognition of the importance of bridging the gap that exists between K-12 and higher education and other postsecondary endeavors. Many state and policy leaders have formed task forces and policy committees such as P-20 councils.

In the *Journal of Advanced Academics* (Summer 2008), Williamson investigated the gap between high school textbooks and various reading materials across several postsecondary domains. As can be seen in *Figure 17*, the resources Williamson used were organized into four domains that correspond to the three major postsecondary endeavors that students can choose – further education, the workplace, or the military – and the broad area of citizenship, which cuts across all postsecondary endeavors. Williamson discovered a substantial increase in reading expectations and text complexity from high school to postsecondary domains – a gap large enough to help account for high remediation rates and disheartening graduation statistics (Smith, 2011).

Figure 17. A continuum of text difficulty for the transition from high school to postsecondary experiences (box plot percentiles: 5th, 25th, 50th, 75th, and 95th).¹



Expanding on Williamson’s work, Stenner, Sanford-Moore, and Williamson (2012) aggregated the readability information across the various postsecondary options available to a high school graduate to arrive at a standard of reading needed by individuals to be considered “college and career ready.” In their study, they included additional citizenship materials beyond those examined by Williamson (e.g., national and international newspapers and other adult reading materials such as Wikipedia articles). Using a weighted mean of the medians for each of the postsecondary options

¹ Reprinted from Williamson, G. L. (2008). A text readability continuum for postsecondary readiness. *Journal of Advanced Academics*, 19(4), 602-632.

(education, military, work place, and citizenship), a measure of 1300L was defined as the general reading demand for postsecondary options and could be used to judge a student's "college and career readiness."

In Texas, two studies were conducted to examine the reading demands in various postsecondary options – technical college, community college, and 4-year university programs. Under Commissioner Raymond Paredes, THECB conducted a research study in 2007 (and extended in 2008) which addressed the focal question of "how well does a student need to read to be successful in community colleges, technical colleges, and universities in Texas?" THECB staff collected a sample of books that first year students in Texas would be required to read in each setting. These books were measured in terms of their text complexity using The Lexile Framework for Reading. Since the TAKS had already been linked with Lexile measures for several years, the THECB study was able to overlay the TAKS cut scores onto the post high school reading requirements. (For a complete description of this report, please visit www.thecb.state.tx.us/index.cfm?objectid=31BFFF6B-BB41-8A43-C76A99EDA0F38B7D.)

Since the THECB study was completed, other states have followed the Texas example and used the same approach in examining the gap from high school to the postsecondary world. In 2009, a similar study was conducted for the Georgia Department of Education; and in 2010, a study was conducted for the Tennessee Department of Education. In terms of mean text demand, the results across the three states produced similar estimates of the reading ability needed in higher-education institutions: Texas, 1230L; Georgia, 1220L; and Tennessee, 1260L. When these results are incorporated with the reading demands of other postsecondary endeavors (military, citizenship, workplace, and adult reading materials [national and international newspapers] and Wikipedia articles) used by Stenner, Koons, and Swartz (2010), the college and career readiness standard for reading is 1293L. These results are based on more than 105,000,000 words from approximately 3,100 sources from the adult text space.

The question for educators becomes how to determine if a student is "on track" for college and career as previously defined in the Common Core State Standards and described above. "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" (Smith, 2012). Implementing the Common Core will require districts and schools to develop new instructional strategies and complementary resources that are not only aligned with these national college- and career-readiness standards, but also utilize and incorporate proven and cost-effective tools that are universally accessible to all stakeholders.

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).

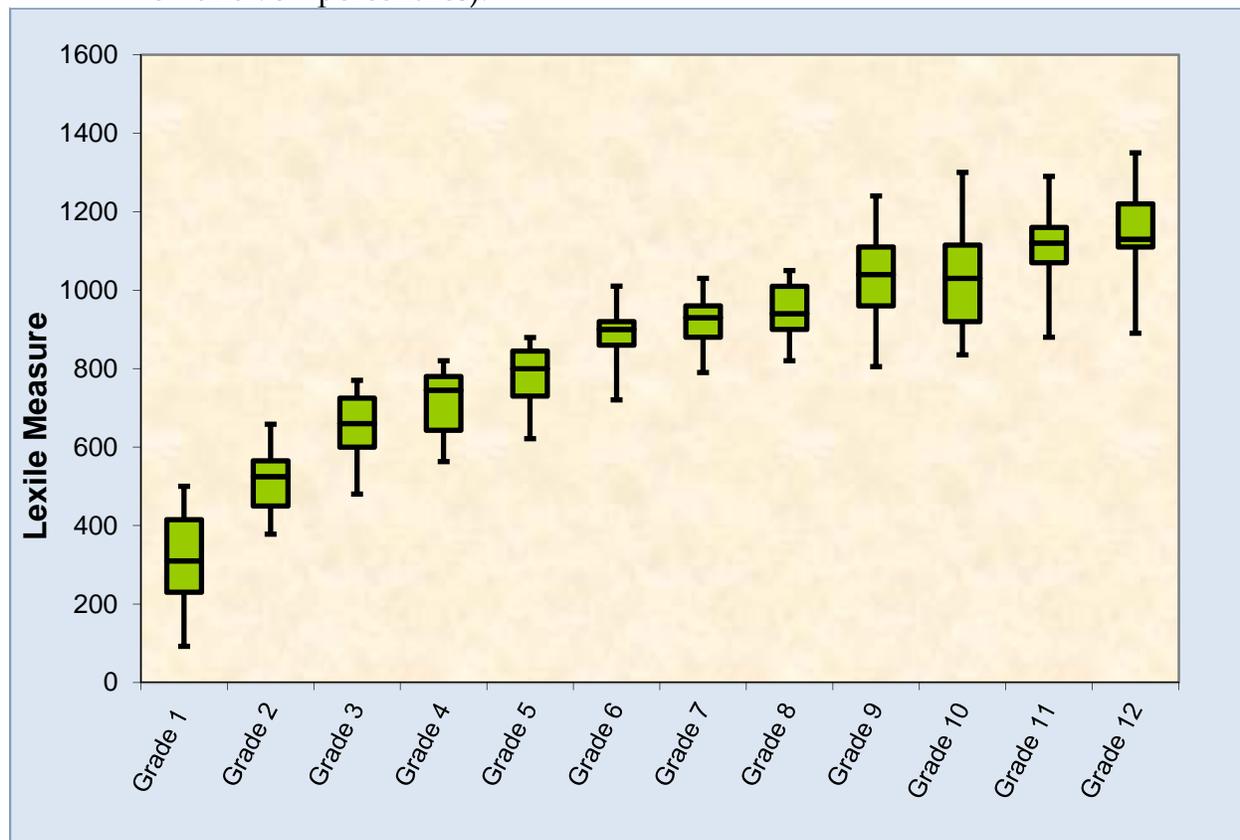
The Common Core State Standards recommends a three-part model for evaluating the complexity of a text that takes into account its qualitative dimensions, quantitative measure, and reader and task considerations. It describes text complexity as “the inherent difficulty of reading and comprehending a text 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 simpler terms, *text complexity is a transaction between text, reader, and task*. The quantitative aspect of defining text complexity consists of a stair-step progression of increasingly difficult text by grade levels (Common Core State Standards for English Language Arts, Appendix A, NGA Center and CCSSO, 2010, p. 8).

