Linking the NC Ready EOG Reading/EOC English II with The Lexile® Framework for Reading

November 2013
Updated September 2021

Prepared by MetaMetrics for:

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

MetaMetrics. *Linking Assessment with Instruction*

LEXILE® QUANTILE®

www.MetaMetricsInc.com | www.Lexile.com | www.Quantiles.com

MetaMetrics • 1000 Park Forty Plaza Drive, Suite 120 • Durham, North Carolina 27713

Preface

Updating the NC EOG Reading and NC EOC English II Reading Scale link with the Lexile Framework for Reading

Maintenance of the focal scale (i.e., NC EOG Reading/EOC English II reading scale) is critical to the validity of any link with an auxiliary scale (i.e., The Lexile scale). If an update occurs to the focal scale, the integrity of the link should be re-evaluated and additional linking studies may be needed to accommodate fundamental changes to the focal scale. Such updates may include, but are not limited to, incorporating new item types into the assessment; revising item calibrations; or revising the assessment program and the reported scale scores.

The North Carolina Standard Course of Study (NCSCOS) defines the appropriate content standards for each grade or proficiency level and each high school course to provide a uniform set of learning standards for every public school in North Carolina. These standards define what students are expected to know and be able to do by the end of each school year or course. The NC State Board of Education policy, SCOS-012, requires that each content area's standards be reviewed every five-to-seven years to ensure the NCSCOS consists of clear, relevant standards and objectives. In 2010, the NC State Board of Education adopted the Common Core State Standards (National Governors Association & CCSSO, 2010). In 2017, the state Board of Education adopted revisions to the Standard Course of Study in math (K-8) and English Language Arts that best aligned with the appropriate content for career and college readiness, with required implementation for schools in 2018-2019 (NCDPI, 2021b).

With the implementation of the North Carolina Extended Content Standards for English Language Arts (NCDPI, 2021a), MetaMetrics and the North Carolina Department of Public Instruction sought to link NC EOG Reading and NC EOC English II scales with the Lexile scale and determine the feasibility of a new field-based linking study. Based on this evaluation, the decision was made to perform an editions linking procedure which leveraged pre-equated item parameters from the 2019 NC Ready EOG Reading/EOC English II to establish a link with edition 5 NC EOG Reading and NC EOC English II. Details on the methodology and results of the study are outlined in the Appendix A.

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, is 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.

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.

Linking the NC Ready EOG Reading/EOC English II with The Lexile® Framework for Reading

Technical Report

Prepared by MetaMetrics for the North Carolina Department of Public Instruction (https://www.dpi.nc.gov/) under Contract No. NC10025818 dated December 17, 2012; and Request for Quote #: 40- PR11813253 dated .

MetaMetrics
800 Taylor Street, Suite 102
Durham, North Carolina 27701
https://lexile.com/

November 2013 – Updated August 2021

Table of Contents

List of Tables

List of Figures

Introduction

Often it is desirable to convey more information about test performance than can be gleaned from a raw score or percentage correct. When items from an assessment are linked to the Lexile scale, the linkage can be used to provide context for understanding the results of the assessment. It is often hard to explain what a student can read based on the results of a reading test. Students may ask, "Based on my test results, what can I read and how well?" 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 can read.

Auxiliary score scales can be used to "convey additional normative information, test-content information, and information that is jointly normative and content based (Petersen, Kolen, and Hoover, 1989, p. 222). One such auxiliary scale is The Lexile® Framework for Reading, which was developed to appropriately match students with text at a level that provides challenge but not frustration.

Linking assessment results with the Lexile Framework for Reading 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 regarding test results. In addition, the Lexile Framework for Reading provides a common way to monitor if students are "on track" for the reading demands of various postsecondary endeavors. By using the Lexile Framework for Reading, 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. By linking NC READY EOG Reading/EOC English II scores with the Lexile Framework for Reading, educators and parents will be able to answer these questions. In addition, they will be better able to use the Lexile reading measures produced from the NC READY EOG Reading/EOC English II assessments 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 to students so that they can be matched with text based on their NC READY EOG Reading/EOC English II test scores. The study was conducted by MetaMetrics for the North Carolina Department of Public Instruction (https://www.dpi.nc.gov/) under Contract No. NC10025818 dated December 17, 2012.

The primary purposes of this study were to:

- link the NC READY EOG Reading/EOC English II scales to the Lexile Framework for Reading;

- develop a correspondence table for converting NC READY EOG Reading/EOC English II scale scores to Lexile reading measures;
- present a solution for matching students with text;
- provide tools (e.g., Lexile Find A Book) and information that can be used to answer questions related to standards, test score interpretation, and test validation; and
- produce a report that describes the linking analysis procedures.

The Lexile Framework for Reading

A reader's comprehension of text is dependent on many factors—the purpose for reading, the ability of the reader, and the text being read. The reader can be asked to read a text for many purposes including entertainment (literary experience), to gain information, or to perform a task. Each reader brings to the reading experience a variety of important factors: reading ability, prior knowledge, interest level, and developmental readiness. For any text, there are three factors associated with the readability of the text: complexity, support, and quality. All of these reader and text factors are important considerations when evaluating the appropriateness of a text for a reader. The Lexile Framework for Reading focuses primarily on two features: reader ability and reading text complexity.

The Lexile Framework for Reading measures for both texts and readers typically range from above 200L to below 1600L, but measures can range from below 0L for beginning reader materials (e.g., BR150L) or above 1600L for advanced materials. Within any single classroom, there will be a range of reading materials to reflect the student range of reading ability and interest in different topics and types of text.

Reading Text Complexity

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 text is dominated by the familiarity of the semantic units and by the complexity of the syntactic structures used in constructing the text. The Lexile Framework for Reading utilizes these two dominant features of language to measure reading text complexity by examining the characteristics of word frequency and sentence length. In addition, when measuring early reader texts, the Lexile Framework for Reading utilizes characteristics found to be important to the complexity of early reader text such as word decodabilty, patterning, and repetition.

Variables that Impact the Text Complexity of Upper Level Text

Semantic Component. Most operationalizations of the semantic component 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 for Reading currently employs a 1.4 billion-word corpus when examining the semantic component of text. This corpus was assembled from the more than 90,000 texts that were measured by MetaMetrics for publishers from 1998 through 2012.

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 for Reading.

Variables that Impact the Text Complexity of Early Reader Texts

Texts designed for early readers are distinct from texts designed for more accomplished readers because they are usually designed specifically to facilitate reading development. For all readers, making meaning of the texts is always the focus, but for early readers, developing an understanding of how to "crack the code" requires specific attention. Early readers must develop the ability to hear sounds in words, develop sight words, and acquire word recognition strategies (Fitzgerald and Shanahan, 2000) as they develop the comprehension and fluency characteristic of more advanced readers. A number of studies support the finding that the presence of specific text features support the development of skills associated with code cracking. For example, word repetition reinforces sight-word learning and development of the sounds associated with spelling patterns (e.g., Vadasy, Sanders, & Peyton, 2005). Repeated phrases also reinforce scaffolding development of a variety of word recognition strategies (e.g., Ehri & McCormick, 1998). The use of words familiar in oral language enhances readers' ability to make meaning from words and permits more attention to word recognition (e.g., Muter, Hulme, Snowling, & Stevenson, 2004). Inclusion of several types of text-characteristic support may further support students' growth as readers. Research suggests that to appropriately describe early reader reading text complexity it is necessary to consider several text characteristics at multiple linguistic levels (Graesser & McNamara, 2011; Graesser, McNamara, & Kulikowich, 2011; Kintsch, 1998; and Snow, 2002). In general, levels of text characteristics include word level (e.g., word structure, word frequency), within-sentence level (e.g., syntax), and across-sentence/discourse level (e.g., referential cohesion). The research base supporting the importance of multiple levels of text characteristics for early phases of learning to read is extensive (Mesmer, Cunningham, & Hiebert, 2012) and has identified the importance of considering the impact of interaction between the features (Merlini Barbaresi, 2003; and Biber, 1988).

In order to determine which text characteristics had the greatest impact on reading text complexity for early readers, MetaMetrics identified 22 unique text characteristics at four linguistic levels: sounds-in-words, words (structure and meaning), within-sentence syntax, and across-sentence/discourse.

- *Sounds-in-Words*—number of phonemes in words, phonemic Levenshtein Distance, and mean internal phonemic predictability
- *Word Structure*—decoding demand, orthographic Levenshtein Distance, number of syllables in words, and mean internal orthographic predictability
- *Word Meaning*—age of acquisition, abstractness, and word rareness
- *Within-Sentence Syntax*—sentence length and grammar
- *Across-Sentence/Discourse*—linear edit distance, linear word overlap, cohesion triggers, type-token ratio, longest common string, edit distance, Cartesian word overlap, information load, and compression ratio

From these characteristics, 238 operationalizations were developed to capture the varied ways in which the characteristics could be quantified in terms of their presence in the text. Three hundred and fifty early reader texts designed for readers in Kindergarten through Grade 2 were selected to represent the range of text types early readers are likely to encounter. These included decodable books, phonics readers, leveled books, high-frequency readers, and various trade books. Two separate sub-studies were conducted to determine the relative challenge of the texts. One study collected primary-grade educators' ratings of the complexity of the 350 texts and the other gathered Grade 1 and 2 students' responses to a subset of 89 texts from the full set of 350 study texts. From these studies a text-complexity logit scale was created so that each text could be assigned a measure (Fitzgerald, Elmore, Koons, Hiebert, Bowen, Sanford-Moore & Stenner, 2016).

Calibration of Text Difficulty of Upper Level Texts

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 eight other standardized tests. The resulting correlation between the observed logit difficulties and the theoretical calibrations across the nine 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 result of the research was a regression equation linking the syntactic and semantic features of text to the difficulty of text. This equation can now be used to calibrate test items and text within the Lexile Framework for Reading.

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 $1/100^{th}$ of the difference between freezing (0 degrees) and boiling (100 degrees) water. For the Lexile Scale, the unit size (a Lexile) was defined as $1/1000^{th}$ 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 $1/1000^{th}$ 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:

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

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 Framework for Reading Map, a graphic

representation of the Lexile Scale from 200L to 1500L+, provides a context for understanding reading comprehension (see Appendix A).

Calibration of Text Difficulty of Early Reader Texts

To bring the observed difficulties (logit scores) of early reader texts from the two studies previously described (Fitzgerald, Elmore, Koons, Hiebert, Bowen, Sanford-Moore & Stenner, 2016) onto the Lexile scale, a theory-based linking procedure was conducted. First, Lexile text measures were calculated based only on the syntactic and semantic features of the text as done with upper level texts. Next, for approximately 10% of the texts the discrepancy between the observed difficulty and the theoretical Lexile reading measure was large, so the texts were flagged and not used in subsequent analyses. Finally, using the remaining 90% of the texts in the study, a linear linking function (SD line) was calculated. In linear linking, a transformation is chosen such that scores on two sets of data are considered to be linked if they correspond to the same number of standard deviations above (or below) the mean in some group of data elements (Angoff, 1984, cited in Petersen, Kolen, and Hoover, 1989; Kolen and Brennan, 2014).
The result of the linear linking function was that the early reader observed difficulties were transformed to Lexile text measures while still maintaining the relative ordering of the difficulty of the texts derived from the educator judgments and student performances.

Once observed Lexile reading measures were calculated, a random forest regression technique was employed to evaluate the importance of the 238 operationalizations of characteristics that research suggests affect reading text complexity of early reader texts. This process was conducted in several stages and is described in detail by Fitzgerald and Elmore and their colleagues (2015). The first step in the analysis was to set baseline performance. Eighty percent of the texts were selected for this training process and 20% were held as a validation sample. Three separate random forest regressions were conducted, one each for: (1) the 80% of the 350 texts that the teachers ordered (n = 279); (2) the 80% of the texts that the students were presented (n = 71), and (3) the two sets of texts combined (N = 350). Each random forest regression produced importance values for each of the 238 variables in relation to the text-complexity logit scale.

The next step in the analysis involved an iterative variable-selection procedure in which the variables with the smallest importance values were systematically removed and the effect on the model re-calculated. This process determined whether fewer variables could predict reading text complexity as well or nearly as well as the 238-variable model. The result was a set of nine variables:

- Word level variables—monosyllable decoding, syllable count, age of acquisition, word rareness, and abstractness
- Within-sentence and across-sentence/discourse level variables—intersentential complexity, phrase diversity, non-compressibility, and text density

Lastly, a final set of three random forest regression models was trained using the nine variables with the teacher text set, the student text set, and the two text sets combined. The resulting

correlations for the teacher, student, and combined models were 0.89, 0.71, and 0.88, respectively. The validation samples, 20% of the teacher texts (n = 71) and 20% of the student texts (n = 19), were combined and a final random forest regression was run with the nine selected variables as predictors. The model was validated with a correlation of 0.85 and RMSE of 9.68. The final model can now be used to calibrate texts intended for early-readers.

The nine variables have been grouped into four Early Reading Indicators based on the linguist level addressed:

- Decoding Demand (Decoding)—syllable count and monosyllable decoding demand
- Semantic Demand (Vocabulary)—abstractness, word rareness, and age of acquisition
- Syntactic Demand (Sentences)—intersentential complexity
- Structure Demand (Patterns)—non-compressibility, phrase diversity, and text density

The Enhanced Lexile Text Analyzer®

When text is analyzed by MetaMetrics, all electronic files are initially edited according to established guidelines used with the enhanced Lexile Text Analyzer software. These guidelines include the removal of all incomplete sentences, chapter titles, and paragraph headings; and running of a spell check. The text is then submitted to the enhanced Lexile Text Analyzer that examines the lengths of the sentences and the frequencies of the words for upper-level texts and the nine early-reader variables for lower-level texts. The enhanced Lexile Text Analyzer first looks at the text features of a piece of text and attempts to determine if the text is written for early readers (early reader texts) or for more advanced readers (upper level texts). Based on the results of the examination, the enhanced Lexile Text Analyzer applies the most appropriate word and sentence/discourse variables to the measurement process. The enhanced Lexile Text Analyzer then reports a Lexile reading measure for the text. If the measure of the text is 650L or below, the four Early Reading Indicators are also reported.

Validity of The Lexile Framework for Reading

The 2014 *Standards for Educational and Psychological Testing* (America Educational Research Association, American Psychological Association, and National Council on Measurement in Education) states that "validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests" (p. 11). In applying this definition to the Lexile Framework for Reading, the question that should be asked is "What evidence supports the use of the Lexile Framework for Reading to describe reading text complexity and reader ability?" Because the Lexile Framework for Reading addresses reading comprehension, an important aspect of validity evidence that should be brought to bear is evidence showing that the construct being addressed is indeed, reading comprehension. This type of validity evidence has traditionally been called construct validity. One source of construct validity evidence for the Lexile Framework for Reading can be evaluated by examining how well Lexile reading measures relate to other measures of reading ability and reading comprehension.

The Lexile Framework for Reading and other Measures of Reading Comprehension. The Lexile Framework for Reading has been linked to numerous standardized tests of reading comprehension. When assessment scales are linked, a common frame of reference can be used to interpret the test results. This frame of reference can be "used to convey additional normative information, test-content information, and information that is jointly normative and content-based. For many test uses, … [this frame of reference] conveys information that is more crucial than the information conveyed by the primary score scale" (Petersen, Kolen, and Hoover, 1989, p. 222). Linking the Lexile Framework for Reading with other measures of reading comprehension produces a common frame of reference: the Lexile reading measure.

Table 1 presents the results from linking studies conducted with the Lexile Framework for Reading. In these studies, students were administered a Lexile reading assessment and another assessment of reading comprehension. There is a strong relationship between reading comprehension ability as measured by the Lexile Framework for Reading and reading comprehension ability as measured by other assessments. For each of the tests listed, student reading comprehension scores can also be reported as Lexile reading measures. This dual reporting provides a rich, criterion-related frame of reference for interpreting the standardized test scores. When a student takes one of the standardized tests, in addition to receiving his norm-referenced test information, the student can receive a reading list consisting of texts (books and articles) targeted to his or her specific reading level.

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, and 10	4,644	0.90
Metropolitan Achievement Test (8th ed.)	2, 4, 6, and 8	2,713	0.92
The Iowa Assessments (Iowa Tests of Basic Skills and Iowa Tests of Educational Development)	3, 5, 7, 9, and 11	4,146	0.91
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	0.87
Oklahoma Core Competency Tests (OCCT)	3 – 8	8,437	0.81 to 0.86*
Wyoming Performance Assessment for Wyoming Students (PAWS)	3, 5, and 8 11	2,293 442	0.91 0.84
Arizona Instrument to Measure Progress (AIMS)	3, 5, 7, and 10	5,599	0.89
Comprehensive Testing Program (CPT 4 – ERB)	2, 4, 6, and 8	644	0.88
TOEFL iBT	NA	2,867	0.65
TOEIC	NA	2,770	0.74
Kentucky Performance Rating for Educational Progress (K-PREP)	3 – 8	6,480	0.71 to 0.79*
North Carolina ACT	11	2,675	0.84
North Carolina READY End-of-Grades/End-of-Course Tests (NC READY EOG/EOC)	3, 5, 7, and 8 E2	7,709 2,068	0.92 0.89
Georgia Milestones EOG/EOC Assessments	3 – 9, and AME	12,415	0.82 to 0.86*
State of Texas Assessments of Academic Readiness (STAAR™)	3 – 8 English I English II	5,856 620 1,063	0.86 0.87 0.87
ACT Aspire	3, 5, 7, and EHS	1,264	0.85
South Carolina READY Reading	3 – 8	10,951	0.94
ISIP Early Reading test Advanced Reading test	1 – 3 4, 6, and 8	5,471 6,479	0.87 0.65

*Notes: * Tests were not vertically scaled; separate linking equations were derived for each grade/course.*

The Lexile Framework for Reading and the Difficulty of Basal Readers. Lexile reading measures are organized in a sequential manner, so a lower Lexile reading measure for a text indicates that the text is less complex than text with a higher Lexile reading measure. Validity evidence for the internal structure (the sequential structure) of the Lexile Framework for Reading was obtained through a study that examined the relationship of basal reader sequencing to Lexile reading measures. In a study conducted by Stenner, Smith, Horabin, and Smith (1987b) Lexile reading 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 reading calibration of each unit in each series. After correction for range restriction and measurement error, the average disattenuated correlation between the Lexile reading 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

