Manzanita Charter Middle School
English Language Learner
Reclassified Fluent English Proficiency Criteria

Reclassification

The local process used by LEAs to determine whether a student has acquired sufficient ELP to perform successfully in academic subjects without EL support. California Education Code (EC) Section 313(f) specifies the four criteria that must be used when making reclassification decisions locally.

Reclassification Criteria Guidelines

The reclassification criteria set forth in California Education Code (EC) Section 313 and California Code of Regulations, Title 5 (5 CCR), Section 11303 remain unchanged. Pursuant to 5 CCR Section 11308 (c)(6), any local reclassification procedures must be reviewed by the school district advisory committee on programs and services for English learners (ELs). LEAs should continue using the following four criteria to establish reclassification policies and procedures:

Criterion 1: Assessment of ELP

Use the Summative ELPAC results as the primary assessment of ELP. LEAs shall use overall PL 4 as the determination that a student has met the ELP assessment criterion.

Criterion 2: Teacher Evaluation

Use the student's academic performance as evidence of curriculum mastery. (Note that incurred deficits in motivation and academic success unrelated to ELP do not preclude a student from reclassification.)

Criterion 3: *Parent Opinion and Consultation*

Provide notice to parents/guardians of their right to consult with the LEA regarding their child's ELP status and encourage them to participate in the reclassification process. Offer opportunities for in-person meetings with parents/guardians, as needed.

Criterion 4: *Comparison of Performance in Basic Skills*

- Identify local or state assessments that the LEA will use to determine whether ELs are meeting academic measures that indicate they are ready for reclassification. EC Section 313(f)(4) calls for a comparison of student performance in basic skills against an empirically established range of performance in basic skills based on the performance of English proficient students of the same age. The following definitions of related terms may be helpful:
  - performance in basic skills. The score and/or performance level resulting from a recent administration of an objective assessment of basic skills in English (e.g., Smarter Balanced assessments, district benchmarks).
  - range of performance in basic skills. A range of scores on the assessment of basic skills in English that corresponds to a performance level or a range within a performance level.
  - students of the same age. English-proficient students who are enrolled in the same grade as the student who is being considered for reclassification.

For 2018–19, LEAs should identify cut scores, or a range of scores, on the selected assessment instrument to determine the skill levels comparable to English-proficient students. Keep the following in mind: Students with scores at or above the cut point selected by LEAs should be considered for reclassification.

- For students scoring below the cut point, LEAs should attempt to determine whether factors other than ELP are responsible for low performance on the test of basic skills and whether it is reasonable to reclassify the student.
- LEAs must monitor student performance for four years after reclassification, in accordance with existing California regulations and Title III of the Elementary and Secondary Education Act (ESEA).

[CDE: ELPAC Information Guide, 2019.]

Manzanita Charter Middle School Protocol for Redesignation English Language Proficient (RFEP)

Criterion 1: Assessment of ELP
Overall Performance Level of 4 on Summative ELPAC.

Criterion 2: Teacher Evaluation
ELA Grade of C- or Higher OR Teacher written recommendation based on evidence of basic grade level proficiency.

Criterion 3: Parent Opinion and Consultation
Parent Notification Letter

Criterion 4: Comparison of Performance in Basic Skills
Score Basic or Higher on Reading Inventory OR Galileo Assessment OR Score Minimum Level 3 Scale Score or Higher on ELA Smarter Balanced Assessment.

RFEP Cycle

Grade	BOY/Winter 2019-2020 Reclassification	Midyear/Spring 2019-2020 Reclassification	EOY/Spring 2019-2020 Reclassification
6	Basic/730 Lexile	Basic/750 Lexile	Basic/770 Lexile
7	Basic/770 Lexile	Basic/875 Lexile	Basic/790 Lexile
8	Basic/790 Lexile	Basic/950 Lexile	Basic/1000 Lexile

(Dates Dependent Upon School Assessment Schedule.)

Manzanita Charter Middle School Protocol for Reclassification English Language Proficient (RFEP)

Criterion	ED Code Criteria Description	Manzanita Criteria
1	Overall PL of 4 on Summative ELPAC	Overall PL of 4 on Summative ELPAC
2	Teacher Evaluation	ELA Grade of C- or Higher OR Teacher Appeal to Override Grade
3	Parent Recommendation and Consultation	Parent Notification Letter (from TOMS)
4	Comparison of Scores of Basic Performance	Score Basic or Higher on Reading Inventory OR Galileo Assessment OR Score Minimum Level 3 Scale Score or Higher on ELA Smarter Balanced Assessment

NOTE: Upon the 2020-2021 CDE release of the OPTEL Tool, Manzanita shall also offer a passing "score" as an option to meet the criteria of both Criteria 2 & 3.