Table 21. Lexile ranges aligned to college- and career-readiness expectations, by grade.

Grade	2012 “Stretch” Text Measure
1	190L to 530L
2	420L to 650L
3	520L to 820L
4	740L to 940L
5	830L to 1010L
6	925L to 1070L
7	970L to 1120L
8	1010L to 1185L
9	1050L to 1260L
10	1080L to 1335L
11-12	1185L to 1385L

Between 2004 and 2008, MetaMetrics (Williamson, Koons, Sandvik, and Sanford-Moore, 2012) collected and measured textbooks across the K-12 educational continuum. The box-and-whisker plot in *Figure 4* shows the Lexile measures (*y*-axis) across grades as defined in the US. For each grade, the box refers to the interquartile range. The line within the box indicates the median. The end of each whisker shows the 5th and 95th percentile text complexity measures in the Lexile metric for each grade. This information can provide a basis for defining at what level students need to be able to read to be ready for various postsecondary endeavors such as further education beyond high school and entering the work force.

Figure 18. Text complexity distributions, in Lexile units, by grade (whiskers represent 5th and 95th percentiles).



This continuum can be “stretched” to describe the reading demands expected of students in Grades 1-12 who are “on track” for college and career (Sanford-Moore and Williamson, 2012). The quantitative aspect of defining text complexity consists of a stair-step progression of increasingly difficult text by grade levels (Common Core State Standards for English Language Arts, Appendix A, NGA Center and CCSSO, 2010, p. 8).

MetaMetrics’ research on the typical reading demands of college and careers contributed to the Common Core State Standards as a whole and, more specifically, to the Lexile-based grade bands in Figure 19. Figure 19 shows the relationship between the “Level 3” performance standard for each grade level established on the NC READY EOG Reading/EOC English II Assessment and the “stretch” reading demands. This shows that the NC READY EOG Reading/EOC English II performance standards for “Level 3” at each grade level is set at a level that is consistent with being “on track” for college and career readiness at the end of Grade 12.

Figure 19. Comparison of NC READY EOG Reading/EOC English II “Level 3” standards with college and career reading levels described by the CCSS.

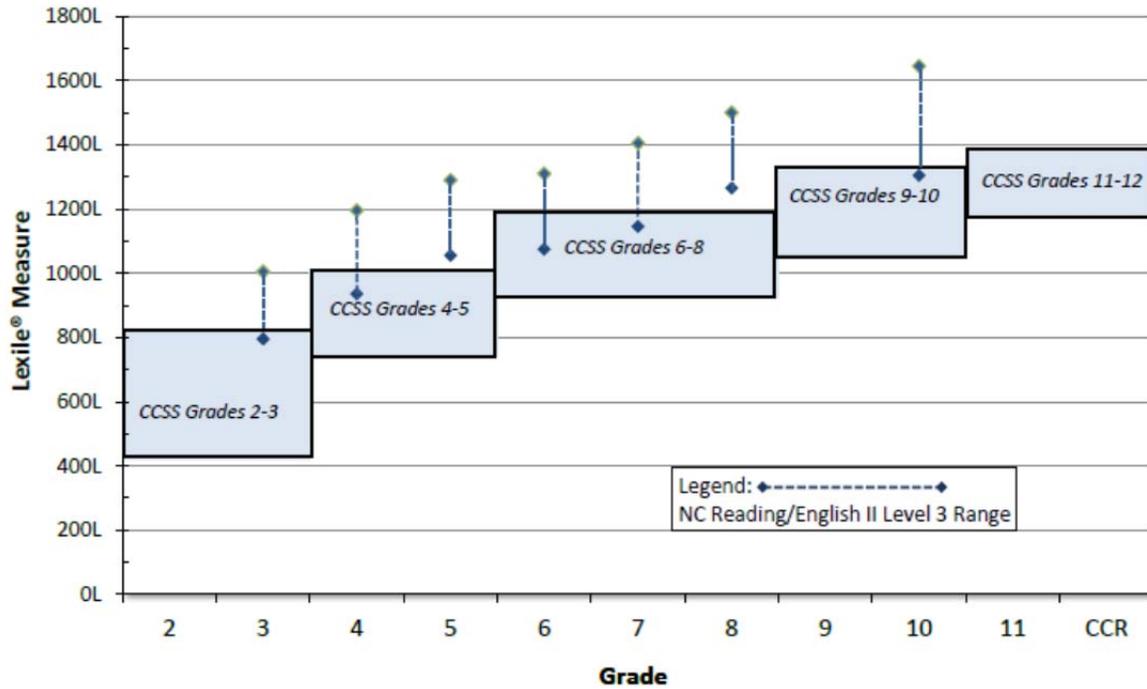
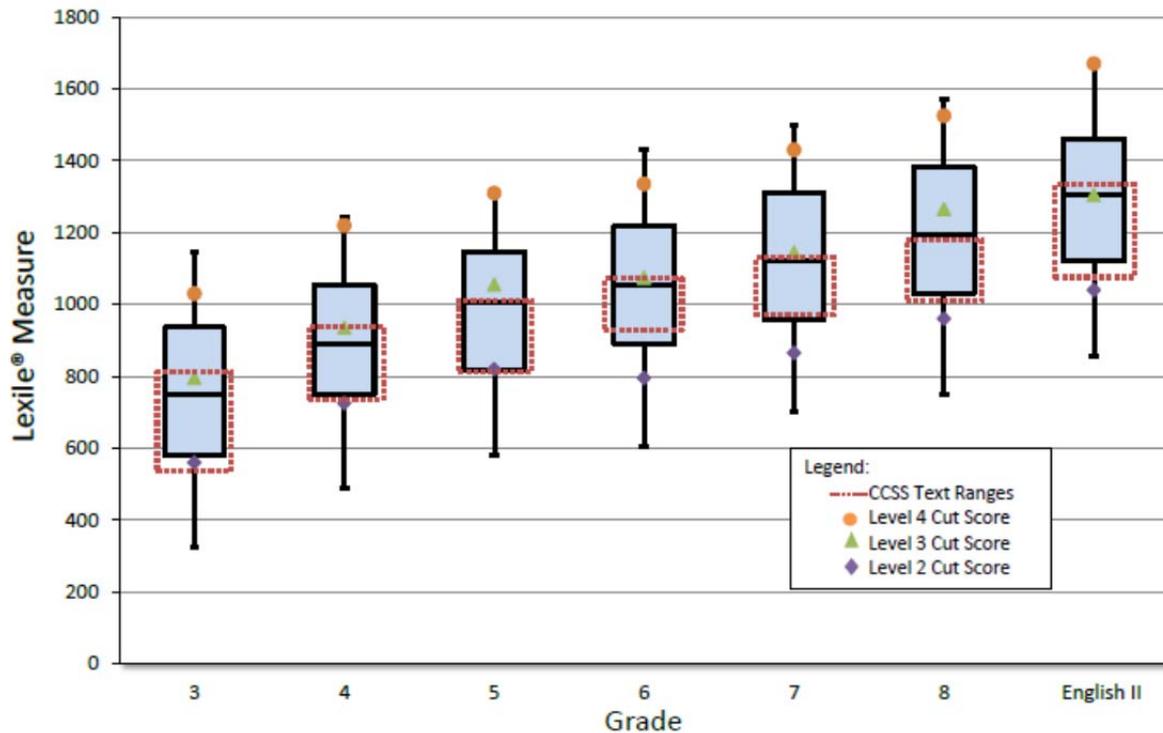