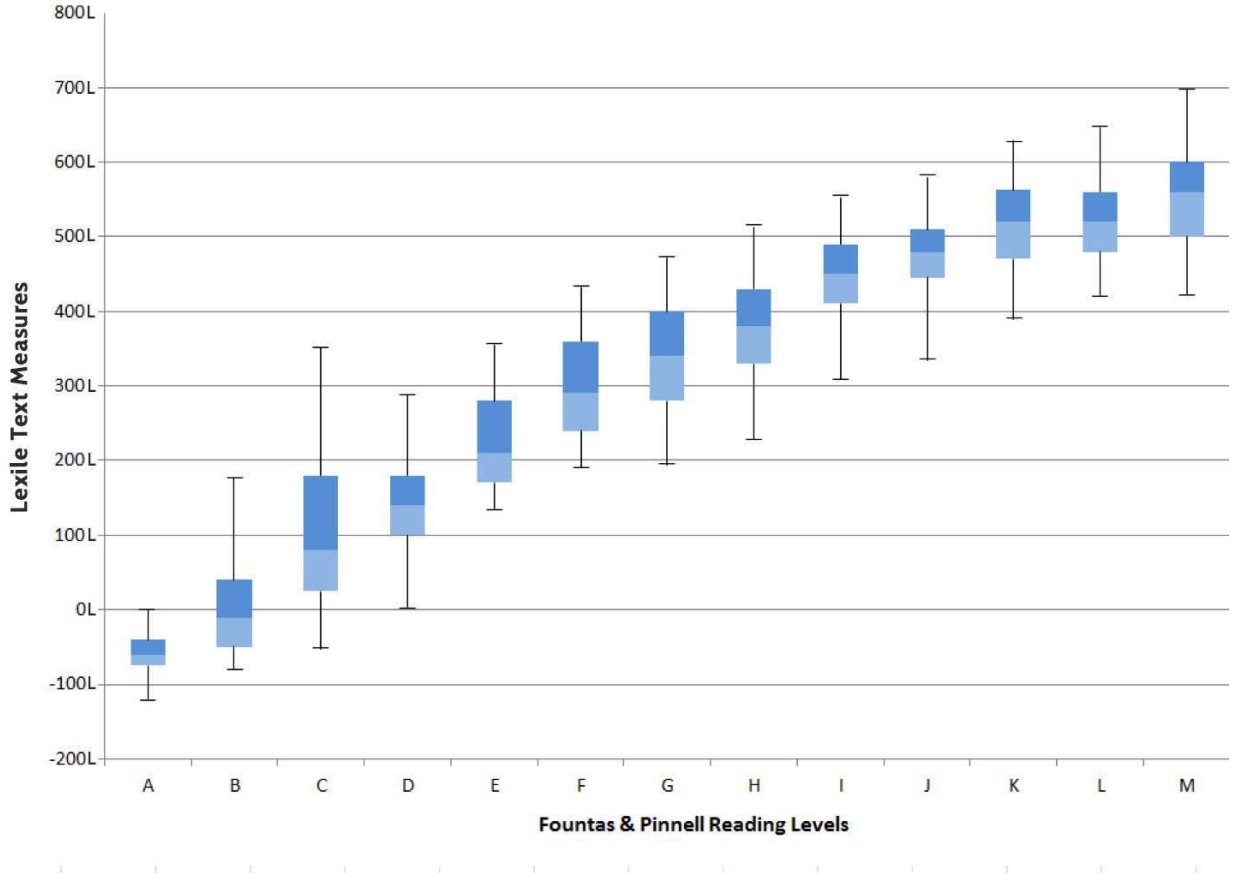
r_{OT} = raw correlation between observed difficulties (O) and theory-based calibrations (*T*).
R_{OT} = correlation between observed difficulties (O) and theory-based calibrations (*T*) corrected for range restriction.
R'_{OT} = correlation between observed difficulties (O) and theory-based calibrations (*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 reading 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.

The Lexile Framework for Reading and Fountas & Pinnell Reading Levels. Koons, Elmore, Sanford-Moore, and Stenner (2017) explored the relationship between Fountas & Pinnell reading levels for a set of texts A through M (i.e. Kindergarten through Grade 2) and their corresponding Lexile reading measures to obtain construct validity evidence for the measurement of early reader texts. The Spearman correlation coefficient between the two text sets was 0.84, indicating a strong positive relationship. Because Fountas & Pinnell reading levels are "larger grained" than the Lexile reading measures, some variation of Lexile reading measures within each Fountas & Pinnell reading level was expected. *Figure 1* shows a series of box-and-whisker plots of the results. The box in each box-and-whisker plot depicts the interquartile range (IQR) with the bottom of the box at the 25th percentile of the distribution of Lexile reading measures, the line between the shaded portions at the median (50th percentile), and the top of the box at the 75th percentile. The bottom whisker depicts the text measure at the 5th percentile of the distribution and the top whisker depicts the text measure at the 95th percentile. *Figure 1* shows steadily increasing Lexile text reading measures across Fountas & Pinnell reading levels for each represented percentile except the 95th percentile of Level C (351L), which has a greater value than the 95th percentile of the two following levels (D: 288L; and E: 350L).

Figure 1. Progression of Lexile text measures and Fountas & Pinnell reading levels, Levels A through M.



The Lexile Framework for Reading and the Difficulty of Reading Test Items. Additional construct validity evidence was obtained by exploring the relationship between Lexile reading calibrations of item difficulties and actual item difficulties of reading comprehension tests. 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 reading 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 Ƶ mean of the raw correlations (r_{OT}) is 0.84. When corrections are made for range restriction and measurement error, the Fisher Ƶ 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 or objectives assessed, or item type used, measure a common comprehension factor specified by the Lexile reading theory.

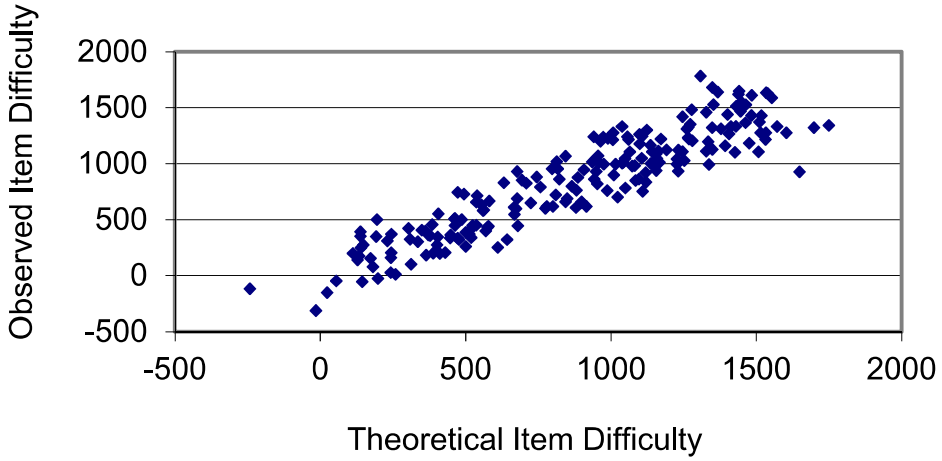
Text Measure Error Associated with the Lexile Framework for Reading

To determine a Lexile reading 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 the enhanced Lexile Reading Analyzer software (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 reading 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 enhanced Lexile Reading Analyzer automates this process, but what "certainty" can be attached to each text measure?

Using a 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 Framework for Reading Map. The Lexile reading 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 2*.

Figure 2. 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
The Stories Julian Tells	46	520L	26
Bunnicula	102	710L	18
The Pizza Mystery	137	620L	15
Meditations of First Philosophy	206	1720L	12
Metaphysics of Morals	209	1620L	12
Adventures of Pinocchio	294	780L	10
Red Badge of Courage	348	900L	10
Scarlet Letter	597	1420L	7
Pride and Prejudice	904	1100L	6
Decameron	2431	1510L	4
War and Peace	4082	1200L	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 tool (hub.lexile.com/find-a-book/search) contains information about each book analyzed: author, Lexile reading measure and Lexile Code, awards, ISBN, and developmental level as determined by the publisher. For some books, Find a Book also provides Lexile text measures by chapter along with selected vocabulary words.

Lexile Item Bank

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

Passage selection. The passages used for item development are excerpted from authentic text, authored by MetaMetrics' staff, or commissioned by MetaMetrics' staff. Excerpted authentic text passages are selected from "real world" reading materials that students 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. Passages authored or commissioned by MetaMetrics staff are created to model "real world" reading materials. The following criteria are used to select passages from authentic and authored passages:

- The passage consists of one main idea or contains one complete piece of information.
- Understanding the passage is independent of the information that comes before or after the passage in the source text.
- Understanding the passage is independent of prior knowledge not contained in the passage.

When writing items based on published text, item writers examine blocks of text that have Lexile reading measures within 100L of the source text (source targeting). Item writers select four to five source-targeted text blocks for potential item development. If it is necessary to shorten or lengthen a 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 to within 100L of the complete text. Items are then developed in conjunction with their associated passages.

When writing original passages, MetaMetrics staff who are experienced in item development and have experience with the everyday reading ability of students at various levels, write original content calibrated to specific Lexile reading zones. Please see "Item Writer Training" in the next section for a detailed description of MetaMetrics' item development process.

Item format. The native-Lexile reading item format is an embedded completion task. 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 125 words in length. The passages are shorter for early 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 options presented, 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, 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.

Items used to assess the reading ability of early readers include picture items, picture/word audio enhanced items, one-sentence items, and two-sentence items. These items are designed using Lexile appropriate vocabulary, sight words, images, and other text characteristics typically associated with early reading. More information on foundational reading items is provided in the next section.

The components of the Lexile Item Bank reading comprehension items and their descriptions are included below.

> *Passage*—the ancillary text for which an item is written. For most items, the Lexile reading measure of the passage is considered the Lexile reading measure of the item. Each passage is used for only one item. For picture items, an image is in place of the passage. For one-sentence items, the passage consists of the stem only. And for two-sentence items, one sentence acts as the passage.
>
> *Stem*—the question or embedded completion statement. For embedded completion statements, they should appear as if they were written as part 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.
>
> *Correct answer*—the correct response. The correct answer (key) typically has a Lexile reading measure similar to the measure of the passage.

Distractor(s)—the three wrong responses that are semantically and syntactically correct. These should be attractive responses if the reader has not read the passage. The distractors have similar Lexile reading measures as the correct answer.

Foundational reading items. Early in their pathway to reading, students develop foundational reading skills which are associated with improved reading outcomes in later stages of reading development and ultimately reading comprehension (National Governors Association & CCSSO, 2010; National Reading Panel, 2000). To support teachers with evaluating the foundational reading skills of students during their early literacy development, and inform instruction appropriate assessment items are needed. In 2019, MetaMetrics conducted research to expand the Lexile Item Bank to include items on the Lexile scale that measure foundational reading skills for children ages 3 to 7/Pre-K to Grade 2 (Webb, Sanford-Moore, Koons, Baker, Hinson, Pringle, and Thorpe, 2021). This research led to the development of a foundational reading framework consisting of three primary domains — Alphabet Knowledge, Phonological Awareness, and Phonics. Each domain is further divided into two or more subdomains (see *Table 5*).

Table 5. Foundational reading domains and subdomains, by grade.

Domain	Subdomain	Grade		
		PK	K	G1
Alphabet Knowledge	Alphabetic Awareness	x	x	
	Letter Sequence		x	
Phonological Awareness	Syllables		x	
	Onsets and Rimes		x	
	Phoneme Isolation		x	x
	Phoneme Blending		x	x
	Phoneme Segmenting		x	x
	Phoneme Manipulation			x
	Rhyme			x
Phonics	Consonant Sounds		x	x
	Consonant Digraphs and Blends			x
	Letter Sound Correspondence		x	
	Vowel Sounds		x	x
	Word Families			x

Targeting each of the foundational reading domains and subdomains in *Table 1*, MetaMetrics developed items ($N = 102$) which were then reviewed by subject matter experts, teachers, and test development researchers in summer 2019 and field-tested in fall 2019. The participants in the field-test study included a total of 1,738 students in Pre-kindergarten ($n = 222$), Kindergarten ($n = 901$) and Grade 1 ($n = 615$) across 30 U.S. states representative of all geographical regions. The students were from 111 classrooms in 73 different schools. Analysis of the resulting data placed each item on the Lexile scale.

Item writer training. Item writers are professional writers, classroom teachers, and other educators who have had experience with the everyday reading ability of students at various

levels. Experienced item writers help to ensure that all Lexile Item Bank reading comprehension items are valid measures of reading comprehension. New item writers practice item writing and reviewing over one to two months so that senior curriculum specialists can provide them with specific and individualized feedback to ensure proper training. 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 training materials also contain examples of poorly constructed items to illustrate the criteria used to evaluate items and corrections based on those criteria. Item writers are also provided vocabulary lists to use during statement and option development. The vocabulary lists were assembled from word lists compiled by MetaMetrics based on vocabulary research related to determining the Lexile reading measures (difficulty) of words (MetaMetrics, 2006). 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 given extensive training related to "sensitivity" issues. Item writing training materials provide examples of sensitivity issues and identify areas to avoid when selecting or writing passages and developing items. The following areas are covered: violence and crime, sources of common phobias, negative emotions such as death and family issues, offensive language, drugs/alcohol/tobacco, sex/attraction, race/ethnicity, class, gender, religion, supernatural/magic, parent/family, politics, animal cruelty and hunting, environmental issues, brand names, and junk food. These materials were developed based on material published by McGraw-Hill (Guidelines for Bias-Free Publishing, 1983) related to universal design and fair-access—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 multi-stage review process. First, items are reviewed and edited by item writers and reviewers according to the 25 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 reviewed and edited by a group of specialists that represent various perspectives—curriculum specialists, content editors, fact checkers, sensitivity reviewers, and test developers. These individuals examine each item for sensitivity issues, grammar and spelling, and item quality (stem, key, and distractors).

During the second stage of the item review process, items are either "approved as presented," "approved with edits," or "rejected." Approximately 10% of the items written are "approved with edits" or "rejected" at this stage. When necessary, item writers receive additional 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 reading measure] and theoretical Lexile reading measure), internal consistency (point-biserial or point-measure correlation), and bias (ethnicity and gender where possible). Where necessary, items are deleted from the item bank or revised and recalibrated.

In addition to content and sensitivity reviews during the development process, Lexile Item Bank items are field-tested as part of MetaMetrics on-going research. Lexile Item Bank items may be

field-tested as part of stand-alone research field tests or they may be embedded within research tests for concurrent projects. During the spring of 1999, 8 levels of a Lexile reading 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 Rasch analysis. The resulting item difficulties (in logits) were assigned Lexile reading measures by multiplying by 180 and anchoring each set of items to the mean theoretical difficulty of the items on the form.

MetaMetrics continues to add new items to its item bank and regularly evaluates items for potential use on linking studies. Each time items are administered, their empirical data are evaluated to determine whether they should be removed from the item bank, revised and retested, or kept for future use on tests developed for MetaMetrics' partners, linking studies, and research studies.

The NC READY EOG/EOC English II–Lexile Framework for Reading 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 for Reading is a tool that helps teachers, parents, and students locate appropriate reading materials. Text complexity (difficulty) and reader ability are measured in the same unit—the Lexile measure. 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 reading measures can range from below zero (e.g., BR 150L) to above 1600L.

The Lexile Framework for Reading measures reading ability by using multiple-choice items focused on the skills readers use to comprehend written materials sampled from various content areas including both literary and informational text. Lexile items do not require prior knowledge of ideas outside of the passage, vocabulary taken out of context, or formal logic. 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.

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 6*.

Table 6. 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 7*.

Table 7. Item statistics from the administration of the Lexile Linking Tests.

Grade	Test Form	*N* (Persons)	*N* (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 Brennen, 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 8 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 8. Number of student tests received and number of students in the matched sample.

Grade	NC READY EOG Reading/EOC English II Received *N*	Lexile Linking Test *N*	Matched *N*	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 9 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 9. Comparison of matched sample and final sample and the reason for student removal.

Matched Sample		***N* Removed by Reason**				**Final Sample**	
Grade	***N***	**Accommodated Students**	**Misfit to Rasch**	**Scores Removed***	**Percentile Rank Difference**	***N***	**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 10 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 10. 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
EC Code	Autism	0.5	0.6	0.6

Student Characteristic	Category	State Sample *N*=541,124	Matched Sample *N*=12,356	Final Sample *N*=9,777
	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 11 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 11,* it can be concluded that the two tests are measuring similar reading comprehension constructs.

Table 11. 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 12 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 12. 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 3 through *12* 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 3. 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).

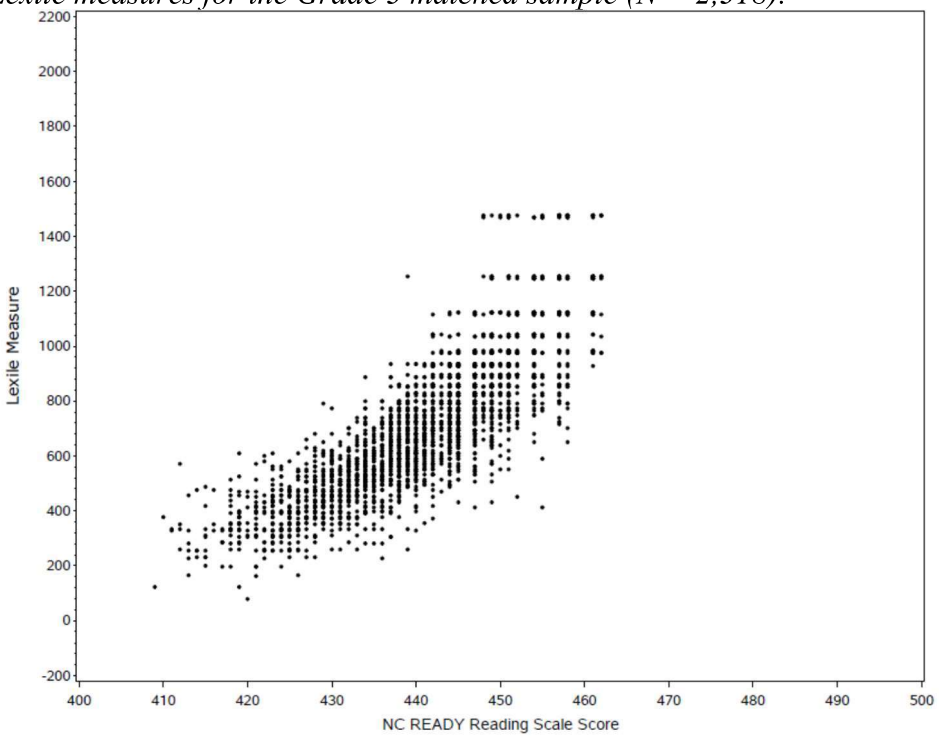


Figure 4. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 3 final sample (N = 1,903).