Figure 20 shows that the spring 2013 student performance on the NC READY EOG Reading/EOC English II assessments at each grade level is “on track” for college and career readiness. Students can be matched with reading materials that are at or above the recommendations in Appendix A of the CCSS for ELA for each grade level.

Figure 10. NC READY EOG Reading/EOC English II 2012-2013 student performance expressed as Lexile measures.



In 2008, MetaMetrics and the North Carolina Department of Public Instruction conducted a study to link the NCEOG Reading Test with the Lexile scale (MetaMetrics, 2008). The minimum score considered “proficient” (Level 3) at each grade level on the NCEOG Reading is presented in *Table 22*. In 2013, NCDPI transitioned their assessment program to the NC READY EOG Reading Assessment to align with the Common Core State Standards in English/Language Arts and to describe student reading performance in relation to college and career readiness. One outcome of this change was to set the performance standards for NC READY EOG Reading at a higher level. For comparison purposes, the minimum “proficient” score for the NC READY EOG Reading assessment is also repeated from *Table 17*. The Lexile scale can be used as an external “yardstick” to evaluate this change in reading demand on the North Carolina reading assessment. The information in *Table 22* shows that the NC READY EOG Reading standards are demanding more of students in terms of reading ability in 2013.

Table 22. Minimum “Level 3” Lexile measure on NCEOG Reading (2008) and NC READY EOG Reading (2013).

Grade	“Proficient” Level 3 Cut Score (2008)	“Proficient” Level 3 Cut Score (2013)
3	665L	795L
4	790L	935L
5	940L	1055L
6	990L	1075L
7	1115L	1145L
8	1165L	1265L

Next Steps. To utilize the results from this study, Lexile measures need to be incorporated into the NC READY EOG Reading/EOC English II results processing and interpretation frameworks. This information can then be used in a variety of areas within the educational system – instruction, assessment, communication to name a few.

Within the *instructional area*, suggested book lists can be developed for ranges of readers. Care must be taken to ensure that the books on the lists are also developmentally appropriate for the readers. The Lexile measure is one factor related to comprehension and is a good starting point in the selection process of a book for a specific reader. Other factors such as student developmental level, motivation, and interest; amount of background knowledge possessed by the reader; and characteristics of the text such as illustrations and formatting also need to be considered when matching a book with a reader.

In this era of student-level accountability and high-stakes assessment, differentiated instruction – the attempt “on the part of classroom teachers to meet students where they are in the learning process and move them along as quickly and as far as possible in the context of a mixed-ability classroom” (Tomlinson, 1999) – is a means for all educators to help students succeed. Differentiated instruction promotes high-level and powerful curriculum for all students, but varies the level of teacher support, task complexity, pacing, and avenues to learning based on student readiness, interest, and learning profile. One strategy for managing a differentiated classroom suggested by Tomlinson is the use of multiple texts and supplementary materials.

The Lexile Framework is an objective tool that can be used to determine a student’s readiness for a reading experience; the Lexile Framework “targets” text (books, newspapers, periodicals) for readers at a 75-percent comprehension level – a level that is challenging, but not frustrating (Schnick and Knickelbine, 2000).

Within the *communication* area, Lexile measures can be used to communicate with students, parents, teachers, educators, and the community by providing a common language to use to talk about reading growth and development. By aligning all areas of the educational system, parents can be included in the instructional process. With a variety of data related to a student’s reading level a more complete picture can be formed and more informed decisions can be made concerning reading-group placement, amount of extra instruction needed, and promotion/retention decisions.

It is much easier to understand what a national percentile rank of 50 means when it is tied to the reading demands of book titles that are familiar to adults. Parents are encouraged to help their children achieve high standards by expecting their children to succeed at school, communicating with their children’s teachers and the school, and helping their children keep pace and do homework.

Through the customized reading lists and electronic database of titles, parents can assist their children in the selection of reading materials that are at the appropriate level of challenge and monitor the reading process at home. A link can be provided to the “Find a Book” website. This site provides a quick, free resource to battle “summer slide” – the learning losses that students often experience during the summer months when they are not in school. Lexile measures make it easy to help students read and learn all summer long and during the school year. This website can help build a reading list of books at a young person’s reading level that are about subjects that interest him or her. This website can be viewed at <http://www.lexile.com/findabook/>.