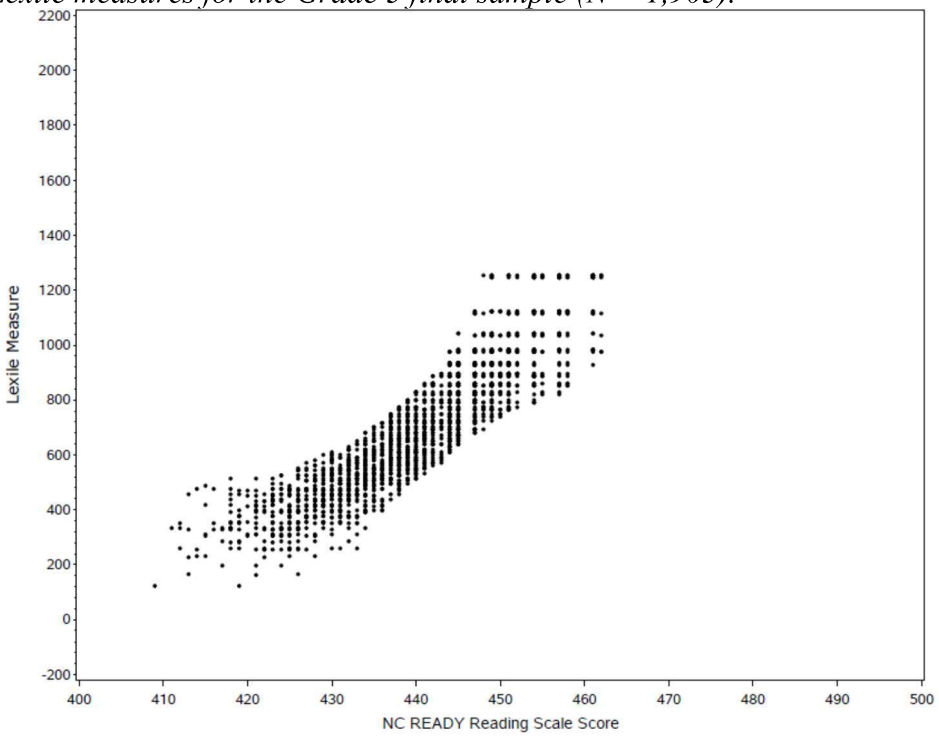


Figure 5. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 5 matched sample (N = 2,260).

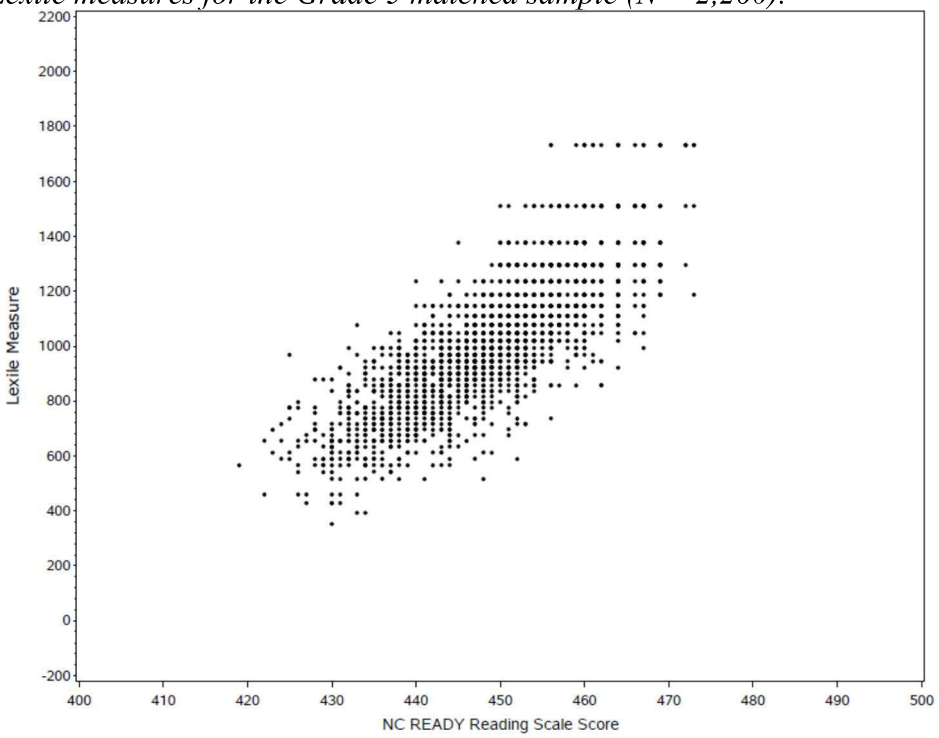


Figure 6. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 5 final sample (N = 1,727).

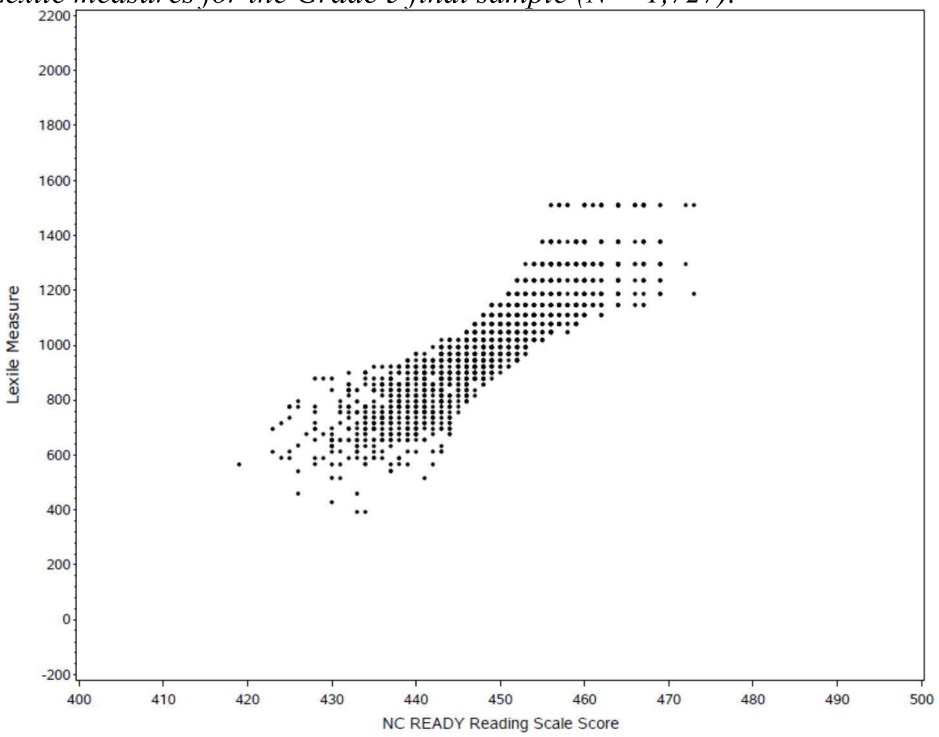


Figure 7. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 7 matched sample (N =2,224).

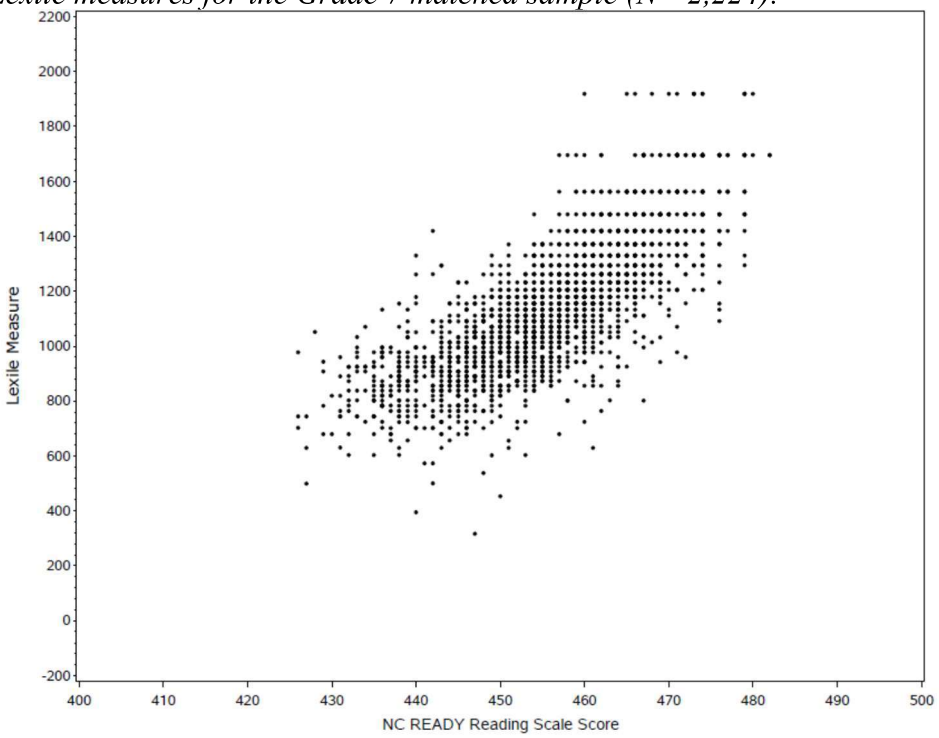


Figure 8. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 7 final sample (N = 1,770).

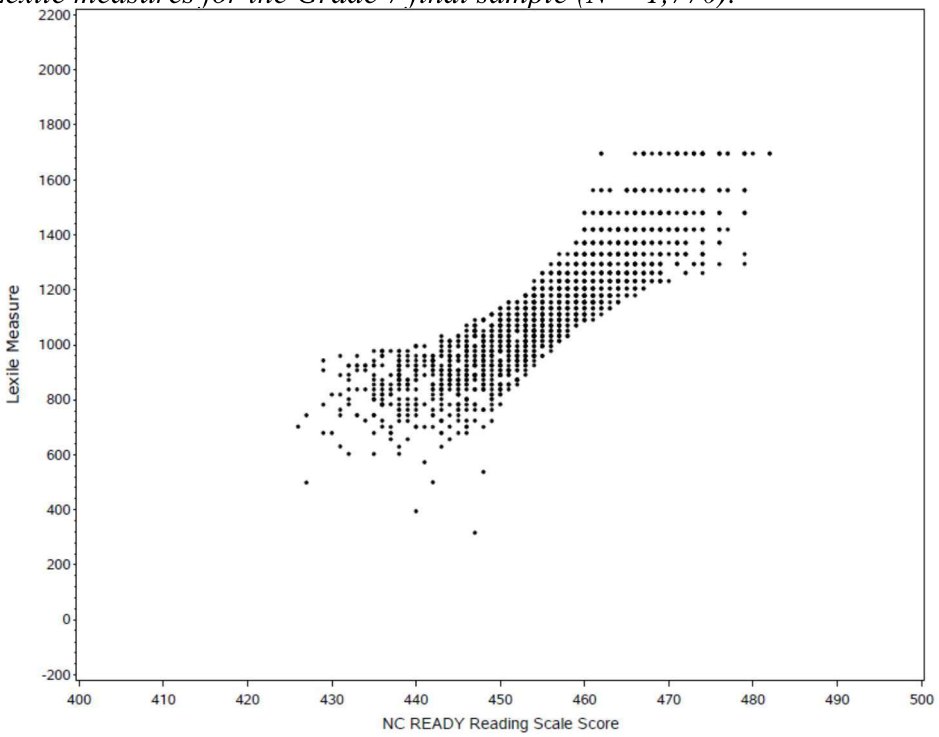


Figure 9. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 8 matched sample (N = 2,939).

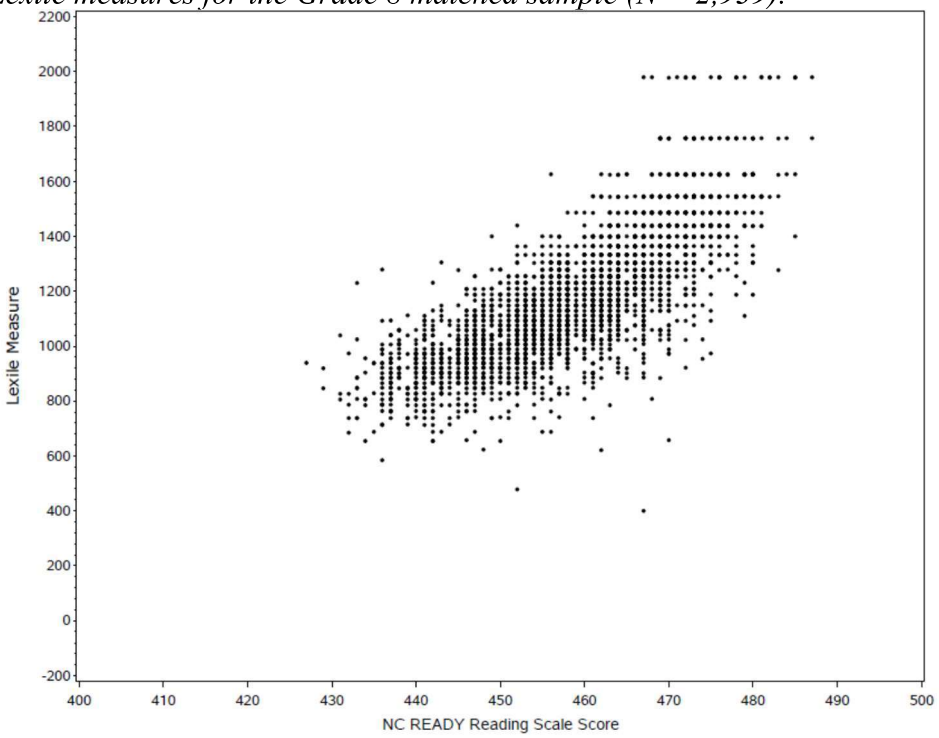


Figure 10. Scatter plot of the NC READY EOG Reading scale scores and the Lexile Linking Test Lexile measures for the Grade 8 final sample (N = 2,309).

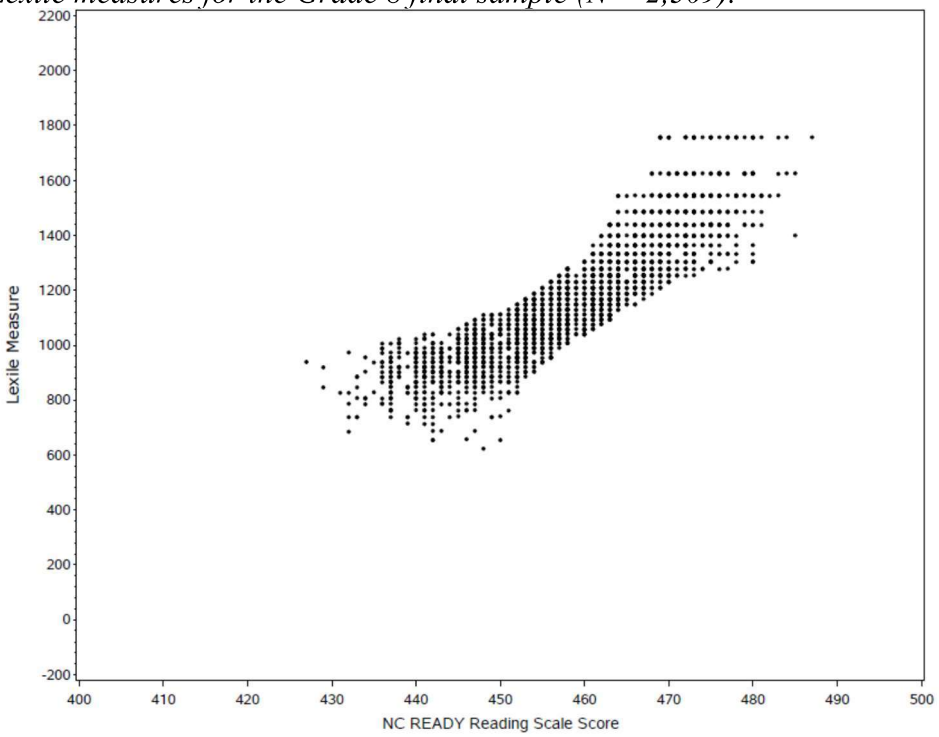


Figure 11. Scatter plot of the NC READY EOC English II scale scores and the Lexile Linking Test Lexile measures for the English II matched sample (N = 2,615).

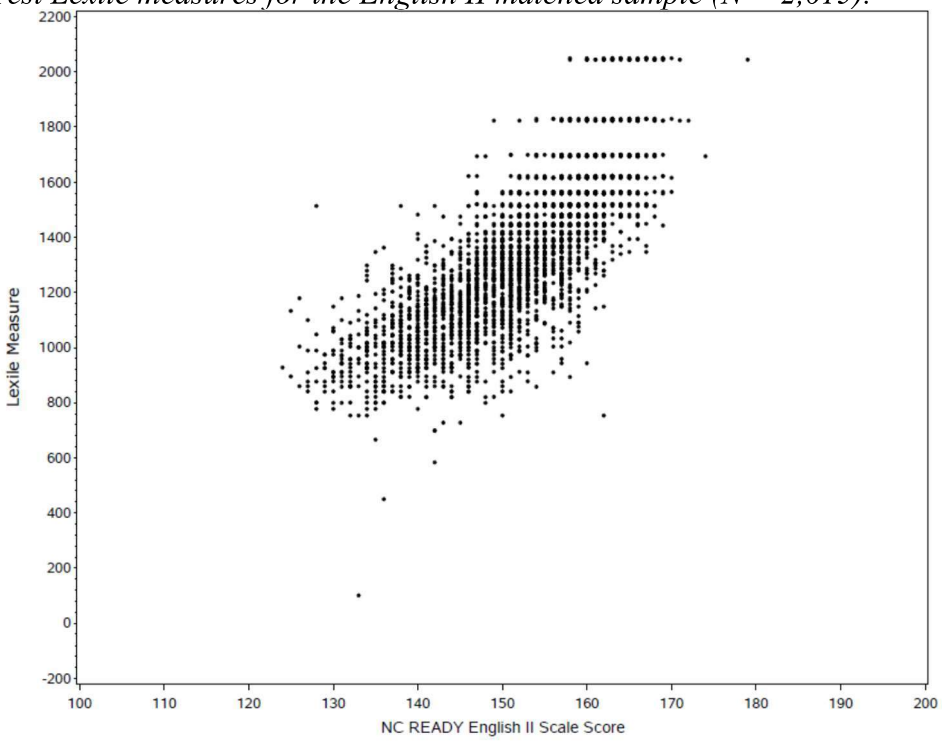
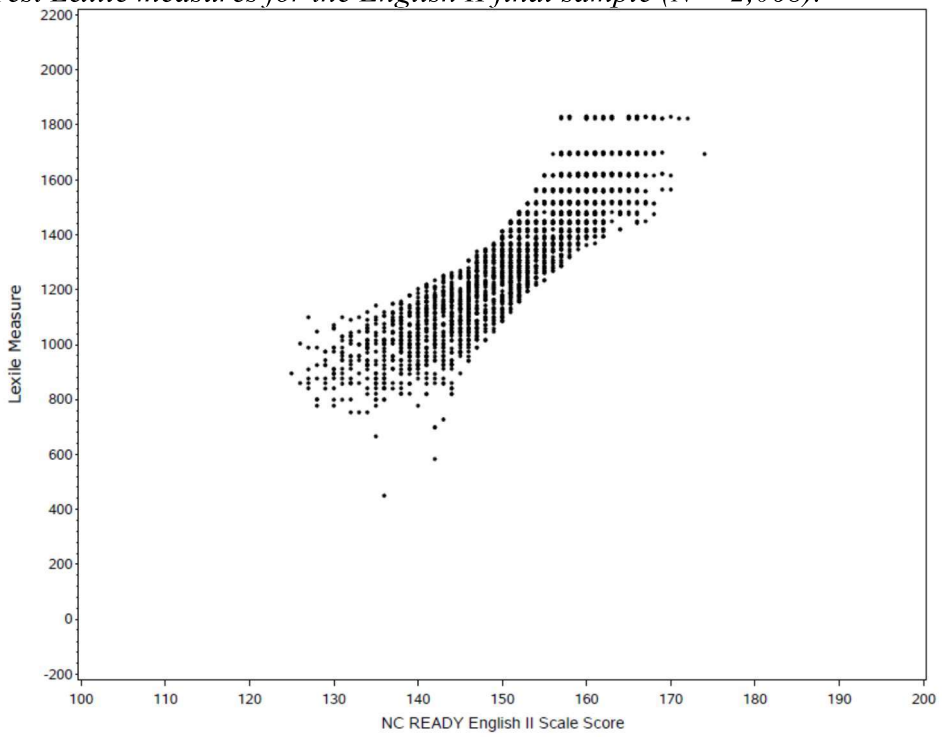


Figure 12. Scatter plot of the NC READY EOC English II scale scores and the Lexile Linking Test Lexile measures for the English II final sample (N = 2,068).



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 12,* 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

$$\frac{(x-\mu_X)}{\sigma_X} = \frac{(y-\mu_y)}{\sigma_y} \qquad \text{(Equation 2)}$$

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

$$x = l_x(y) = \left(\frac{\sigma_X}{\sigma_y}\right)y + \left(\mu_x - \frac{\mu_y\sigma_X}{\sigma_y}\right) \qquad \text{(Equation 3)}$$

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 12*.

Using the final sample data described in *Table 12*, 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 13*. 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 13. Linear linking equation coefficients used to predict Lexile measures from the NC READY EOG Reading and the EOC English II scale scores.

Group (*g*)	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 14* 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 14. Capped values of the Lexile measure by grade/course.

Grade/Course	Capped Lexile
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 15 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 15. 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	*N*	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 16. Interpretation chart for effect size.

Small	0.20
Medium	0.50
Large	0.80

In *Table 15*, 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 17 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 17. 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

Table 17 (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.

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

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 18 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.

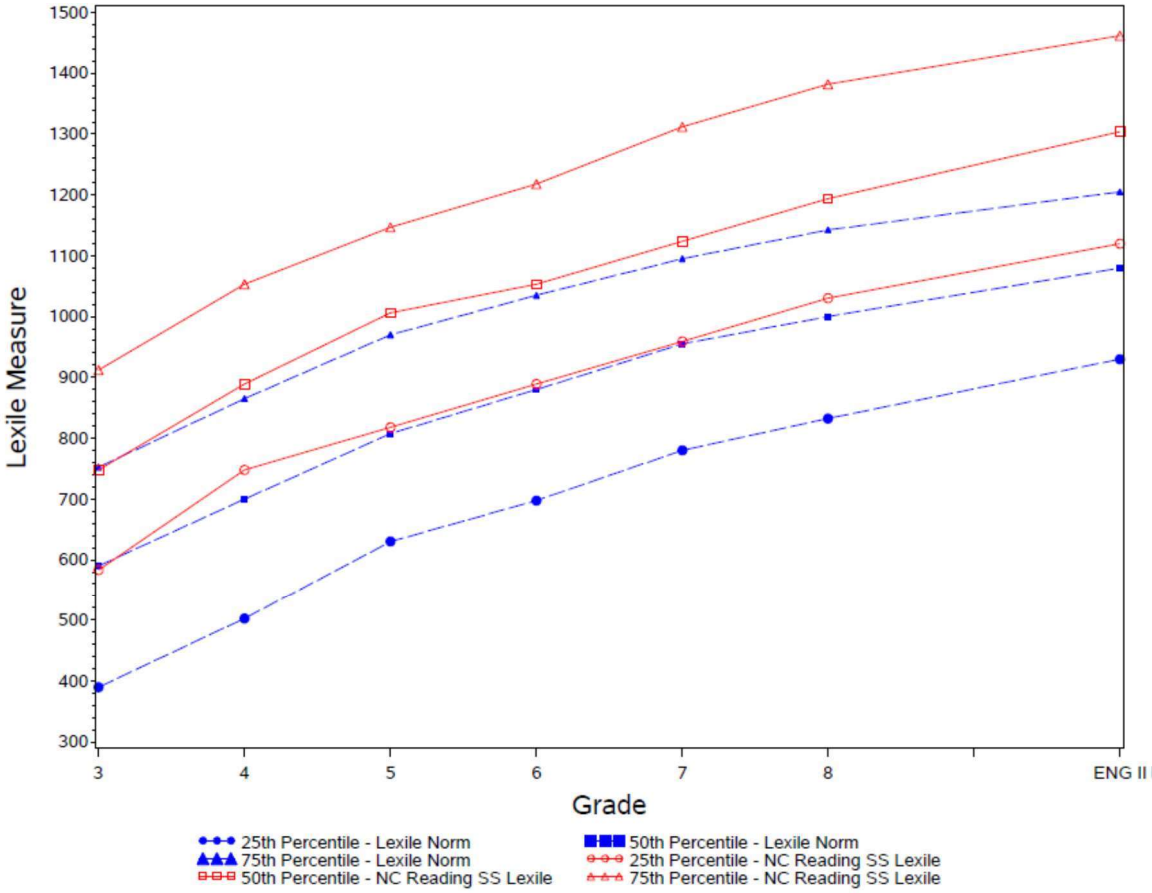
Table 18. NC READY EOG Reading/EOC English II performance level cut scores and the associated Lexile measures.

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 13 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 13. 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 14, 15,* and *16*) 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 14, 15,* and *16* 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 14. Box and whisker plot of the Lexile Linking Tests Lexile measures by grade, final sample (N =9,777).

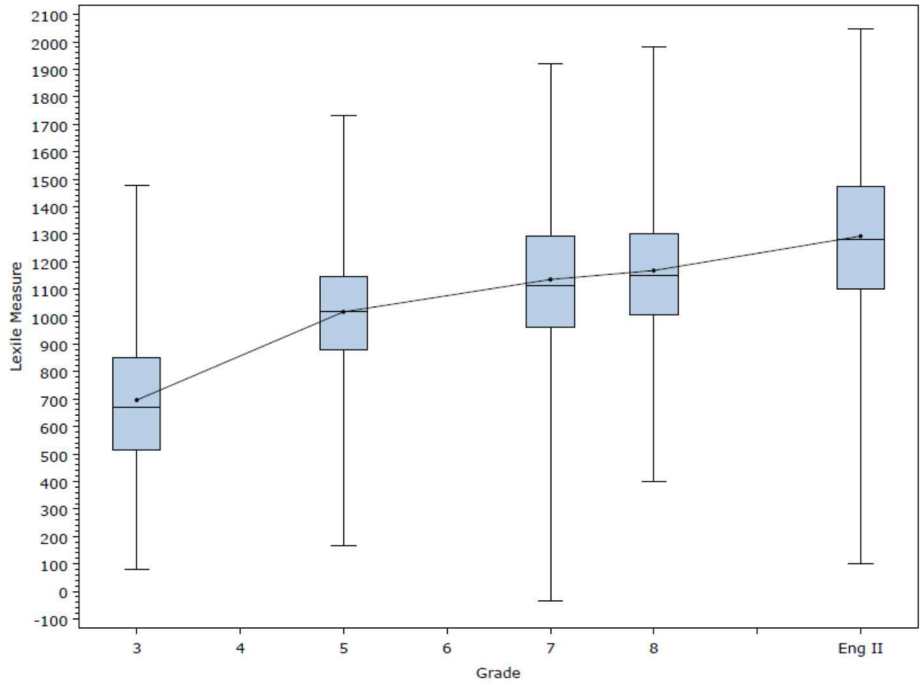


Figure 15. Box and whisker plot of the NC READY EOG Reading/EOC English II Lexile measures by grade, matched sample (N = 12,356).

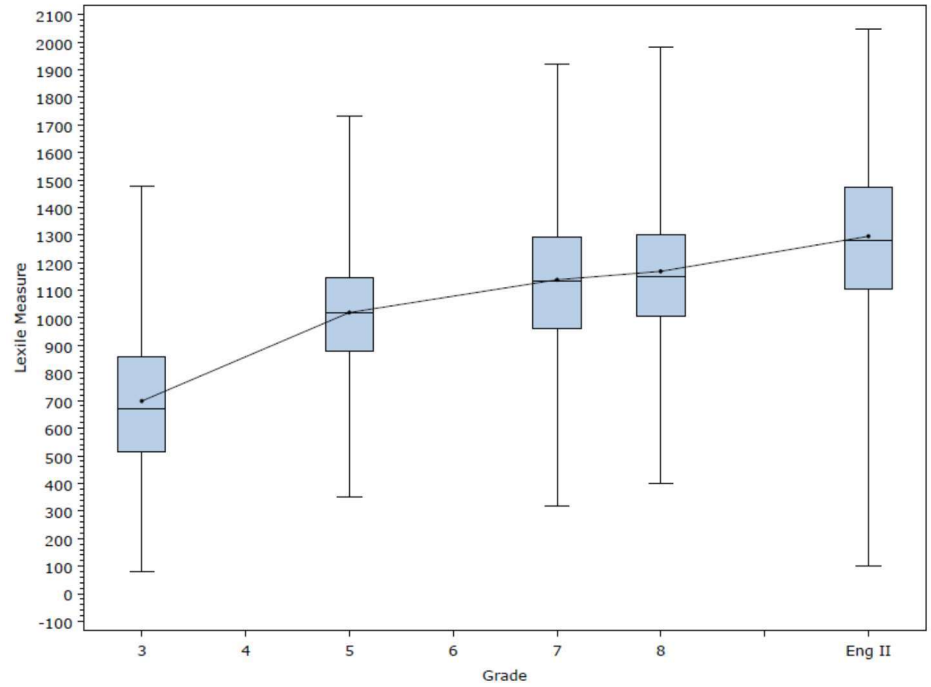
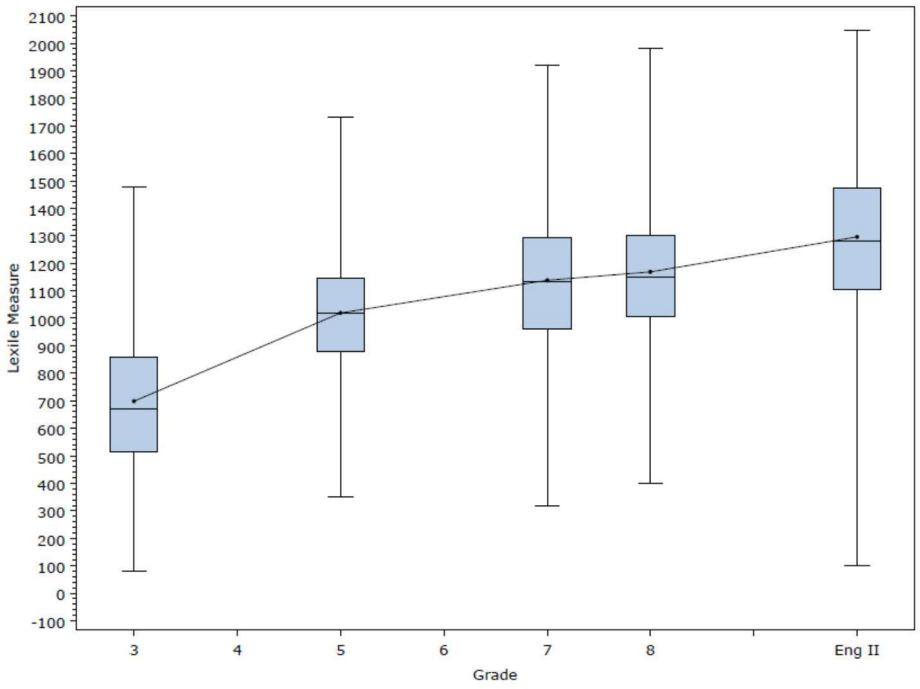


Figure 16. Box and whisker plot of the NC READY EOG Reading/EOC English II Lexile measures by grade, final sample (N = 9,777).



The Lexile Framework for Reading and Forecasted Comprehension Rates

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

The 75% comprehension rate for an examinee-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. An individual who answers three-fourths of the questions correctly has a 75% comprehension rate.

Suppose instead that the text and the examinee measures are not the same. It is the difference in Lexile reading measures between the examinee and text that governs comprehension. If the text measure is less than the examinee 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 individual reads a 350L text?

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

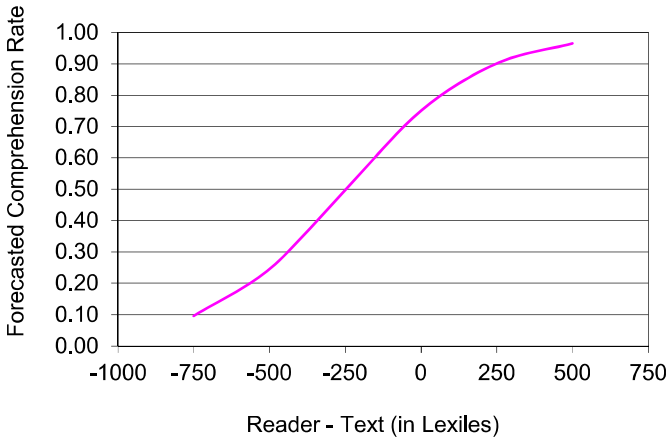
$$Rate = \frac{e^{ELD+1.1}}{1+e^{ELD+1.1}} \qquad \text{Equation (5)}$$

where ELD is the "effective logit difference" given by

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

Figure 17 shows the general relationship between examinee-text discrepancy and forecasted comprehension rate. When the examinee 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 examinee measure of 600L and the text calibration of 350L is 250L. Referring to *Figure 17* and using +250L (examinee minus text), the forecasted comprehension rate for this examinee-text combination would be 90 percent.

Figure 17. Relationship between examinee-text discrepancy and forecasted comprehension rate.



Tables 19 and *20* show comprehension rates calculated for various combinations of examinee measures and text calibrations.

Table 19. Comprehension rates for the same individual with materials of varying comprehension difficulty.

Examinee Reading Measure	Text Measure	Sample Titles	Forecast Comprehension
1000L	500L	*Tornado* (Byars)	96%
1000L	750L	*The Martian Chronicles* (Bradbury)	90%
1000L	1000L	*Reader's Digest*	75%
1000L	1250L	*The Call of the Wild* (London)	50%
1000L	1500L	*On the Equality Among Mankind* (Rousseau)	25%

Table 20. Comprehension rates of different examinee abilities with the same material.

Examinee Reading Measure	Calibration for a Grade 10 Biology Textbook	Forecasted Comprehension Rate
500L	1000L	25%
750L	1000L	50%
1000L	1000L	75%
1250L	1000L	90%
1500L	1000L	96%

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

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

Examinee Forecasted Comprehension Rate. Using Equation 6 with different combinations of examinee measure and text difficulty, a forecasted comprehension rate can be determined. *Table 21* shows the changes in the forecasted comprehension rate for different combinations of examinee and text interactions.

Table 21. Effect of examinee-text discrepancy on forecasted comprehension rate.

Examinee Lexile Reading 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

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.

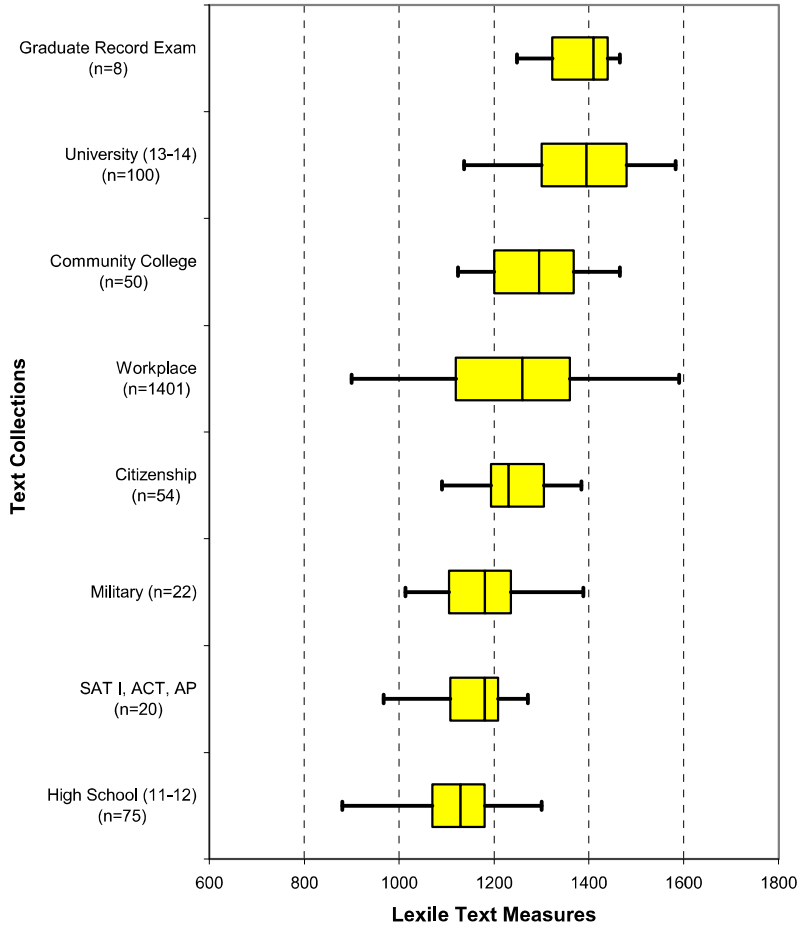
Some assessments report a Lexile range for each student, which is 50L above and 100L below the student's actual Lexile measure. This range represents the boundaries between the easiest kind of reading material for the student and the level at which the student will be more challenged, yet can still read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

College and Career Reading Demands

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. 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 reading 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 18. A continuum of text difficulty for the transition from high school to postsecondary experiences (box plot percentiles: 5th, 25th, 50th, 75th, and 95th).[1]



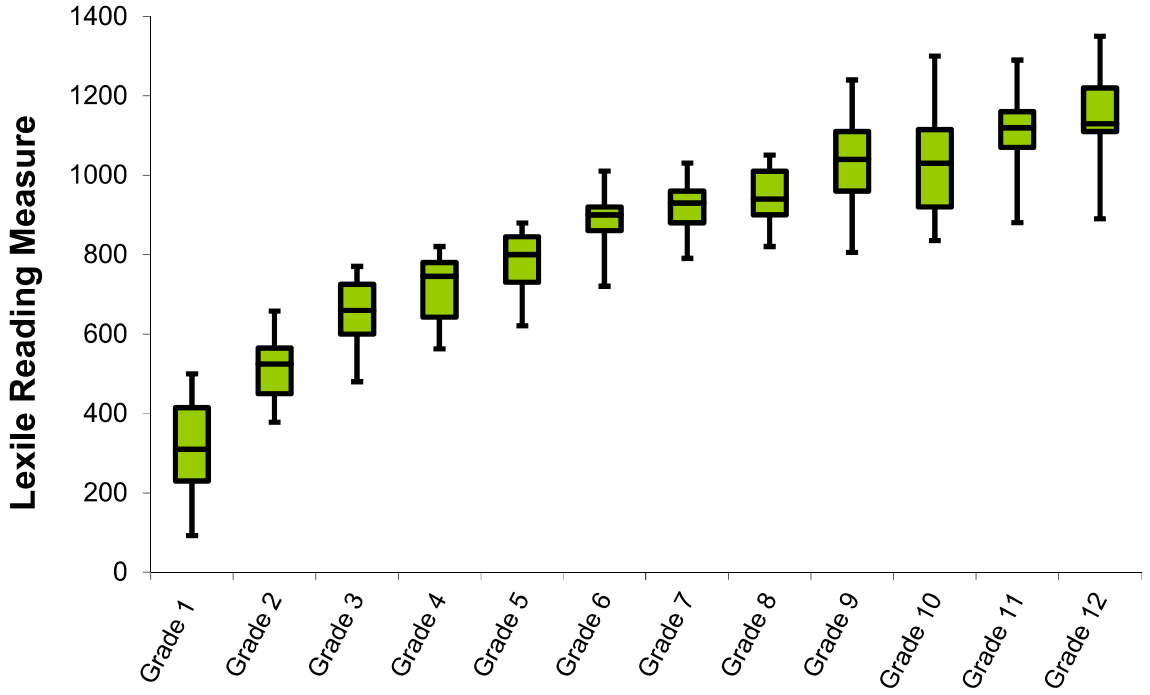
[1] Reprinted from Williamson, G. L. (2008). A text readability continuum for postsecondary readiness. *Journal of Advanced Academics, 19*(4), 602-632.