In one large school district, the end-of-year testing results are sent home to parents in a folder. The folder consists of a Lexile Map on one side and a letter from the superintendent on the other side. The school district considers this type of material as “refrigerator-friendly.” They encourage parents to put the Lexile Map on the refrigerator and use it to monitor and track the reading progress of their child throughout the school year.

The community-at-large (business leaders, citizens, politicians, and visitors) sees the educational system as a reflection of the community. Through the reporting of assessment results (after all, that is what the community is most interested in – results), people can understand what the community values and see the return for its investment in the schools and its children.

One way to involve the community is to work with the public libraries and local bookstores when developing reading lists. The organizations should be contacted early enough so that they can be sure that the books will be available. Often books can be displayed with their Lexile measures for easy access.

Many school districts make presentations to civic groups to educate the community as to their reading initiatives and how the Lexile Framework is being utilized in the school. Conversely, many civic groups are looking for an activity to sponsor, and it could be as simple as “donate-a-book” or “sponsor-a-reader” campaigns.

Notes

1. A T-parallel test is a test that is designed to be “theoretically parallel” to another test in that it has the same number of items/points, the same overall level of difficulty in terms of raw score means and standard deviations, and assesses the same construct domain (MetaMetrics, Inc. 1998).

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bormuth, J.R. (1966). Readability: New approach. *Reading Research Quarterly*, 7, 79-132.
- Carroll, J.B., Davies, P., & Richman, B. (1971). *Word frequency book*. Boston: Houghton Mifflin.
- Carver, R.P. (1974). Measuring the primary effect of reading: Reading storage technique, understanding judgments and cloze. *Journal of Reading Behavior*, 6, 249-274.
- Chall, J.S. (1988). "The beginning years." In B.L. Zakaluk and S.J. Samuels (Eds.), *Readability: Its past, present, and future*. Newark, DE: International Reading Association.
- Crain, S. & Shankweiler, D. (1988). "Syntactic complexity and reading acquisition." In A. Davidson and G.M. Green (Eds.), *Linguistic complexity and text comprehension: Readability issues reconsidered*. Hillsdale, NJ: Erlbaum Associates.
- Davidson, A. & Kantor, R.N. (1982). On the failure of readability formulas to define readable text: A case study from adaptations. *Reading Research Quarterly*, 17, 187-209.
- Dunn, L.M. & Dunn, L.M. (1981). *Manual for Forms L and M of the Peabody Picture Vocabulary Test – Revised*. Circle Pines, MN: American Guidance Service.
- Dunn, L.M. & Markwardt, F.C. (1970). *Peabody Individual Achievement Test*. Circle Pines, MN: American Guidance Service.
- Efron, B. (1981). Nonparametric estimates of the standard error: The Jackknife, the Bootstrap, and other resampling techniques. *Biometrika*. 68, 589-599.
- Gorsuch, R.L. (1983). *Factor analysis*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Grolier, Inc. (1986). *The electronic encyclopedia*. Danbury, CT: Author.
- Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Bulletin*. 9, 139-164.

- Klare, G.R. (1963). *The measurement of readability*. Ames, IA: Iowa State University Press.
- Kolen, M.J. & Brennan, R.L. (2004). *Test equating, scaling, and linking: Methods and practices*. 2nd edition. New York: Springer Science + Business Media, LLC.
- Liberman, I.Y., Mann, V.A., Shankweiler, D., & Westelman, M. (1982). Children's memory for recurring linguistic and non-linguistic material in relation to reading ability. *Cortex*, 18, 367-375.
- Linacre, J.M. (2011). WINSTEPS (Version 3.73) [Computer Program]. Chicago: Author.
- Lord, F.M. (1980). *Applications of item response theory to practical testing problems*. New York: Erlbaum Associates.
- McGraw-Hill Book Company. (1983). *Guidelines for bias-free publishing*. Monterey, CA: Author.
- MetaMetrics, Inc. (2008). *Linking the North Carolina EOG Reading and EOC English I Tests with the Lexile Framework*. Durham, NC: Author.
- Miller, G.A. & Gildea, P.M. (1987). How children learn words. *Scientific American*, 257, 94-99.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects, Appendix A*. Washington, DC: Author.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2012). *Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects, Revised Appendix A*. Washington, DC: Author.
- National Research Council. (1999). *Uncommon measures: Equivalence and linkage among educational tests*. Washington, D.C.: National Academy Press.
- North Carolina Department of Public Education. (2013a). *Achievement Level Descriptors for North Carolina End-of-Course Tests*. Retrieved on November 6, 2013 from <http://sbepolicy.dpi.state.nc.us/Policies/GCS-C-036.asp?Acr=GCS&Cat=C&Pol=036>

- North Carolina Department of Public Education. (2013b). *Achievement Level Descriptors for North Carolina End-of-Grade Tests*. Retrieved on November 6, 2013 from <http://sbepolicy.dpi.state.nc.us/Policies/GCS-C-033.asp?Acr=GCS&Cat=C&Pol=033>
- North Carolina Department of Instruction. (2013c). *Common Core State Standards (CCSS) for English Language Arts: North Carolina Assessment Specifications Summary, READY EOG Assessments, Grades 3-8 READY EOC English II Assessments*. Retrieved on October 31, 2013 from <http://www.ncpublicschools.org/docs/acre/assessment/ela.pdf>
- North Carolina Department of Instruction. (2013d). *North Carolina READY End-of-Course Assessments*. Retrieved on October 24, 2013 from <http://www.ncpublicschools.org/docs/accountability/policyoperations/assessbriefs/assessbriefeoc13.pdf>
- North Carolina Department of Instruction. (2013e). *North Carolina READY End-of-Grade Assessments*. Retrieved on October 24, 2013 from <http://www.ncpublicschools.org/docs/accountability/policyoperations/assessbriefs/assessbriefeog13.pdf>
- Petersen, N.S., Kolen, M.J., & Hoover, H.D. (1989). "Scaling, Norming, and Equating." In R.L. Linn (Ed.), *Educational Measurement* (Third Edition) (pp. 221-262). New York: American Council on Education and Macmillan Publishing Company.
- Poznanski, J.B. (1990). A meta-analytic approach to the estimation of item difficulties. Unpublished doctoral dissertation, Duke University, Durham, NC.
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics*, 4, 207-230.
- Sanford-Moore, E., & Williamson, G. L. (2012). *Bending the text complexity curve to close the gap* (MetaMetrics Research Brief). Durham, NC: MetaMetrics, Inc.
- Schnick, T. & Knickelbine, M. (2000). *The Lexile Framework: An introduction for educators*. Durham, NC: MetaMetrics, Inc.
- Shankweiler, D. & Crain, S. (1986). Language mechanisms and reading disorder: A modular approach. *Cognition*, 14, 139-168.
- Sitter, R.R. (1992). Comparing three bootstrap methods for survey data. *The Canadian Journal of Statistics*, 20(2), 135-154.