In Texas, two studies (MetaMetrics, 2007; MetaMetrics, 2008) were conducted to examine the reading demands in various postsecondary options – technical college, community college, and 4-year university programs. Under Commissioner Raymond Paredes, the Texas Higher Education Coordinating Board (THECB), in conjunction with MetaMetrics, 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. The reading text complexity of these books was measured using the Lexile Framework for Reading. Since the TAKS (Texas Assessment of Knowledge and Skills) had already been linked with Lexile reading measures for several years, the THECB study was able to overlay the TAKS cut scores onto the post high school reading requirements.

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

Between 2004 and 2008, MetaMetrics (Williamson, Koons, Sandvik, and Sanford-Moore, 2012) conducted research to describe the typical reading demands and develop a text continuum of reading materials across Grades 1-12. The grade-by-grade text distributions are presented in *Figure 19*.

Figure 19. Reading text complexity distributions, in Lexile reading units, by grade (whiskers represent 5th and 95th percentiles).



This continuum can be "stretched" to describe the reading demands students will likely encounter in Grades 1-12 when "on track" for college and career (Sanford-Moore and Williamson, 2012). 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. *Table 22* provides the stretch text measure ranges for Grades 1 through 12. Combining student results with criterion referenced indicators provides information to reference when matching students with reading materials that are at or above the recommendations in Appendix B for each grade level.

Table 22. Lexile reading ranges aligned to college- and career-readiness reading 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

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 20*. *Figure 20* 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 20. Comparison of NC READY EOG Reading/EOC English II "Level 3" standards with college and career reading levels described by the CCSS.

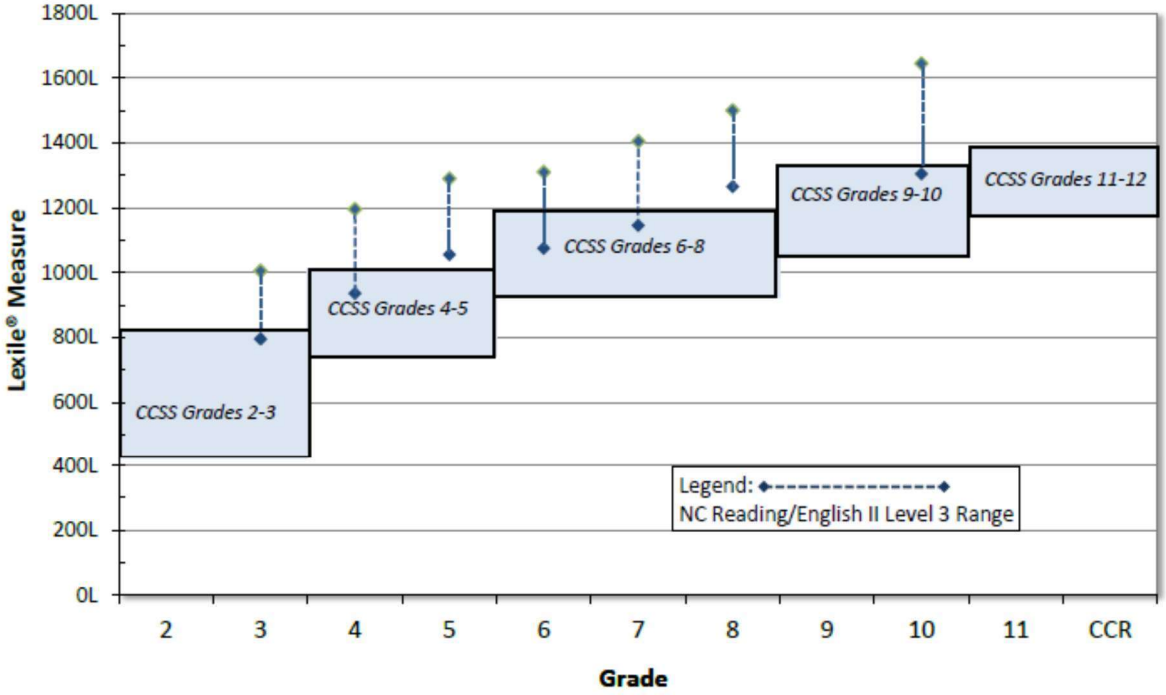
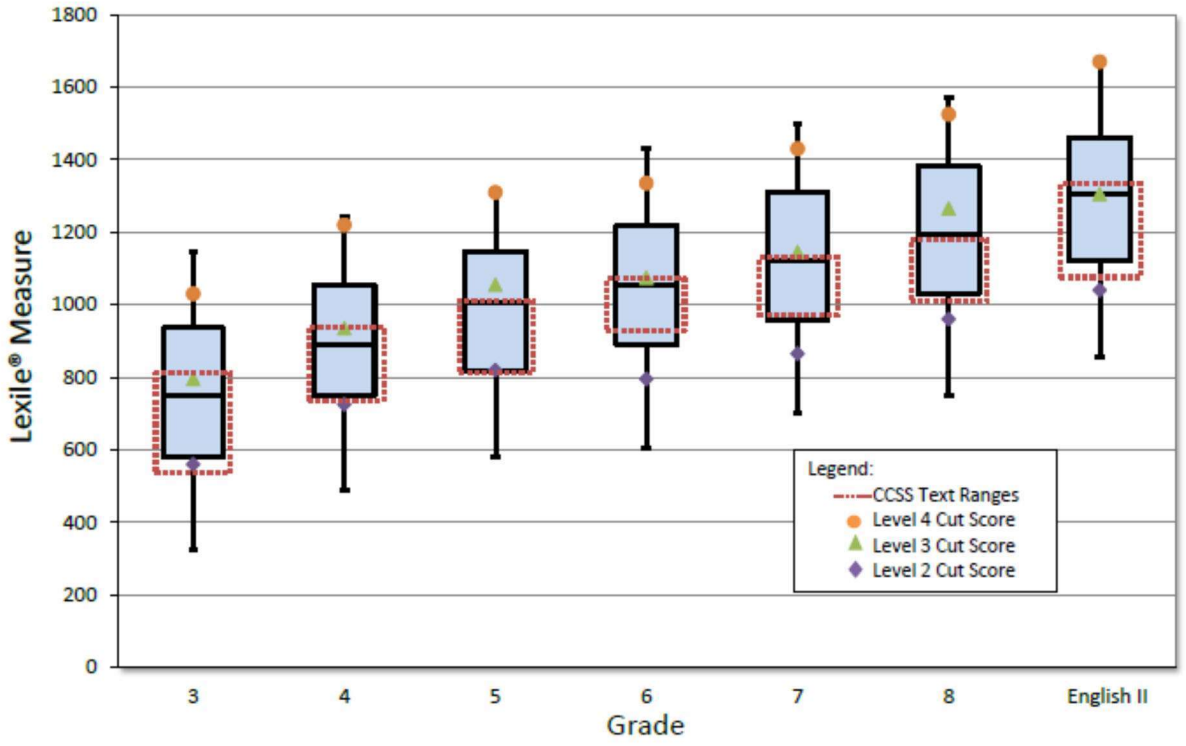


Figure 21 shows that the spring 2013 student performance on the NC READY EOG Reading/EOC English II assessments at each grade level is "on track" for college and career readiness. Students can be matched with reading materials that are at or above the recommendations in Appendix B of the CCSS for ELA for each grade level.

Figure 21. 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 23*. 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 18*. 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 23* shows that the NC READY EOG Reading standards are demanding more of students in terms of reading ability in 2013.

Table 23. 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

To utilize the results from this study, Lexile reading measures need to be incorporated into the NC READY EOG Reading/EOC English II results processing and interpretation frameworks. When that occurs, educators can use the measures and tools available through the Lexile Framework to link the assessment results with subsequent instruction. The following sections provide a description of the caveats and limitations associated with the study and recommended uses of the Lexile Framework and other associated tools.

Caveats and Limitations

Lexile Reading Measures and Grade Levels. Lexile reading measures do not translate specifically to grade levels. Within any grade, there will be a range of students and a range of materials to be read. In an eighth-grade classroom there will be some students who are far ahead of the others and there will be some students who are behind the others in terms of reading ability. To say that some books are "just right" for eighth graders assumes that all eighth graders are reading at the same level. The Lexile Framework for Reading can be used to match students with texts at whatever level the student is reading.

Simply because a student is an excellent reader, it should not be assumed that the student would necessarily comprehend a text typically found at a higher grade level. Without adequate background knowledge, the words may not have sufficient meaning to the student. A high Lexile reading measure for a grade indicates that the student can read grade-appropriate materials at a higher comprehension level (90%, for example).

Maintenance of the NC READY EOG Reading/EOC English II reading scale. Maintenance of the focal scale (i.e., NC READY EOG Reading/EOC English II reading scale) is critical to the validity of any link with an auxiliary scale (i.e., The Lexile scale). If an update occurs to the focal scale, the integrity of the link needs to be re-evaluated and additional linking studies may be needed to incorporate fundamental changes to the focal scale. Such updates include, but are not limited to, the incorporation of new item types into the scale; revision of item calibrations and the resulting scale; or revision of the assessment program and, therefore, the reported scale scores.

Linking error. Error in estimating the linking relationship of two scales is present whenever linking is conducted. Not all error associated with a study can be accounted for; however, error should be continually investigated to ensure scores are as accurate and reliable as possible. The two sources of error present are random error and systematic error. Random linking error occurs when directly estimating the linking relationship because a sample is collected to perform the study. Systematic error occurs when estimation methods introduce bias, statistical assumptions for the methods are not met, improper sampling techniques were used to collect the data for the linking study, or different placement of items impacts scale scores. To the extent possible, MetaMetrics and the North Carolina Department of Public Instruction worked to minimize systematic error through the design of the linking study.

Recommendations for Using the Lexile Framework for Reading

Use the Lexile Framework for Reading to Select Books. Teachers can use the tools provided by the Lexile Framework for Reading to select materials to develop individualized reading lists that are tailored to individual students. 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. A student's Lexile reading measure can be leveraged to aid comprehension and is a good starting point in the selection process of a book for a specific reader.

The Lexile Framework for Reading is an objective tool that can be used to determine a student's readiness for a reading experience; the Lexile Framework for Reading "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).

Another feature of the Lexile Framework for Reading is that it makes provisions for students who read below or beyond their grade level, because the reporting scale is not bounded by grade level. See the Lexile Framework for Reading Map for literary and informational titles, leveled reading samples, and approximate grade ranges (Appendix B).

However, it is important to note that the Lexile reading measure should never be the only piece of information used when selecting a text for a reader. When matching a book with a reader, one must also consider other factors that may affect the relationship between a reader and a book. These factors include student developmental level, motivation, and interest; amount of background knowledge possessed by the reader; and suitability of the text and text difficulty. For example, if a student is highly motivated for a particular reading task (e.g., self-selected free reading), the teacher may suggest books higher in the student's Lexile reading range. If the student is less motivated or intimidated by a reading task, material at the lower end of his or her Lexile reading range can provide the basic comprehension support to keep the student from feeling overwhelmed.

The Lexile Framework for Reading does not prescribe a reading program, but it gives educators more knowledge of the variables involved when they design reading instruction. The Lexile Framework for Reading facilitates multiple opportunities for use in a variety of instructional activities. After becoming familiar with the Lexile Framework for Reading, teachers are likely to think of a variety of additional creative ways to use this tool to match students with books that students find challenging, but not frustrating.

Target Instruction to Students' Abilities. To encourage optimal progress with the use of any reading materials, teachers need to be aware of the complexity level of the text relative to a student's reading level. A text that is too difficult may serve to undermine a student's confidence and diminish learning. Frequent use of text that is too easy may foster poor work habits and unrealistic expectations that will undermine the later success of the best students.

When students confront new kinds of texts and texts containing new content, the introduction can be softened and made less intimidating by guiding the student to easier reading. On the other hand, students who are comfortable with a particular genre or format or the content of such texts can be challenged with more difficult reading levels, which will reduce boredom and promote the greatest rate of development of vocabulary and comprehension skills.

Similarly, teachers can use Lexile reading measures to guide a struggling student by selecting texts at the lower end of the student's Lexile reading range (e.g., 50L below his or her Lexile reading measure). At the same time, teachers can also motivate advanced students by challenging them with reading texts at the midpoint of their Lexile reading range or slightly above (i.e., 25L above to 100L above his or her Lexile reading measure).

Teach Learning Strategies by Controlling Comprehension Match. The Lexile Framework for Reading permits the teacher to target readers with challenging text and to systematically adjust text targeting when the teacher wants fluency and automaticity (i.e., reader measure is well above text measure) or wants to teach strategies for attacking "hard" text (i.e., reader measure is well below text measure). For example, metacognitive ability has been well documented to play an important role in reading comprehension performance. Once teachers know the kinds of texts that would likely be challenging for a group of readers, they can systematically plan instruction that will allow students to encounter difficult text in a controlled fashion and make use of instructional scaffolding to build student success and confidence with more challenging text. The teacher can model appropriate learning strategies for students, such as rereading or rephrasing

text in one's own words, so that students can then learn what to do when comprehension breaks down. Students can then practice these metacognitive strategies on selected text while the teacher monitors their progress.

Apply Lexile Reading Measures Across the Curriculum. Over 600 publishers provide Lexile reading measures for their trade books and textbooks, enabling educators to make connections among all of the different components of the curriculum to plan instruction more effectively. With a student's Lexile reading measure, teachers can connect him or her to hundreds of thousands of books. Using periodical databases, teachers and students can also find appropriately challenging newspaper and magazine articles that have Lexile reading measures.

Use the Lexile Framework for Reading to facilitate communicating with stakeholders. Lexile reading 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 an appropriate level of challenge and monitor the reading process at home. The "Find A Book" website also 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 reading 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 https://hub.lexile.com/find-a-book/search.

In one large school district, the end-of-year testing results are sent home to parents in a folder. The folder consists of the Lexile Framework for Reading 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 Framework for Reading 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, stakeholders can understand what the community values and more readily 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 reading 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 for Reading 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.

There are numerous ways to incorporate the Lexile Framework for Reading including:

- Building text sets that include texts at varying levels to enhance thematic teaching. These texts might not only support the theme, but also provide a way for all students to successfully learn about and participate in discussions about the theme, building knowledge of common content for the class while building the reading skills of individual students. Such discussions can provide important collaborative brainstorming opportunities to fuel student writing and synthesize the curriculum.
- Sequencing materials in a reading program to encourage growth in reading ability. For example, an educator might choose one article a week for use as a read-aloud. In addition to considering the topic, the educator could increase the complexity of the articles throughout the course. This approach is also useful when utilizing a core program or textbook that is set up in anthology format. (The order in which the readings in anthologies are presented to the students may need to be rearranged to best meet student needs).
- Developing a reading folder that goes home with students and comes back for weekly review. The folder can contain a reading list of texts within the student's Lexile reading range, reports of recent assessments, and a form to record reading that occurs at home. This is an important opportunity to encourage individualized goal setting and engage families in monitoring the progress of students in reaching those goals.
- Selecting texts lower in the student's Lexile reading range when factors make the reading situation more challenging or unfamiliar. Select texts at or above the student's range to stimulate growth when a topic is of extreme interest to a student, or when adding additional support such as background teaching or discussion.
- Enhancing a student's experience with exposure to differentiated, challenging text at least once every two to three weeks.
- Leveraging the free Find a Book website (at https://hub.lexile.com/find-a-book/search) to support book selection and create booklists within a student's Lexile reading range to help the student make more informed choices when selecting texts.
- Utilizing database resources to infuse research into the curricula while tailoring reading selections to specific Lexile reading levels. In this way, students can explore new content at an appropriate reading level and then demonstrate their assimilation of that content through writing and/or presentations. A list of the database service providers that have their collections measured can be found at https://metametricsinc.com/products/library-products/.

- Using the Lexile® WordBank (https://hub.lexile.com/wordlists) to identify subsets of words that are relevant to the context or application. The WordBank is a new resource containing approximately 50,000 unique words from the top four best-selling textbook programs (published after 2011) in science, math, social studies, and reading/English language arts. Some common uses include: identifying grade appropriate words to target vocabulary instruction and assessment; identifying words to include in instructional materials for domain-specific content; and selecting important academic words by grade and domain to highlight in reading passages, books or other instructional materials.

Use the Lexile Framework for Reading in the Library. Augmenting libraries provides even more ways to leverage the Lexile Framework for Reading including:

- Making the Lexile reading measures of books available to students to better enable them to find books of interest at their appropriate reading level.
- Enabling comparison of student Lexile reading levels with the Lexile reading levels of the books and periodicals in the library to analyze and develop the collection to more fully meet the needs of all students.
- Leveraging the database resources to search for articles at specific Lexile reading levels to support classroom instruction and independent student research. A list of the database service providers that have had their collections measured can be found at https://metametricsinc.com/products/library-products/).
- Using the free Find a Book website (at https://hub.lexile.com/find-a-book/search) to support book selection and help students make informed choices when selecting texts.

Set and Monitor Reading Program Goals. Schools often write grant applications in which they are required to state how they will monitor progress of the intervention or program funded by the grant. Schools that receive funds targeted to assist students with improving their reading skills can use the Lexile Framework for Reading for evaluation purposes. Schools can use student-level and school-level Lexile reading information to monitor and evaluate interventions designed to improve reading skills. Progress tests throughout the year can be conducted to help monitor students' progress toward their goals.

Students' Lexile reading measures can also be used to identify reading materials that students are likely to comprehend with 75% accuracy. Students can set goals of improving their reading comprehension and plan clear strategies for reaching those goals using literature from the appropriate Lexile reading ranges. Measurable goals can be clearly stated in terms of Lexile reading measures. Examples of measurable goals and clearly related strategies for reading intervention programs might include:

> *Example Goal 1:* At least half of the students will improve reading comprehension abilities by 100L after one year of use of an intervention.
>
> *Example Goal 2: S*tudents' attitudes about reading will improve after reading 10 books at their 75% comprehension level.

These examples of goals emphasize the fact that the Lexile Framework for Reading is not an intervention, but a tool to help educators plan instruction and measure the success of the reading program.

Summary

This report shows how a link has been established between NC READY EOG Reading/EOC English II scale scores and Lexile reading measures, permitting students to be matched with books and texts that provide an appropriate level of challenge while avoiding frustration. Students can be matched with texts that they are forecasted to read with 75% comprehension. It is anticipated that as a result of this purposeful match, students will read more, and thereby, read better. Wherever the student may be in the development of his or her reading skills, the Lexile Framework for Reading can be used to examine his or her growth. As a student grows, he or she can be matched with more demanding texts, thus facilitating additional growth.

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. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Biber, D. (1988). *Variation across speech and writing.* Cambridge, England: Cambridge University Press.

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.

Boyer, M. & Keng, L. (2021). *Getting ahead of the curve: Planning for accurate equating in 2021.* Center for Assessment. https://www.nciea.org/blog/school-disruption/getting-ahead-curve-planning-accurate-equating-2021

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.

Dorans, N. J., & Holland, P. W. (2000). Population invariance and the equatability of tests: Basic theory and the linear case. *Journal of Educational Measurement*, *37*, 281–306.

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.

Ehri, L. C., & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading and Writing Quarterly: Overcoming Learning Difficulties*, *14*, 135-163.

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychology*, *93*, 3-22.

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

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

Graesser, A. C., & McNamara, D. S. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher, 40,* 223-234.

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher*, *40*, 223-234.

Grolier, Inc. (1986). *The electronic encyclopedia*. Danbury, CT: Author.

Haladyna, T. M., (1994). *Developing and Validating Multiple-Choice Test Items*. Hillsdale, NJ. Lawrence Erlbaum Associates.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.

Klare, G.R. (1963). *The measurement of readability*. Ames, IA: Iowa State University Press.

Kolen, M. (1991). Smoothing methods for estimating test score distributions. *Journal of Educational Measurement, 28*(3), 257-282.

Kolen, M. J. & Brennan, R. L. (2004). *Test equating, scaling, and linking: methods and practices* (2nd ed.). New York: Springer-Verlag.

Kolen, M.J. & Brennan, R.L. (2014). *Test equating, scaling, and linking: Methods and practices*. 3rd edition. New York: Springer Science + Business Media, LLC.

Koons, H., Elmore, J., Sanford-Moore, E., Stenner, A.J. (2017). The relationship between Lexile text measures and early grades Fountas & Pinnell reading levels. MetaMetrics: Durham, NC.

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 3.70) [Computer Program]. Chicago: Author.

Lord, F. (1965). A strong true score theory, with applications. *Psychometrika, 30*(3), 239-270.

McGraw-Hill Book Company. (1983). *Guidelines for bias-free publishing*. Monterey, CA: Author.

Merlini Barbaresi, L. M. (2003). Towards a theory of text complexity. In L. Merlini Barbaresi (Ed.), *Complexity in language and text* (pp. 22-66). Pisa, Italy: Edizioni Plus.

Mesmer, H. A., Cunningham, J. W., & Hiebert, E. H. (2012). Toward a theoretical model of text complexity for the early grades: Learning from the past, anticipating the future. *Reading Research Quarterly*, *47*, 235-258.

MetaMetrics, Inc. (2006). *Lexile Vocabulary Analyzer. Technical report*. Durham, NC: Author.

MetaMetrics, Inc. (2007). *Texas Higher Education Coordinating Board: Text Measurement and Analysis. Technical report*. Durham, NC: Author.

MetaMetrics, Inc. (2008). *Texas Higher Education Coordinating Board: Text Measurement and Analysis (Update). Technical report*. Durham, NC: Author.

Miller, G.A. & Gildea, P.M. (1987). How children learn words. *Scientific American*, *257*, 94-99.

Muter, V., Hulme, C., Snowling, M. J., Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology*, *40*, 665-681.

National Council on Measurement in Education. (2021). *COVIDS impact on assessments – An NCME webinar series. https://www.ncme.org/meetings/covidwebinarseries*

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 Institute of Child Health and Human Development. (2000). *Report of the National Reading Pane. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). Washington, CD: U.S. Government Printing Office.

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

North Carolina Department of Instruction. (2021a). *North Carolina Extended Content Standards English Language Arts K-12*. Retrieved on June 28, 2021 from https://www.dpi.nc.gov/media/3996/open

North Carolina Department of Instruction. (2021b). *Standard Course of Study*. Retrieved on June 28, 2021 from https://www.dpi.nc.gov/districts-schools/classroom-resources/k-12-standards-curriculum-and-instruction/standard-course-study

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.

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.

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica: RAND Corporation.

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.

Vadasy, P. F., Sanders, E. A., Peyton, J. A. (2005). Relative effectiveness of reading practice or word-level instruction in supplemental tutoring: How text matters. *Journal of Learning Disabilities, 38*, 364-382.

Webb, X., Sanford-Moore, E., Koons, H., Baker, R., Hinson, A., Pringle, K. & Thorpe, R. (2021). *Extending the measurement of the reading construct on the Lexile scale: Foundational reading skills item study (under review)*. Durham, NC: MetaMetrics, Inc.

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

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.

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.

Appendix A

NC EOG Reading and NC EOC English II Scale-Lexile Link

This addendum specifies the procedures used to link scores from the North Carolina End-of-Grade (NC EOG) Reading Tests and NC End-of-Course (EOC) English II Test with the Lexile Framework for Reading. A description of the NC EOG Reading and NC EOC English II scale and its relationship with North Carolina's previous edition of the state assessment program is provided. Then, the methods used to establish the link are described. Finally, results from the 2021 administration are examined.

The study was conducted by MetaMetrics with the North Carolina Department of Public Instruction (contract dated July 27, 2018). The primary purposes of this study were to:

- link the NC EOG Reading and NC EOG English II scales with the Lexile Framework for Reading;
- develop a correspondence table for converting NC EOG Reading/NC EOC English II scale scores to Lexile reading measures;
- present a solution for matching students with text;
- provide tools (e.g., Lexile Find A Book) and information that can be used to answer questions related to standards, test score interpretation, and test validation; and
- produce a report that describes the linking analysis procedures.

The NC EOG Reading Tests and NC EOC English II Test were revised to reflect the North Carolina State Board of Education adoption of revisions to the Standard Course of Study in 2017 (NCDPI, 2021b). The NC EOG Reading Tests and EOC NC English II Test are in their 5th edition. Implementation of the 5th edition was scheduled for the school year 2019-2020. Due to the COVID-19 pandemic, full scale administration was delayed until school year 2020-2021. During the inaugural administration a study was planned to link the NC EOG Reading Test and the NC EOC English II Test. A delay of the full scale administration of the revised tests resulted in a delay in the linking study planned with the Lexile Framework for Reading.

Study Design and Linking the NC EOG Reading/EOC English II Scale Scores with the Lexile Scale

A single-group/non-equivalent anchor test design was originally chosen for this study (Dorans and Holland, 2000). 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 (Kolen and Brennen, 2014, pp. 16–17). During the 2019-2020 school year testing was postponed due to the pandemic and the study design was reconsidered.

Student educational experiences during the 2020-2021 pandemic school year were atypical compared to any other school year. A study conducted during this time may have different outcomes compared to the previous 2013 linking study solely due to this factor. In the planned

2021 study, the linking items were to be appended. This has the potential to lead to item sequence effects, such as motivation or fatigue. Due to multiple confounds, a study conducted as a single groups design might have different outcomes and may not generalize beyond this particular school year and was therefore re-evaluated.

MetaMetrics and NC DPI designed and conducted the linking study using pre-equated item parameters from the 2019 NC Ready EOG Reading/EOC English II under advisement from the NC DPI technical advisory committee (TAC; NCDPI, 2020). The 2019 NC Ready EOG Reading and NC EOC English II had an established link from the 2013 linking study. In addition to recommendations from the NC DPI TAC, the use of pre-equated measures was in accordance with recommendations from multiple professional organizations such as National Council of Measurement in Education and The Council of Chief State School Officers (Boyer & Keng, 2021; NCME, 2021).

NC EOG Reading Tests in Grades 3 through 5 consist of 40 operational items, and in Grades 5 through 8 consist of 44 operational items. The NC EOC English II Test consists of 51 operational items. Each NC EOG Reading Test and NC English II Test are scaled horizontally ranging from 500 to 600. It is important to note that, even though the reported NC EOG Reading Test scale ranges for the 2021 version are similar to the previous edition, the reported scale scores do not have the same meaning between editions.

Using the pre-equated item parameters, the 2021 tests were provided to MetaMetrics on both the 2019 and 2021 reporting scales. A one-to-one relationship between these scales did not exist at every scale score point. In these occurrences, to link the 2021 edition scale to the 2019 edition scale, MetaMetrics found the average 2019 score value associated with each 2021 reported score. This was done separately for each grade-based scale and for the NC EOC English II scale. The Lexile linking formula established in 2013 was applied to these average 2019 scores. This provided a direct correspondence of Lexile reading measures between the 2019 NC READY EOG Reading and NC EOC English II scales with the 2021 NC EOG Reading and NC EOC English II scales, respectively. Then a concordance table was established between the 2021 NC EOG Reading/EOC English II scale scores and the Lexile scale. The concordance table is an optimal solution in this scenario as the property of symmetry between the NC EOG Reading and NC EOC English II scale scores and the Lexile scale is maintained.

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

This section provides evidence to support the link between the NC EOG Reading and NC EOC English II scales and the Lexile scale. All results originated from the statewide administration and are population statistics. A total of 651,028 student records across all test levels were provided by NC DPIT. A basic description of the population is provided. The results from the concordance are applied to the student population for the NC EOG Reading and NC EOC English II. Then descriptive information is displayed to illustrate the benefit of portraying the NC horizontal scale in terms of a vertical scale, the Lexile scale. A comparison is made to the Lexile user norms. Lastly, NC EOG Reading and NC EOC English II achievement level

descriptors are translated into Lexile reading measures and compared with MetaMetrics research on text complexity across grades.

Table A. 1 describes the demographics of the student population administered the NC EOG Reading or NC English II Tests. In total, 651,028 students were administered a test as part of the state testing program. Each test level represents similar percentages of the population ranging from 11.7% to 16.5% with a generally increasing enrollment as the Grade or Test Level increase.

Table A. 1. Percentage of students in the NC EOG Reading/EOC English II for selected demographic characteristics, N = 651,028.

Student Characteristic	Category	Percent of Population
Test Level	Grade 3	11.7
	Grade 4	12.1
	Grade 5	13.0
	Grade 6	15.3
	Grade 7	15.7
	Grade 8	15.8
	English II	16.5
Gender	Female	40.9
	Male	42.6
	Not Available	16.5
Ethnicity	American Indian or Alaska Native	1.1
	Asian	2.7
	Black	20.2
	Hispanic/Latino	17.1
	Native Hawaiian or Other Pacific Islander	0.1
	Two or more races	4.2
	White	37.8
	Not Available	16.7
Economically Disadvantaged Student	Yes	38.6
	No	60.9
	Not Available	0.6
English Language Development	1	0.2
	2	0.6
	Yes	7.4
	No	88.7
	Not Available	3.0
Students with Disabilities	Yes	11.4
	No	84.6
	Not Available	4.0
Geographic Location	City	37.1
	Not Available	0.2
	Rural	47.0
	Suburb	8.3
	Town	7.4

Grade-Level Progressions. *Table A. 2* presents the descriptive statistics of the NC EOG Reading/EOC English II test scale scores as well as the linked Lexile reading measures for the state testing population. The NC EOC Reading/EOG English II is on a horizontal scale as illustrated by the similar means and standard deviations (SD) observed across test levels, while the Lexile scale is on a vertical scale illustrated by the continually increasing mean as test level increases.

Table A. 2. Descriptive statistics for the NC EOG Reading and NC EOC English II scale scores and the Lexile reading measures, N = 651,028.

Test Level	*N*	NC EOC Reading/EOG English II Scale Score Mean (SD)	Lexile Reading Measure Mean (SD)
Grade 3	75,989	537 (9.8)	647L (237.9)
Grade 4	78,622	542 (10.0)	785L (236.0)
Grade 5	84,566	547 (9.7)	911L (227.5)
Grade 6	99,736	550 (9.7)	980L (239.5)
Grade 7	102,181	552 (9.9)	1031L (247.5)
Grade 8	102,729	556 (9.8)	1119L (244.9)
English II	107,205	550 (9.5)	1263L (251.3)

The box-and-whisker plots in *Figure A. 1* shows the horizontal nature of the NC EOG Reading and NC EOC English II scale scores (y-axis) across test levels (*x*-axis). *Figure A. 2* shows the progression of the NC EOG Reading and NC English II scale scores expressed as Lexile reading measures (the *y*-axis) across Grades 3-8 and English II (the *x*-axis). For each grade, the box refers to the interquartile range, the line within the box indicates the median, the plus symbol indicates the mean, and the trend line (*Figure A. 2*) connects each box at the median. The end of each whisker represents the 5th and 95th percentile values of the scores (the *y*-axis).

The NC EOG Reading and NC EOC English II scale is a horizontal scale, while the Lexile scale is a vertical scale. *Figure A. 1* demonstrates the horizontal nature of the NC EOG Reading and NC EOC English II scale as seen by the similar score ranges observed across grades. *Figure A. 2* demonstrates the vertical nature of the Lexile scale. The scores in *Figure A. 2* increase as grade level increases and the score distributions for adjacent grades overlap. The “overlap across grades” is characteristic of vertical scales (Kolen & Brennan, 2014).

Figure A. 1. Box-and-whisker plots of Grades 3–8 NC EOG Reading and NC EOC English II scale scores.

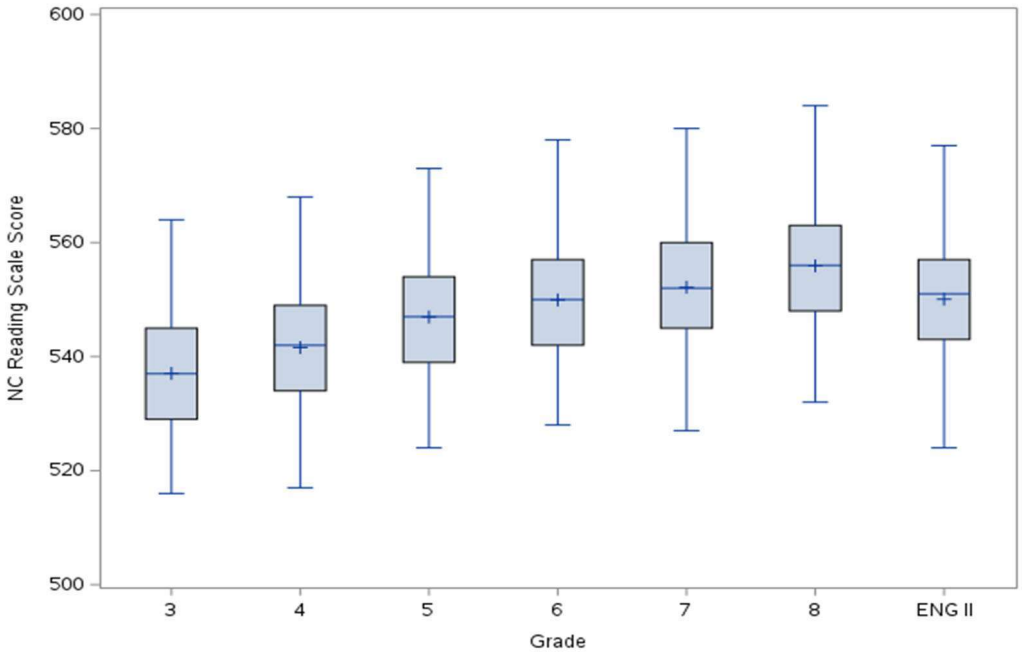
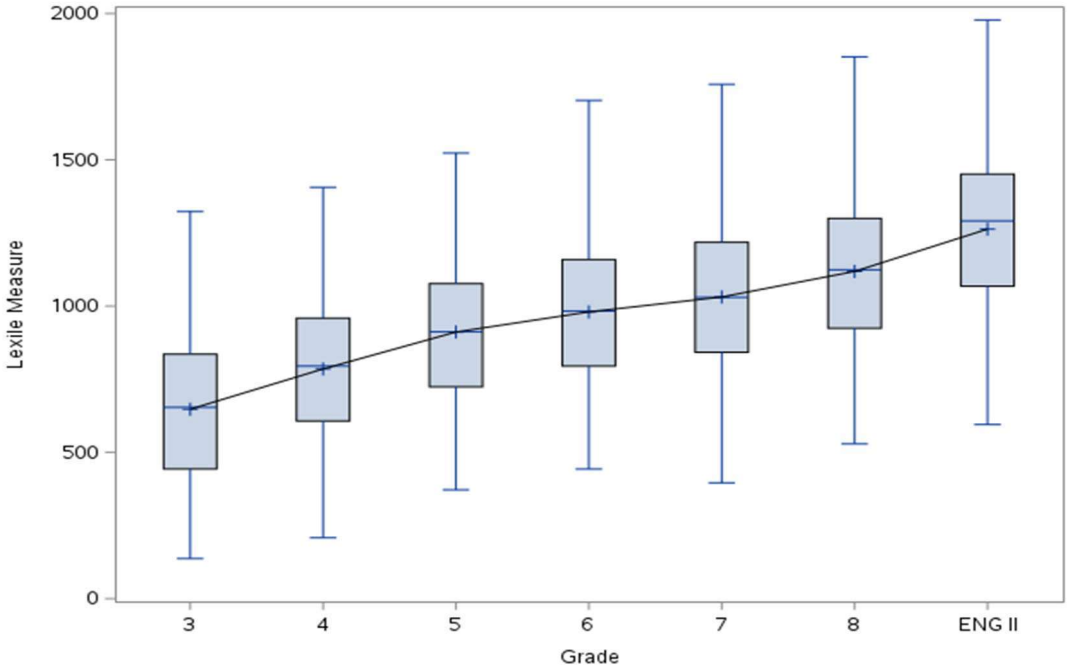