- Smith, M. (2011, March 30). *Bending the reading growth trajectory: Instructional strategies to promote reading skills and close the readiness gap*. MetaMetrics Policy Brief. Durham, NC: MetaMetrics, Inc.
- Smith, M. (2012, February). *Not so common: Comparing Lexile® measures with the standards' other text complexity tools*. MetaMetrics White Paper. Durham, NC: MetaMetrics, Inc.
- Stenner, A.J. (1990). Objectivity: Specific and general. *Rasch Measurement Transactions*, 4, 111.
- Stenner, A. J., Sanford-Moore, E., & Williamson, G. L. (2012). *The Lexile® Framework for Reading quantifies the reading ability needed for "College & Career Readiness."* MetaMetrics Research Brief. Durham, NC: MetaMetrics, Inc.
- Stenner, A.J., Smith, M., & Burdick, D.S. (1983). Toward a theory of construct definition. *Journal of Educational Measurement*, 20(4), 305-315.
- Stenner, A.J., Smith, D.R., Horabin, I., & Smith, M. (1987a). Fit of the Lexile Theory to item difficulties on fourteen standardized reading comprehension tests. Durham, NC: MetaMetrics, Inc.
- Stenner, A.J., Smith, D.R., Horabin, I., & Smith, M. (1987b). Fit of the Lexile Theory to sequenced units from eleven basal series. Durham, NC: MetaMetrics, Inc.
- Tomlinson, C.A. (1999). *The differentiated classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Williamson, G. L. (2008). A text readability continuum for postsecondary readiness. *Journal of Advanced Academics*, 19(4), 602-632.
- Williamson, G. L. (2011, March 16-17). *Growth and growth norms: Standards for academic growth*. Presentation at the first semi-annual meeting of MetaMetrics' Technical Advisory Committee, held at MetaMetrics, Inc., Durham, NC.
- 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.
- Williamson, G. L., Thompson, C. L., and Baker, R. F. (2007, April). *North Carolina's growth in reading and mathematics*. AERA Distinguished Paper Presentation at the 2007 annual meeting of the American Educational Research Association (AERA), Chicago, IL.

Wright, B.D. & Linacre, J.M. (1994, August). *The Rasch model as a foundation for the Lexile Framework*. Unpublished manuscript.

Wright, B.D. & Stone, M.H. (1979). *Best Test Design*. Chicago: MESA Press.



THE LEXILE[®] FRAMEWORK FOR READING MAP

Matching Readers with Text

Imagine getting students excited about reading while also improving their reading abilities. With the Lexile[®] Map, students have a chance to match books with their reading levels, and celebrate as they are able to read increasingly complex texts!

Let your students find books that fit them! Build custom book lists for your students by accessing our “Find a Book” tool at Lexile.com/fab.

HOW IT WORKS

The Lexile[®] Map provides examples of popular books and sample texts that are matched to various points on the Lexile[®] scale, from 200L for emergent reader text to 1600L for more advanced texts. The examples on the map help to define text complexity and help readers identify books of various levels of text complexity. Both literature and informational texts are presented on the Lexile Map.

HOW TO USE IT

Lexile reader and text measures can be used together to forecast how well a reader will likely comprehend a text at a specific Lexile level. A Lexile reader measure is usually obtained by having the reader take a reading comprehension test. Numerous tests report Lexile reader measures including many state end-of-year assessments, national norm-referenced assessments, and reading program assessments. A Lexile reader measure places students on the same Lexile scale as the texts. This scale ranges from below 200L to above 1600L. The Lexile website

also provides a way to estimate a reader measure by using information about the reader’s grade level and self-reported reading ability.

Individuals reading within their Lexile ranges (100L below to 50L above their Lexile reader measures) are likely to comprehend approximately 75 percent of the text when reading independently. This “targeted reading” rate is the point at which a reader will comprehend enough to understand the text but will also face some reading challenge. The result is growth in reading ability and a rewarding reading experience.

For more guidance concerning targeting readers with books, visit www.Lexile.com/fab to access the “Find a Book” tool. “Find a Book” enables users to search from over 150,000 books to build custom reading lists based on Lexile range and personal interests and to check the availability of books at the local library.





1500L+ ▶

1500L **Don Quixote**** CERVANTES

The Words were to me so many Pearls of Eloquence, and his Voice sweeter to my Ears than Sugar to the Taste. The Reflection on the Misfortune which these Verses brought on me, has often made me applaud Plato's Design of banishing all Poets from a good and well governed Commonwealth, especially those who write wantonly or lasciviously. For, instead of composing lamentable Verses, like those of the Marquiss of Mantua, that make Women and Children cry by the Fireside, they try their utmost Skill on such soft Strokes as enter the Soul, and wound it, like that Thunder which hurts and consumes all within, yet leaves the Garment sound. Another Time he entertained me with the following Song.



SAMPLE TITLES

LITERATURE	1640L	The Plot Against America (ROTH)
	1560L	Rob Roy (SCOTT)
	1530L	The Good Earth (BUCK)
	1520L	A Fable (FAULKNER)
	1500L	The Decameron (BOCCACCIO)
INFORMATIONAL	1600L	Sustaining Life: How Human Health Depends on Biodiversity (CHIVIAN & BERNSTEIN)
	1550L	The Art of War (SUN TZU)
	1560L	The United States' Constitution
	1520L	Fair Play: The Ethics of Sport (SIMON)
	1500L	Critique of Pure Reason (KANT)

1400L ▶ 1495L

1400L **Nathaniel's Nutmeg** MILTON

Setting sail once again they kept a sharp look-out for Busse Island, discovered thirty years previously by Martin Frobisher, but the rolling sea mists had grown too thick. Storms and gale—force winds plagued them for days on end and at one point grew so ferocious that the foremast cracked, splintered and was hurled into the sea. It was with considerable relief that the crew sighted through the mist the coast of Newfoundland—a vague geographical term in Hudson's day—at the beginning of July. They dropped anchor in Penobscot Bay, some one hundred miles west of Nova Scotia.