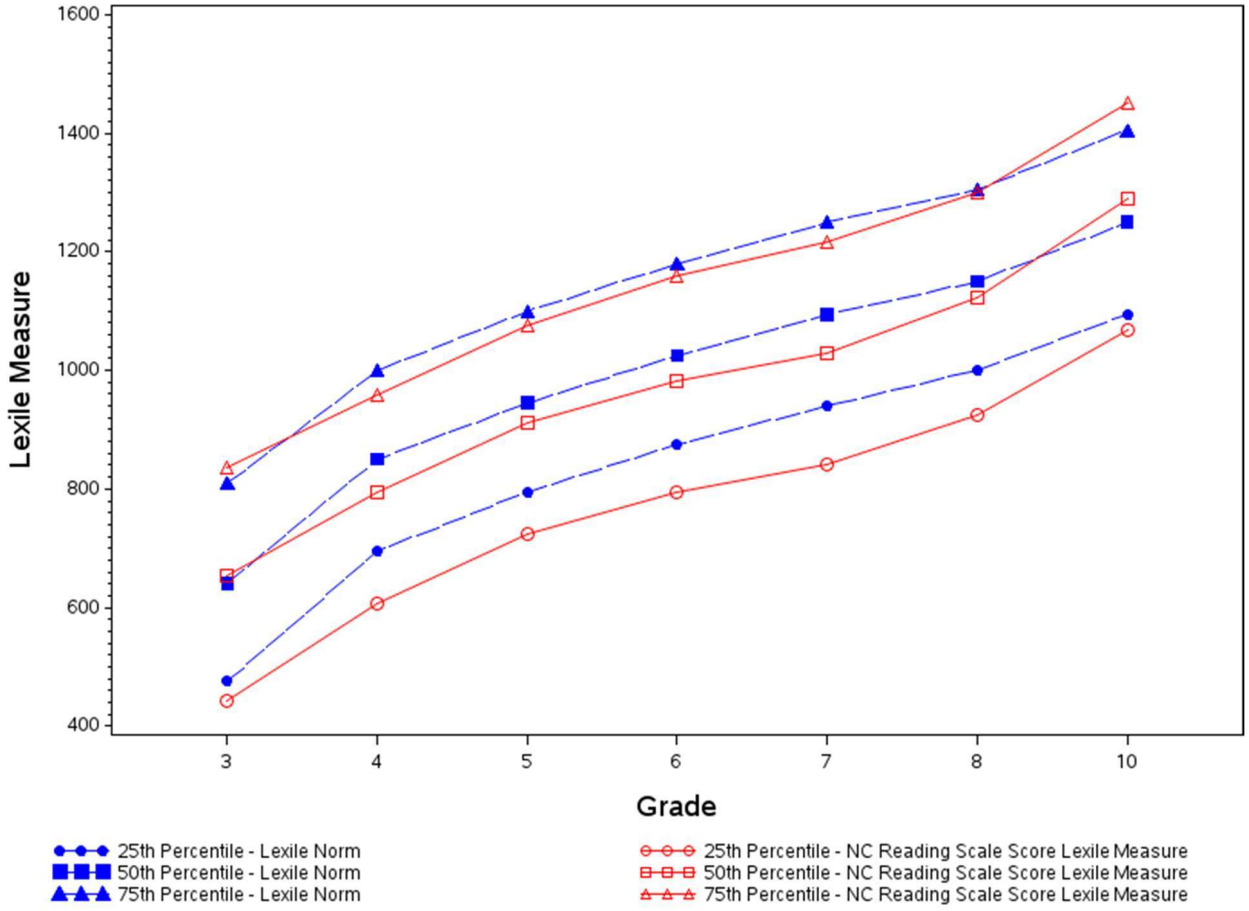
Figure A. 2. Box-and-whisker plots of Grades 3–8 NC EOG Reading and NC EOC English II expressed as Lexile reading measures.



The Lexile Framework for Reading Norms. *Figure A. 3* shows the NC EOG Reading and NC EOC English II Lexile reading measures and the Lexile reading use norms. The normative information for the Lexile Framework for Reading is based on linking studies conducted with the Lexile Framework and the results of assessments that report directly in the Lexile metric (N = 3,888,110). The sample included students in Kindergarten through Grade 12 from 51 states, districts, or territories and who were tested from 2010 to 2016 (Grades 1-12) and 2016 to 2019 (Kindergarten). Of the students with gender information (45%), 51.6% of the students were male and 48.4% of the students were female. Of the students with race or ethnicity information (30.2%), the majority of the students in the norming sample were White 56.3%, with 5.8% African-American, 2.0% American Indian/Alaskan Native, 14.7% Hispanic, 14.7% Asian, and 5.2% Other. Of the students with data indicating English proficiency and/or special education status (2.9), 7.0% of the students were classified as "limited English proficient"; and 9.1% of the students were classified as "Needing Special Education Services." Of the students with data on eligibility for free or reduced price lunch (2.9%); 45.9% of the students were eligible for the free or reduced-price lunch program.

The NC EOG Reading and NC EOC English II scale scores as expressed in the Lexile metric are very similar to the 25th, 50th, and 75th percentiles trends across the grade ranges. In Grade 3 and English II the 75th and 50th percentiles are slightly above the Lexile user norms and the 25th percentile is slightly lower than the Lexile user norms. For the remainder of Grades, the selected percentiles for NC EOG Reading are slightly below the Lexile user norms. Overall, the NC EOG Reading and NC English II show very similar patterns with that of the Lexile user norms.

Figure A. 3. Selected percentiles (25th, 50th, and 75th) plotted for the NC EOG Reading and NC EOC English II Lexile reading measures in relation to the Lexile reading measure norms.



NC EOG Reading and NC EOC English II Achievement Level Descriptors. The NC EOG Reading and NC English II scales are divided into four achievement levels. These four achievement levels and associated cut points are used to describe student results: Not Proficient, Level 3, Level 4, and Level 5. Students whose score is at or above the Level 3 cut point are identified as having demonstrated proficiency in grade-level skills and grade-appropriate materials. Student performances that do not reach the cut score for Level 3 are not considered to have met grade level proficiency.

Each achievement level descriptor is defined as the following (NC DPI, 2020; pp. 1-5)

> **Not Proficient**: Students who are not proficient demonstrate **inconsistent** understanding of grade level content standards and will need support at the next grade/course.
>
> **Level 3**: Students at Level 3 demonstrate **sufficient** understanding of grade level content standards though *s*ome support may be needed to engage with content at the next grade/course.
>
> **Level 4**: Students at Level 4 demonstrate a **thorough** understanding of grade level content standards and are on track for career and college.
>
> **Level 5**: Students at Level 5 demonstrate **comprehensive** understanding of grade level content standards, are on track for career and college, and are prepared for advanced content at the next grade/course.

Table A. 3 provides the scale score and Lexile reading measure ranges for each of the achievement levels.

Table A. 3. NC EOG Reading and NC EOC English II achievement level scale score ranges and associated Lexile reading measures.

	Not Proficient		Level 3		Level 4		Level 5	
Grade /Test Level	**NC EOG/EOC Scale Score Range**	**Lexile Measure Range**	**NC EOG/EOC Scale Score Range**	**Lexile Measure Range**	**NC EOG/EOC Scale Score Range**	**Lexile Measure Range**	**NC EOG/EOC Scale Score Range**	**Lexile Measure Range**
3	515-539	115L-700L	540-545	725L-835L	546-550	860L-965L	551-564	985L-1200L
4	517-543	210L-820L	544-547	840L-910L	548-555	935L-1100L	556-568	1125L-1300L
5	524-549	370L-960L	550-553	985L-1055L	554-559	1075L-1195L	560-573	1220L-1400L
6	528-551	445L-1005L	552-557	1030L-1160L	558-566	1180L-1380L	567-578	1400L-1500L
7	527-553	395L-1055L	554-558	1075L-1180L	559-565	1195L-1355L	566-580	1370L-1600L
8	532-556	530L-1125L	557-562	1145L-1265L	563-571	1300L-1500L	572-584	1515L-1700L
Eng. II	524-548	595L-1210L	549-554	1240L-1370L	555-564	1405L-1630L	565-577	1655L-1980L

College and Career Reading Demands. MetaMetrics has conducted research on the reading demands that are typically associated with college and career readiness (CCR) and developed a Lexile-based reading text complexity range for each grade band. *Figure A. 4* shows the relationship between the "Level 3" achievement level/proficiency standard for each test level established on the NC EOG Reading and NC EOC English II Tests and the "stretch" reading demands. At each grade, the lowest score in the Level 3 range is the cut point and the highest score in the Level 3 range is the last score before the Level 4 cut point, with a dashed line connecting them. *Figure A. 4* helps contextualize the proficiency level set by the North Carolina Department of Public Instruction by showing that students classified as "Level 3" and above on the NC EOG Reading and NC EOC English II should be able to read text that they are likely to encounter as they prepare for college and careers.

Figure A. 4 Comparison of NC EOG Reading and NC EOC English II "Level 3" achievement level with college and career reading levels.

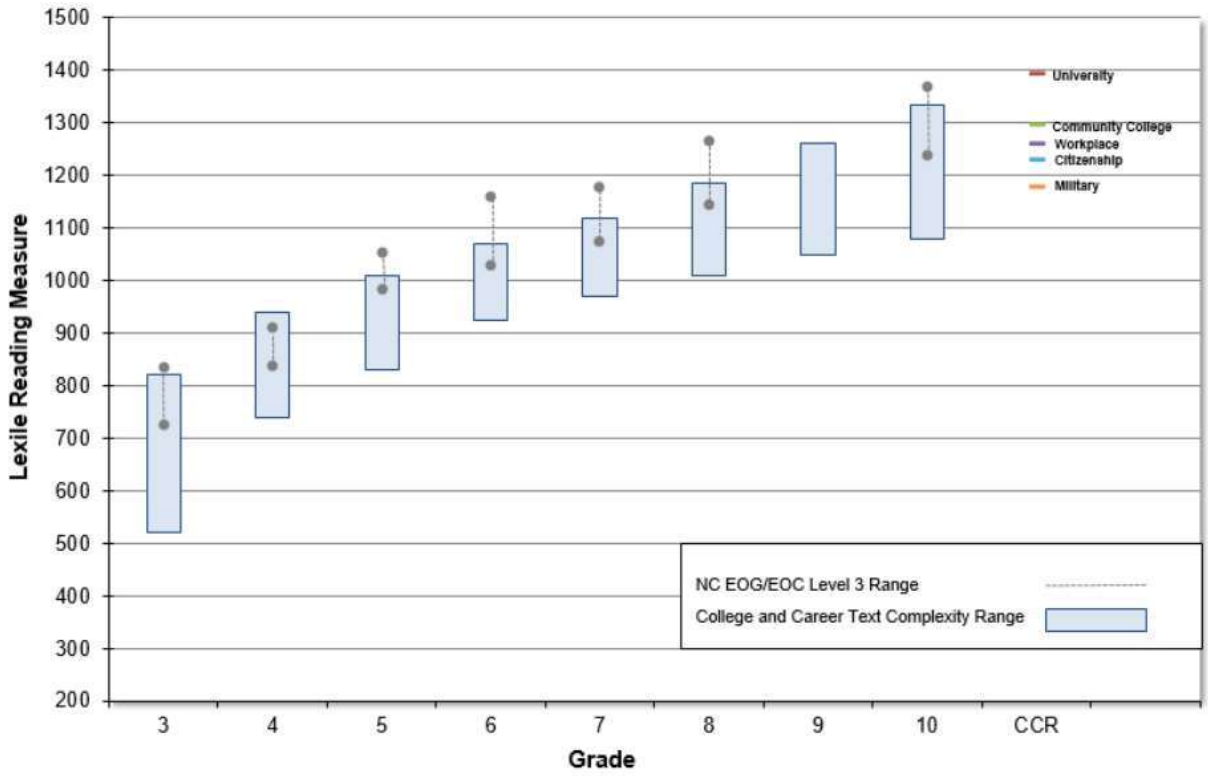
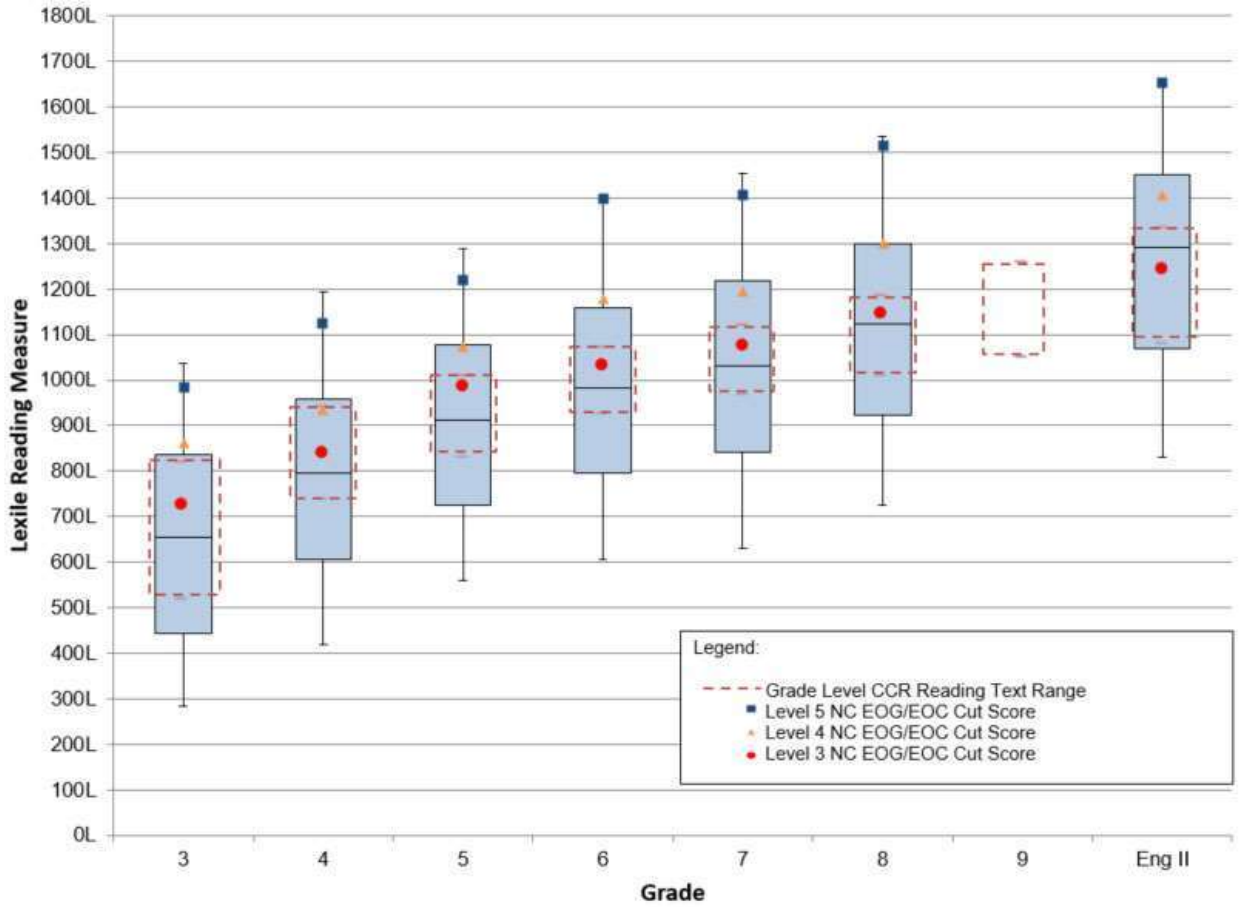


Figure A. 5 compares the distribution of 2020-2021 student performance expressed as Lexile reading measures (blue boxes) to the North Carolina achievement level cut scores and college and career reading levels. For each test level, the blue box refers to the interquartile range. The line within the blue box indicates the median. The end of each whisker represents the 5th percentile at the low end and the 95th percentile at the high end of the distribution of students' Lexile reading measures. For each grade/test level the achievement level cut scores are provided. Across grades, most student scores fall within or above the CCR text demand ranges. For each grade/test level the median student score is below the Level 3 achievement level. Combining student results with criterion referenced indicators provides information to reference when matching students with reading texts. Example texts across the Lexile range are shown on the Lexile Map provided in Appendix B.

Figure A. 5. NC EOG Reading and NC EOC English II 2020-2021 student performance expressed as Lexile reading measures overlayed with the achievement level descriptors and grade level CCR reading text ranges.



Conclusions

The purpose of this study was to establish a linkage between the scores on the NC EOG Reading for Grades 3 through 8 and NC English II scales with the Lexile scale. A concordance table was constructed leveraging the existing link with NC READY EOG and NC EOC English II established in 2013. Due to pandemic interruptions during the 2019-2020 and 2020-2021 school years, pre-equated item parameters were used to extend the established link from edition 4 to edition 5 as recommended by NC DPI's technical advisory committee.

The comparison of NC EOG Reading and NC EOC English II scale scores and the Lexile reading norm values showed similar trajectories throughout the grade level progressions. Grade-level progression box plots showed that Lexile reading measures increased as grade increased, providing evidence for the added interpretation from the vertical scale of the Lexile Framework.

NC EOG Reading and NC EOC English II student performance in the Lexile metric and the NC achievement levels as set by NC DPI were explored in relation to established CCR reading text complexity ranges. The "Level 3" performance range was at or above the CCR reading text complexity ranges for each grade and at English II indicating that students classified as "Level 3" are projected to be able to read text that they are likely to encounter as they prepare for college and careers.

To utilize the results from this study, Lexile reading measures need to be incorporated into the NC EOG Reading and NC English II results processing and interpretation frameworks. This information can then be used in a variety of areas within the educational system—instruction, assessment, and communication to name a few. Once a linkage has been established between a target test and the Lexile scale, educators will be able to utilize the assessment results, reported in Lexile reading measures, to inform classroom instruction. The following sections provide a more detailed description about the caveats associated with the study followed by recommended uses of The Lexile Framework for Reading and associated tools.

Limitations: Lexile Reading Measures and Grade Levels. Lexile reading measures do not translate specifically to grade levels. Within any grade, there will be a range of readers and a range of materials to be read. In a sixth-grade classroom there will be some readers who are far ahead of the others and there will be some readers who are behind the others in terms of reading ability. To say that some books are "just right" for sixth graders assumes that all sixth graders are reading at the same level. The Lexile Framework for Reading can be used to match readers with texts at whatever level the reader is reading.