SAMPLE TITLES

LITERATURE	1460L	The Legend of Sleepy Hollow (IRVING)
	1450L	Billy Budd** (MELVILLE)
	1430L	The Story of King Arthur and His Knights (PYLE)
	1420L	Life All Around Me by Ellen Foster (GIBBONS)
	1420L	The Scarlet Letter** (HAWTHORNE)
INFORMATIONAL	1490L	America's Constitution: A Biography** (AMAR)
	1490L	Gettysburg Address (LINCOLN)
	1480L	The Declaration of Independence
	1410L	Profiles in Courage (KENNEDY)
	1400L	The Life and Times of Frederick Douglass (DOUGLASS)

1300L ▶ 1395L

1300L **1776: America and Britain at War**** MCCULLOUGH

But from this point on, the citizen-soldiers of Washington's army were no longer to be fighting only for the defense of their country, or for their rightful liberties as freeborn Englishmen, as they had at Lexington and Concord, Bunker Hill and through the long siege at Boston. It was now a proudly proclaimed, all-out war for an independent America, a new America, and thus a new day of freedom and equality. At his home in Newport, Nathanael Greene's mentor, the Reverend Ezra Stiles, wrote in his diary almost in disbelief: Thus the Congress has tied a Gordian knot, which the Parl [iament] will find they can neither cut, nor untie. The thirteen united colonies now rise into an Independent Republic among the kingdoms, states, and empires on earth...And have I lived to see such an important and astonishing revolution?



SAMPLE TITLES

LITERATURE	1360L	Robinson Crusoe (DEFOE)
	1350L	The Secret Sharer (CONRAD)
	1340L	The Hunchback of Notre Dame (HUGO)
	1340L	The Metamorphosis** (KAFKA)
	1340L	Fever Pitch (HORNBY)
INFORMATIONAL	1390L	In Defense of Food: An Eater's Manifesto (POLLAN)
	1380L	Politics and the English Language** (ORWELL)
	1370L	Jane Austen's Pride and Prejudice (BLOOM)
	1340L	Walden** (THOREAU)
	1300L	Arctic Dreams: Imagination and Desire in a Northern Landscape (LOPEZ)



1200L ▶ 1295L

1200L *Why We Can't Wait* KING

We sing the freedom songs today for the same reason the slaves sang them, because we too are in bondage and the songs add hope to our determination that "We shall overcome, Black and white together, We shall overcome someday." I have stood in a meeting with hundreds of youngsters and joined in while they sang "Ain't Gonna Let Nobody Turn Me 'Round." It is not just a song; it is a resolve. A few minutes later, I have seen those same youngsters refuse to turn around from the onrush of a police We sing the freedom songs today for the same reason the slaves sang them, because we too are in bondage and the songs add hope to our determination that "We shall overcome, Black and white together, We shall overcome someday."



SAMPLE TITLES

- LITERATURE
 - 1280L *The House of the Spirits* (ALLENDE)
 - 1270L *Tarzan of the Apes* (BURROUGHS)
 - 1270L *Chronicle of a Death Foretold* (GARCÍA MÁRQUEZ)
 - 1220L *Annie John* (KINCAID)
 - 1210L *The Namesake*** (LAHIRI)
- INFORMATIONAL
 - 1290L *A Brief History of Time* (HAWKING)
 - 1280L *Black, Blue, and Gray: African Americans in the Civil War*** (HASKINS)
 - 1240L *Blood Done Sign My Name* (TYSON)
 - 1230L *Stiff: The Curious Lives of Human Cadavers* (ROACH)
 - 1200L *The Dark Game: True Spy Stories* (JANECZKO)

1100L ▶ 1195L

1100L *Pride and Prejudice*** AUSTEN

Lydia was a stout, well-grown girl of fifteen, with a fine complexion and good-humoured countenance; a favourite with her mother, whose affection had brought her into public at an early age. She had high animal spirits, and a sort of natural self-consequence, which the attentions of the officers, to whom her uncle's good dinners and her own easy manners recommended her, had increased into assurance. She was very equal therefore to address Mr. Bingley on the subject of the ball, and abruptly reminded him of his promise; adding, that it would be the most shameful thing in the world if he did not keep it. His answer to this sudden attack was delightful to their mother's ear.



SAMPLE TITLES

- LITERATURE
 - 1180L *The Curious Incident of the Dog in the Night-time* (HADDON)
 - 1170L *The Amazing Adventures of Kavalier & Clay* (CHABON)
 - 1150L *A Wizard of Earthsea* (LE GUIN)
 - 1130L *All the King's Men* (WARREN)
 - 1110L *A Separate Peace* (KNOWLES)
- INFORMATIONAL
 - 1160L *The Longitude Prize*** (DASH)
 - 1160L *In Search of Our Mothers' Gardens* (WALKER)
 - 1140L *Winterdance: The Fine Madness of Running the Iditarod* (PAULSEN)
 - 1130L *The Great Fire*** (MURPHY)
 - 1100L *Vincent Van Gogh: Portrait of an Artist*** (GREENBERG & JORDAN)

1000L ▶ 1095L

1000L *Mythbusters Science Fair Book* MARGLES

There may be less bacteria on the food that's picked up quickly, but playing it safe is the best idea. If it hits the floor, the next thing it should hit is the trash. If putting together petri dishes and dealing with incubation seems like a bigger project than you're ready to take on, there's a simpler way to observe bacterial growth. Practically all you need is some bread and your own two hands. Cut the edges off each slice of bread so that they'll fit into the plastic containers. Put one slice of bread into each container. Measure one tablespoon of water and splash it into the first piece of bread. Put the lid on the container and use your pen and tape to label this your control.



SAMPLE TITLES

- LITERATURE
 - 1080L *I Heard the Owl Call My Name* (CRAVEN)
 - 1070L *Savvy* (LAW)
 - 1070L *Around the World in 80 Days* (VERNE)
 - 1010L *The Pearl* (STEINBECK)
 - 1000L *The Hobbit or There and Back Again* (TOLKIEN)
- INFORMATIONAL
 - 1070L *Geeks: How Two Lost Boys Rode the Internet Out of Idaho*** (KATZ)
 - 1030L *Phineas Gage* (FLEISCHMAN)
 - 1020L *This Land Was Made for You and Me: The Life and Songs of Woody Guthrie* (PARTRIDGE)
 - 1010L *Travels With Charley: In Search of America*** (STEINBECK)
 - 1000L *Claudette Colvin: Twice Toward Justice* (HOOSE)

**Common Core State Standards Text Exemplar



900L ▶ 995L

900L ***We are the Ship: The Story of Negro League Baseball*** NELSON

Rube ran his ball club like it was a major league team. Most Negro teams back then weren't very well organized. Didn't always have enough equipment or even matching uniforms. Most times they went from game to game scattered among different cars, or sometimes they'd even have to "hobo"—which means hitch a ride on the back of someone's truck to get to the next town for a game. But not Rube's team. They were always well equipped, with clean, new uniforms, bats, and balls. They rode to the games in fancy Pullman cars Rube rented and hitched to the back of the train. It was something to see that group of Negroes stepping out of the train, dressed in suits and hats. They were big-leaguers.



SAMPLE TITLES

- LITERATURE
 - 980L **Dovey Coe** (DOWELL)
 - 950L **Bud, Not Buddy** (CURTIS)
 - 940L **Harry Potter and the Chamber of Secrets** (ROWLING)
 - 940L **Heat** (LUPICA)
 - 900L **City of Fire** (YEP)
- INFORMATIONAL
 - 990L **Seabiscuit** (HILLENBRAND)
 - 970L **The Kid's Guide to Money: Earning It, Saving It, Spending It, Growing It, Sharing It**** (OTFINOSKI)
 - 950L **Jim Thorpe, Original All-American** (BRUCHAC)
 - 930L **Colin Powell A & E Biography** (FINLAYSON)
 - 920L **Talking with Artists** (CUMMINGS)

800L ▶ 895L

800L ***Moon Over Manifest*** VANDERPOOL

There wasn't much left in the tree fort from previous dwellers. Just an old hammer and a few rusted tin cans holding some even rustier nails. A couple of wood crates with the salt girl holding her umbrella painted on top. And a shabby plaque dangling sideways on one nail, FORT TREECONDEROGA. Probably named after the famous fort from Revolutionary War days. Anything else that might have been left behind had probably been weathered to bits and fallen through the cracks. No matter. I'd have this place whipped into shape lickety-split. First off, I picked out the straightest nail I could find and fixed that sign up right. Fort Treeconderoga was open for business.



SAMPLE TITLES

- LITERATURE
 - GN840L* **The Odyssey** (HINDS)
 - 830L **Baseball in April and Other Stories** (SOTO)
 - 820L **Maniac Magee** (SPINELLI)
 - 820L **Where the Mountain Meets the Moon**** (LIN)
 - 800L **Homeless Bird** (WHELEN)
- INFORMATIONAL
 - 880L **The Circuit** (JIMENEZ)
 - 870L **The 7 Habits of Highly Effective Teens** (COVEY)
 - IG860L* **Animals Nobody Loves** (SEYMOUR)
 - 860L **Through My Eyes: Ruby Bridges** (BRIDGES)
 - 830L **Quest for the Tree Kangaroo: An Expedition to the Cloud Forest of New Guinea**** (MONTGOMERY)

700L ▶ 795L

700L ***The Miraculous Journey of Edward Tulane*** DICAMILLO

Edward, for lack of anything better to do, began to think. He thought about the stars. He remembered what they looked like from his bedroom window. What made them shine so brightly, he wondered, and were they still shining somewhere even though he could not see them? Never in my life, he thought, have I been farther away from the stars than I am now. He considered, too, the fate of the beautiful princess who had become a warthog. Why had she become a warthog? Because the ugly witch turned her into one—that was why. And then the rabbit thought about Pellegrina. He felt, in some way that he could not explain to himself, that she was responsible for what had happened to him. It was almost as if it was she, and not the boys, who had thrown Edward overboard.



SAMPLE TITLES

- LITERATURE
 - 770L **Walk Two Moons** (CREECH)
 - 760L **Hoot** (HIAASEN)
 - 750L **Esperanza Rising** (RYAN)
 - 720L **Nancy's Mysterious Letter** (KEENE)
- INFORMATIONAL
 - GN720L* **Sherlock Holmes and the Adventure at the Copper Beeches** (DOYLE)
 - 790L **Be Water, My Friend: The Early Years of Bruce Lee** (MOCHIZUKI)
 - 760L **Stay: The True Story of Ten Dogs** (MUNTEAN)
 - IG760L* **Mapping Shipwrecks with Coordinate Planes** (WALL)
 - 720L **Pretty in Print: Questioning Magazines** (BOTZAKIS)
 - 720L **Spiders in the Hairdo: Modern Urban Legends** (HOLT & MOONEY)

*GN denotes Graphic Novel, IG denotes Illustrated Guide
**Common Core State Standards Text Exemplar



600L ▶ 695L

600L ***You're on Your Way, Teddy Roosevelt*** ST. GEORGE & FAULKNER

But from his first workout in Wood's Gymnasium he had been determined to control his asthma and illnesses rather than letting his asthma and illnesses control him. And he had. On that hot summer day in August he had proved to himself—and everyone else—that he had taken charge of his own life. In 1876 Teedie—now known as Teddy—entered Harvard College. He was on his own ...without Papa. That was all right. "I am to do everything for myself," he wrote in his diary. Why not? He was stronger and in better health than he had ever been. And ready and eager for the adventures and opportunities that lay ahead.



SAMPLE TITLES

LITERATURE	680L	Charlotte's Web (WHITE)
	660L	Holes (SACHAR)
	620L	M.C. Higgins, the Great** (HAMILTON)
	610L	Mountain Bike Mania (CHRISTOPHER)
INFORMATIONAL	610L	A Year Down Yonder (PECK)
	690L	Where Do Polar Bears Live?*** (THOMSON)
	680L	An Eye for Color: The Story of Josef Albers (WING)
	660L	Remember: The Journey to School Integration (MORRISON)
	660L	From Seed to Plant*** (GIBBONS)
630L	Sadako and the Thousand Paper Cranes (COERR)	

500L ▶ 595L

500L ***A Germ's Journey*** ROOKE

Excuse me! Let's blow out of this place! In real life, germs are very small. They can't be seen without a microscope. Rudy forgot to use a tissue. His cold germs fly across the room at more than 100 miles an hour. Whee! I can fly! Best ride ever! A few germs land on Ernie. But skin acts like a suit of armor. It protects against harm. The germs won't find a new home there. Healthy skin keeps germs out. But germs can sneak into the body through cuts, scrapes, or cracks in the skin. Most germs enter through a person's mouth or nose. Rudy's germs continue to fall on nearly everything in the room—including Brenda's candy.