Simply because a student is an excellent reader, it should not be assumed that the student would necessarily comprehend a text typically found at a higher grade level. Without adequate background knowledge, the words may not have sufficient meaning to the student. A high Lexile reading measure for a grade indicates that the student can read grade-appropriate materials at a higher comprehension level (90%, for example).

Maintenance of the NC EOG Reading and NC English II scale. Maintenance of the focal scale is critical to the validity of any link with an auxiliary scale (i.e., Lexile scale). If an update occurs to the focal scale, the integrity of the link should be re-evaluated and additional linking studies may be needed to accommodate fundamental changes to the focal scale. Such updates may include, but are not limited to, incorporating new item types into the assessment; revising item calibrations; or revising the assessment program and the reported scale scores.

Linking error. Error in estimating the linking relationship of two scales is present whenever linking is conducted. Not all error associated with a study can be accounted for, however error should be continually investigated to ensure scores are as accurate and reliable as possible. The two sources of error present are random error and systematic error. Random linking error occurs when directly estimating the linking relationship because a sample is collected to perform the study. Systematic error occurs when estimation methods introduce bias, statistical assumptions for the methods are not met, improper sampling techniques were used to collect the data for the linking study, or different placement of items impacts scale scores. To the extent possible MetaMetrics and NC DPI worked to minimize systematic error through the design of the linking study.

Because this linking study was conducted during a school year in which unprecedented challenges faced students and educators due to COVID-19 and the policies put in place to ensure safe learning environments, potential implications were considered throughout the linking process.

Next Steps. To utilize the results from this study, Lexile reading measures need to be incorporated into the NC EOG Reading and NC 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, and communication to name a few. Once a linkage has been established between a target test and the Lexile scale, educators will be able to utilize the assessment results, reported in Lexile reading measures, to inform classroom instruction.

Appendix B

The Lexile® Framework for Reading Map

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 fab.lexile.com.

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 early 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.

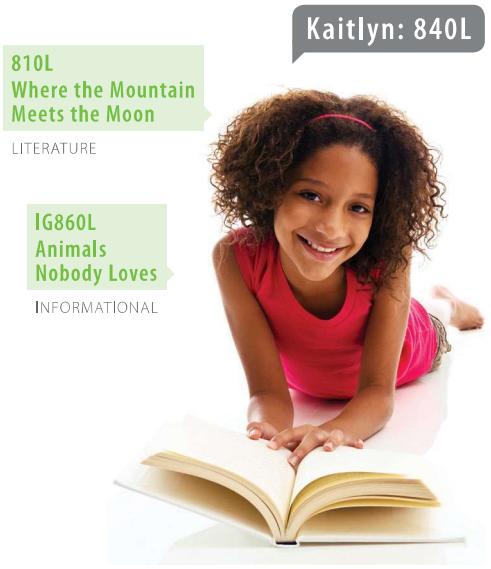
HOW TO USE IT

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 fab.lexile.com to access the "Find a Book" tool. "Find a Book" enables users to search from over 275,000 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+

1630L ***Descartes: Philosophical Essays*** LAFLEUR

But neither should we fall into the error of those who occupy their minds only with deep and serious matters, of which, after much effort, they acquire only a confused knowledge, while they hoped for a profound one. It is therefore in these easier matters that we should first exercise our minds, but methodically, so that we become accustomed to penetrate each time, by open and recognized paths and almost as in a game, to the inner truth of things. In this way, soon afterward, and in less time than one could hope, we will find ourselves able to deduce with equal ease and from self-evident principles, many propositions which appear very difficult and intricate. But perhaps some will be astonished that in this study, where we are inquiring how we can be made more competent to deduce some truths from others, we omit all the rules by which the logicians think they regulate human reason. These prescribe certain forms of argument which involve such necessary implications that the mind which relies upon this method, even though it neglects to give clear and attentive consideration to the reasoning, can nevertheless reach certain conclusions on the strength of the form of the argument alone.

SAMPLE TITLES

LITERATURE

- 1640L **The Plot Against America** (ROTH)
- 1530L **The Good Earth** (BUCK)
- 1520L **A Fable** (FAULKNER)

INFORMATIONAL

- 1650L **Twenty Years at Hull-House** (ADDAMS)
- 1600L **The U.S. Constitution and Other Key American Writings** (ASSORTED)
- 1600L **Sustaining Life: How Human Health Depends on Biodiversity** (CHIVIAN)
- 1590L **Captain John Smith: A Select Edition of His Writings** (SMITH)
- 1520L **Collapse: How Societies Choose to Fail or Succeed** (DIAMOND)
- 1510L **Original Meanings: Politics and Ideas in the Making of the Constitution** (RAKOVE)

1400L–1495L

1440L ***Fordlandia*** GRANDIN

As Ford biographer Robert Lacey put it, the "Five Dollar Day raised the pain threshold of capitalism." But beyond an incentive to make workers stay put, it also became a model for how to respond to another crisis that plagued industrialism. The mechanized factory production that took flight during America's Gilded Age had promised equality and human progress but in reality delivered deepening polarization and misery, particularly in sprawling industrial cities like Detroit. Ford, advised by farsighted company executives such as James Couzens and John Lee, understood that high wages and decent benefits would do more than create a dependable and thus more productive workforce; they would also stabilize and stimulate demand for industrial products by turning workers into consumers.

SAMPLE TITLES

LITERATURE

- 1460L **The Legend of Sleepy Hollow** (IRVING)
- 1450L **Billy Budd** (MELVILLE)
- 1420L **The Life All Around Me by Ellen Foster** (GIBBONS)
- 1420L **The Fall of the House of Usher** (POE)
- 1410L **Death in Venice** (MANN)

INFORMATIONAL

- 1490L **Rousseau's Political Writings** (ROUSSEAU)
- 1430L **America's Constitution: A Biography** (AMAR)
- 1410L **Profiles in Courage** (KENNEDY)
- 1400L **The Mysteries of Beethoven's Hair** (MARTIN & NIBLEY)
- 1400L **Life and Times of Frederick Douglass: His Early Life as a Slave, His Escape From Bondage, and His Complete History to the Present Time** (DOUGLASS)

1300L–1395L

1340L ***Silent Spring*** CARSON

The basic element, carbon, is one whose atoms have an almost infinite capacity for uniting with each other in chains and rings and various other configurations, and for becoming linked with atoms of other substances. Indeed, the incredible diversity of living creatures from bacteria to the great blue whale is largely due to this capacity of carbon. The complex protein molecule has the carbon atom as its basis, as have molecules of fat, carbohydrates, enzymes, and vitamins. So, too, have enormous numbers of nonliving things, for carbon is not necessarily a symbol of life.

SAMPLE TITLES

LITERATURE

- 1390L **The Yellow Wallpaper** (GILMAN)
- 1350L **The Secret Sharer** (CONRAD)
- 1330L **The Jungle** (SINCLAIR)
- 1330L **Silas Marner** (ELIOT)
- 1300L **Gulliver's Travels** (SWIFT)

INFORMATIONAL

- 1390L **In Defense of Food: An Eater's Manifesto** (POLLAN)
- 1360L **Anne Frank: The Book, the Life, the Afterlife** (PROSE)
- 1340L **Walden and Civil Disobedience** (THOREAU)
- 1330L **The Professor and the Madman: A Tale of Murder, Insanity, and the Making of the Oxford English Dictionary** (WINCHESTER)
- 1300L **Arctic Dreams: Imagination and Desire in a Northern Landscape** (LOPEZ)

1200L–1295L

1210L ***The Tortilla Curtain*** BOYLE

He didn't wake America, not yet. He made four trips up to the ledge and back, with the tools, the sacks of vegetables—they could use the empty sacks as blankets, he'd already thought of that—and as many wooden pallets as he could carry. He'd found the pallets stacked up on the far side of the shed, and though he knew the maintenance man would be sure to miss them, it could be weeks before he noticed and then what could he do? As soon as Qindido had laid eyes on those pallets an architecture had invaded his brain and he knew he had to have them. If the fates were going to deny him his apartment, well then, he would have a house, a house with a view.

SAMPLE TITLES

LITERATURE

- 1290L An Old-Fashioned Girl (ALCOTT)
- 1280L The House of the Spirits (ALLENDE)
- 1280L The Castle (KAFKA)
- 1220L The Silent Cry (ŌE)
- 1210L Chronicle of a Death Foretold (GARCÍA MÁRQUEZ)

INFORMATIONAL

- 1290L A Brief History of Time: From the Big Bang to Black Holes (HAWKING)
- 1280L Black, Blue, and Gray: African Americans in the Civil War (HASKINS)
- 1230L Stiff: The Curious Lives of Human Cadavers (ROACH)
- 1230L Knowing Mandela: A Personal Portrait (CARLIN)
- 1200L The Dark Game: True Spy Stories (JANECZKO)

1100L–1195L

1150L ***A Room of One's Own*** WOOLF

The reason perhaps why we know so little of Shakespeare—compared with Donne or Ben Jonson or Milton—is that his grudges and spites and antipathies are hidden from us. We are not held up by some "revelation" which reminds us of the writer. All desire to protest, to preach, to proclaim an injury, to pay off a score, to make the world the witness of some hardship or grievance was fired out of him and consumed. Therefore his poetry flows from him free and unimpeded. If ever a human being got his work expressed completely, it was Shakespeare. If ever a mind was incandescent, unimpeded, I thought, turning again to the bookcase, it was Shakespeare's mind.

SAMPLE TITLES

LITERATURE

- 1180L Sense and Sensibility (AUSTEN)
- 1170L The Amazing Adventure of Kavalier & Clay (CHABON)
- 1150L Great Expectations (DICKENS)
- 1140L Cold Mountain (FRAZIER)
- 1130L Democracy (DIDION)

INFORMATIONAL

- 1160L The Longitude Prize (DASH)
- 1160L In Search of Our Mothers' Gardens (WALKER)
- 1150L The Human Microbiome: The Germs That Keep You Healthy (HIRSCH)
- 1150L In My Place (HUNTER-GAULT)
- 1100L Something to Declare (ALVAREZ)

1000L–1095L

1070L ***Geeks: How Two Lost Boys Rode the Internet out of Idaho*** KATZ

Geeks were the first to grasp just how much information was available on the Web, since they wrote the programs that put much of it there—movie times and reviews, bus and train schedules, news and opinions, catalogues, appliance instructions, plus, of course, software and its upgrades. And of course, music, the liberation of which is considered a seminal geek accomplishment.

Virtually everything in a newspaper—and in many magazines—is now available online. In fact, some things, like the latest weather and breaking news, appear online hours before they hit print.

Yet while Jesse had gone through literally thousands of downloaded software applications, he'd never paid for any of them. He didn't even quite get the concept. The single cultural exception was books. Perhaps as a legacy of his childhood, Jesse remained an obsessive reader. He liked digging through the bins of used bookstores to buy sci-fi and classic literature; he liked books, holding them and turning their pages.

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

- 1030L Phineas Gage: A Gruesome but True Story About Brain Science (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 Harriet Tubman: Conductor on the Underground Railroad (PETRY)
- 1000L Claudette Colvin: Twice Toward Justice (HOOSE)

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: An American Legend (HILLENBRAND)
980L 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 (FINLAYSON)
920L Talking With Artists (CUMMINGS)

800L–895L

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

We tiptoed down the hall to the second classroom on the right. The heavy wooden door opened easily and we stepped in. There is an eerie, expectant feeling to a schoolroom in the summer. The normal classroom items were there: desks, chalkboards, a set of encyclopedias. The American flag with accompanying pictures of Presidents Washington and Lincoln. But without students occupying those desks and their homework tacked on the wall, that empty summer classroom seemed laden with the memory of past students and past learning that took place within those walls. I strained to listen, as if I might hear the whisperings and stirrings of the past. Maybe Ruthanne was right. Maybe there was more here than met the eye.

SAMPLE TITLES

LITERATURE

GN840L* The Odyssey (HINDS)
830L Baseball in April and Other Stories (SOTO)
820L Maniac Magee (SPINELLI)
810L Where the Mountain Meets the Moon (LIN)
800L Homeless Bird (WHELAN)

INFORMATIONAL

880L Volcanoes (SIMON)
880L The Circuit: Stories From the Life of a Migrant Child (JIMÉNEZ)
IG860L* Animals Nobody Loves (SIMON)
860L Through My Eyes: Ruby Bridges (BRIDGES)
830L Quest for the Tree Kangaroo (MONTGOMERY)

700L–795L

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

Edward Tulane waited.

He repeated the old doll's words over and over until they wore a smooth groove of hope in his brain: *Someone will come; someone will come for you.*

And the old doll was right.

Someone did come.

It was springtime. It was raining. There were dogwood blossoms on the floor of Lucius Clarke's shop.

She was a small girl, maybe five years old, and while her mother struggled to close a blue umbrella, the little girl walked around the store, stopping and staring solemnly at each doll and then moving on.

When she came to Edward, she stood in front of him for what seemed like a long time. She looked at him and he looked back at her.

SAMPLE TITLES

LITERATURE

770L Walk Two Moons (CREECH)
760L Hoot (HIAASEN)
750L Esperanza Rising (RYAN)
720L Nancy's Mysterious Letter (KEENE)
GN720L* Sherlock Holmes and the Adventure at the Copper Beeches (DOYLE)

INFORMATIONAL

790L Be Water, My Friend: The Early Years of Bruce Lee (MIOCHIZUKI)
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)

600L–695L

620L ***The Year of Billy Miller*** HENKES

His heart was pounding.

Once again, he forgot every word of his poem, including the title—but this time he didn't have a copy of it to read from.

He saw Ms. Silver in the fringes of his vision. She was smiling and nodding, urging him on with her wide eyes.

Should he walk over to her to get a copy of his poem? She seemed about a mile away. And he didn't think he could make his legs move.

What should he do?

The air felt weird all of a sudden. As if it had sprouted wings and was brushing against him. The air was fluttering against his arm.

How could that be?

He turned around and Mama was there with a copy of his poem, tapping it lightly against his elbow. "Here," she whispered. "You can do it."

SAMPLE TITLES

LITERATURE

690L Firefly Hollow (MCGHEE)
680L Charlotte's Web (WHITE)
670L A Year Down Yonder (PECK)
660L Holes (SACHAR)
610L Mountain Bike Mania (CHRISTOPHER)

INFORMATIONAL

690L Sadako and the Thousand Paper Cranes (COERR)
680L An Eye for Color: The Story of Josef Albers (WING)
680L The Moon (LANDAU)
660L Remember: The Journey to School Integration (MORRISON)
620L Crittercam (EINSPRUCH)

500L–595L

500L ***The Curse of the Cheese Pyramid*** STILTON

Trap winked at me and announced, "Grandfather has hired me to be his personal cook!"

This was ridiculous! I was getting hotter than a bag of cheese popcorn in a microwave. Who would help me run the paper?

At that moment, I felt a tug on the sleeve of my jacket. It was my young nephew Benjamin. "Uncle Geronimo, guess what?" he beamed. "Great-grandfather William has hired me to be his personal assistant!"

Grandfather stroked Ben's tiny ears.

"Ah, the family, there's nothing like the family! The Stilton Family, that is..." I snorted. I could see I was the workmouse of the family. It looked like I would be the only one doing any work!

SAMPLE TITLES

LITERATURE

590L The Great Kapok Tree (CHERRY)
580L Tops and Bottoms (STEVENS)
570L Grace for President (DIPUCCHIO)
540L Ron's Big Mission (BLUE & NADEN)
500L Poppleton in Spring (RYLANT)

INFORMATIONAL

IG590L* Claude Monet (CONNOLLY)
580L What Magnets Can Do (FOWLER & BARKAN)
560L Molly the Pony (KASTER)
550L Martin Luther King, Jr. and the March on Washington (RUFFIN)
510L A Picture for Marc (KIMMEL)

400L–495L

470L ***Frog and Toad Are Friends*** LOBEL

Toad said, "Frog, you are looking quite green."
"But I always look green," said Frog. "I am a frog."
"Today you look very green even for a frog," said Toad.
"Get into my bed and rest."
Toad made Frog a cup of hot tea.
Frog drank the tea, and then he said, "Tell me a story while I am resting."
"All right," said Toad.

SAMPLE TITLES

LITERATURE

480L A Birthday for Frances (HOBAN)
470L Tales of a Fourth Grade Nothing (BLUME)
450L Amelia Bedelia (PARISH)
440L Fox on the Job (MARSHALL)
420L Hey, New Kid! (DUFFEY)

INFORMATIONAL

480L Rally for Recycling (BULLARD)
480L Grand Canyon (GILBERT)
470L Life in China (CHUNG)
460L Half You Heard of Fractions? (ADAMSON & ADAMSON)
440L Abraham Lincoln (HANSEN)

THE LEXILE® FRAMEWORK FOR READING

200L–395L+ LEXILE RANGE

300L–395L

330L ***Seals*** ARNOLD

Earless seals live in oceans.
Thick blubber keeps seals warm.
A seal's back flippers help it swim fast.
A seal on land is slow.
Its claws dig into rocks and ice.
Many seals have dark brown or gray fur.
Some have spots.
Seals molt every year.

SAMPLE TITLES

LITERATURE	370L	Little Bear Book (MINARIK)
	350L	To the Rescue! (MAYER)
	340L	Snow (SHULEVITZ)
	GN320L*	Spotlight Soccer (SANCHEZ)
	310L	I Spy Fly Guy! (ARNOLD)
INFORMATIONAL	370L	Starfish (HURD)
	IG340L*	We Can Be Friends (JORDAN)
	340L	Fernando Exercises!: Tell and Write Time (KAY)
	340L	Simple Machines (RISSMAN)
	310L	Visiting the Beach in Summer (FELIX)

200L–295L

220L ***Put Me in the Zoo*** LOPSHIRE

Look at this, now! One! Two! Three!
I can put them on a tree.
And now when I say "One, two, three"
All my spots are back on me!
Look, now!
Here is one thing more. I take my spots. I make them four.
Oh! They would put me in the zoo, if they could see what I can do.

SAMPLE TITLES

LITERATURE	290L	The Class Pet From the Black Lagoon (THALER)
	280L	Puddle (YUM)
	240L	Are You My Mother? (EASTMAN)
	210L	Green Eggs and Ham (SEUSS)
	200L	Tiny Goes to the Library (MEISTER)
INFORMATIONAL	280L	Whales (LINDEEN)
	260L	Leaves in Fall (SCHUH)
	220L	Plants on a Farm (DICKMANN)
	210L	Counting in the City (STEFFORA)
	210L	The Tractor Race (SCHUH)

*** GN DENOTES GRAPHIC NOVEL, IG DENOTES ILLUSTRATED GUIDE**

Please note:

The Lexile measure (text complexity) of a book is an excellent starting point for a student's book selection. It's important, though, 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 www.Lexile.com.

TEXT LEXILE 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	740–1010L
2–3	420L–820L
1	190L–530L

Common Core State Standards for English Language Arts, Appendix A (Additional Information), NGA and CCSSO, 2012

MetaMetrics®
LINKING ASSESSMENT WITH INSTRUCTION