SAMPLE TITLES

LITERATURE	560L	Sarah, Plain and Tall (MACLACHLAN)
	530L	It's All Greek to Me (SCIESZKA)
	520L	John Henry: An American Legend (KEATS)
	500L	Judy Moody Saves the World (MCDONALD)
	500L	The Curse of the Cheese Pyramid (STILTON)
INFORMATIONAL	IG590L*	Claude Monet (CONNOLLY)
	560L	Lemons and Lemonade: A Book about Supply and Demand (LOEWEN)
	560L	Molly the Pony (KASTER)
	530L	Langston Hughes: Great American Poet (MCKISSACK)
	510L	A Picture for Marc (KIMMEL)

400L ▶ 495L

400L ***How Not to Babysit Your Brother*** HAPKA

I continued to search. I checked under Steve's bed. Then I checked under my bed. I searched the basement, the garage, and my closet. There was no sign of Steve. This was going to be harder than I thought. Where was Steve hiding? CRASH! Uh-oh, I thought. I heard Buster barking in the kitchen. I ran to see what was going on. When I got there, the dog food bin was tipped over. Steve's head and shoulders were sticking out of the top. Dog food was stuck in his hair, on his clothes, and up his nose. He looked like an alien from the planet Yuck. He giggled as Buster licked some crumbs off his ear.



SAMPLE TITLES

LITERATURE	460L	Chrysanthemum (HENKES)
	410L	The Enormous Crocodile (DAHL)
	GN400L*	Pilot And Huxley (MCGUINNESS)
	400L	The Fire Cat*** (AVERILL)
INFORMATIONAL	400L	Cowgirl Kate and Cocoa*** (SILVERMAN)
	480L	Martin Luther King, Jr. and the March on Washington** (RUFFIN)
	460L	True Life Treasure Hunts (DONNELLY)
	460L	Half You Heard of Fractions? (ADAMSON)
	420L	Rally for Recycling (BULLARD)
400L	Animals in Winter (RUSTAD)	

*GN denotes Graphic Novel, IG denotes Illustrated Guide
**Common Core State Standards Text Exemplar



300L ▶ 395L

300L *Princess Posey and the Next-Door Dog* GREENE

"We have to stop now," said Miss Lee. "It's time for reading." "Ohhh..." A disappointed sound went up around the circle. "Here's what we'll do." Miss Lee stood up. "You are all very interested in dogs. So this week, you can write a story about your own dog or pet. Then you can read it to the class." Everyone got excited again. Except Posey. She didn't have a pet. Not a dog. Not a cat. Not a hamster. "Those of you who don't have a pet," Miss Lee said, "can write about the pet you hope to own someday." Miss Lee had saved the day! Now Posey had something to write about, too. Posey told her mom about Luca's puppy on the way home.



SAMPLE TITLES

LITERATURE	380L	<i>Martha Bakes a Cake</i> (BARSS)
	380L	<i>Junie B. Jones is (Almost) a Flower Girl</i> (PARK)
	360L	<i>Poppleton in Winter**</i> (RYLANT)
	340L	<i>Never Swipe a Bully's Bear</i> (APPLEGATE)
INFORMATIONAL	330L	<i>Frog and Toad Together**</i> (LOBEL)
	GN380L*	<i>BMX Blitz</i> (CIENCIN)
	380L	<i>Lemonade for Sale</i> (MURPHY)
	350L	<i>A Snowy Day</i> (SCHAEFER)
	330L	<i>Freedom River</i> (RAPPAPORT)
	300L	<i>From Tree to Paper</i> (MARSHALL)

200L ▶ 295L

200L *Ronald Morgan Goes to Bat* GIFF

He smacked the ball with the bat. The ball flew across the field. "Good;" said Mr. Spano. "Great, Slugger!" I yelled. "We'll win every game. It was my turn next. I put on the helmet, and stood at home plate. "Ronald Morgan," said Rosemary. "You're holding the wrong end of the bat." Quickly I turned it around. I clutched it close to the end. Whoosh went the first ball. Whoosh went the second one. Wham went the third. It hit me in the knee. "Are you all right?" asked Michael. But I heard Tom say, "I knew it. Ronald Morgan's the worst." At snack time, we told Miss Tyler about the team.



SAMPLE TITLES

LITERATURE	280L	<i>Hi! Fly Guy**</i> (ARNOLD)
	260L	<i>The Cat in the Hat</i> (SEUSS)
	GN240L*	<i>Lunch Lady and the Cyborg Substitute</i> (KROSOCZKA)
	200L	<i>Dixie</i> (GILMAN)
INFORMATIONAL	200L	<i>The Best Bug Parade</i> (MURPHY)
	290L	<i>The Story of Pocahontas</i> (JENNER)
	250L	<i>Math in the Kitchen</i> (AMATO)
	230L	<i>What makes Day and Night</i> (BRANLEY)
	220L	<i>I Love Trains!</i> (STURGES)
210L	<i>Sharks!</i> (CLARKE)	

*GN denotes Graphic Novel

**Common Core State Standards Text Exemplar

Please note:

The Lexile measure (text complexity) of a book is an excellent starting point for a student's book selection. It's important to understand that the book's Lexile measure should not be the only factor in a student's book selection process. Lexile measures do not consider factors such as age-appropriateness, interest, and prior knowledge. These are also key factors when matching children and adolescents with books they might like and are able to read.



Lexile codes provide more information about developmental appropriateness, reading difficulty, and common or intended usage of books. For more information on Lexile codes, please visit Lexile.com.

LEXILE TEXT RANGES TO GUIDE READING FOR COLLEGE AND CAREER READINESS

GRADES	CCSS LEXILE TEXT RANGE
11-12	1185L-1385L
9-10	1050L-1335L
6-8	925L-1185L
4-5	740L-1010L
2-3	420L-820L
1	190L-530L

COMMON CORE STATE STANDARDS FOR ENGLISH LANGUAGE ARTS, APPENDIX A (ADDITIONAL INFORMATION), NGA AND CCSSO, 2012

METAMETRICS®, the METAMETRICS® logo and tagline, LEXILE®, LEXILE® FRAMEWORK and the LEXILE® logo are trademarks of MetaMetrics, Inc., and are registered in the United States and abroad. Copyright © 2013 MetaMetrics, Inc. All rights reserved.