Committee Approved on:
Parent Committee Members:

____

____

____

____

____

***5 CCR sections 11303 (Reclassification) and 11308 [c][6] (Advisory Committee)*

The Lexile Framework and GALILEO

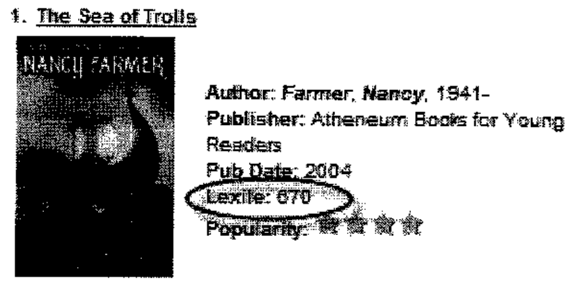
GALILEO Can Help You Find Books and Articles by Their Lexile Score

The Lexile Framework is a tool that links readers and text with a metric called Lexile. A Lexile is a standard score that matches a student's reading ability with difficulty of reading material.

Look for Lexile scores in these GALILEO databases –

NoveList and NoveList K-8

Fiction titles listed in NoveList databases include their assigned Lexile score. You can find the Lexile for a specific book or you can search for books within a specified Lexile range that match the subject, author, series, or other terms of interest to the reader. NoveList folder function allows you to create and save book lists based on various criteria, including Lexile ranges and maturity level.



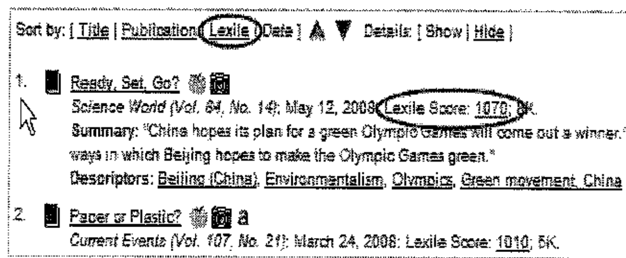
SIRS Databases

- SIRS Discoverer (elementary and middle)
- SIRS Researcher (high school)

In both databases, users can limit a search on any topic to a specific Lexile range using the Advanced Search feature. Once a set of results is retrieved, the results can be sorted by Lexile score. The citation for each article includes the Lexile indicator underlined for easy identification.

- Ppt for Lexiles and ProQuest

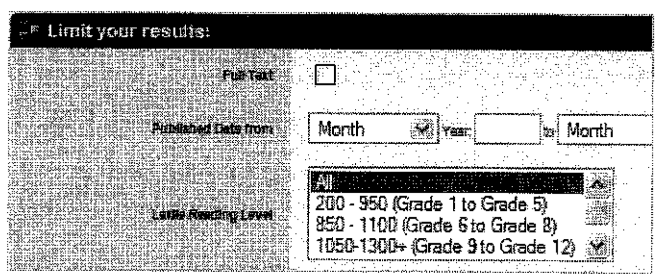
http://www.proquestk12.com/go/LexilesOverview



EBSCO Databases

- MASUltra
- Middle Search Plus
- Primary Search
- TOPICSearch

Portals that access the above databases include the Student Research Center for middle/high, and Kids Search for elementary/middle, and Searchasaurus for primary/ elementary. Students and teachers can search for topics within a specified Lexile range. Each citation in a results list specifies the Lexile indicator, which is hyperlinked to a Lexile rankings map.



From the Lexile.com website:

What is The Lexile Framework for Reading?
The Lexile Framework® for Reading is a scientific approach to reading and text measurement. It includes the Lexile® measure and the Lexile scale. The Lexile measure is a reading ability or text difficulty score followed by an "L" (e.g., "850L"). The Lexile scale is a developmental scale for reading ranging from below 200L for beginning-reader material to above 1700L for advanced text. All Lexile Framework products, tools and services rely on the Lexile measure and scale to match reader and text.

How is a text's Lexile measure determined?
Lexile measures are based on two well-established predictors of how difficult a text is to comprehend: semantic difficulty (word frequency) and syntactic complexity (sentence length). In order to Lexile a book or article, text is split into 125-word slices. Each slice is compared to the nearly 600-million word Lexile corpus – taken from a variety of sources and genres – and words in each sentence are counted. These calculations are put into the Lexile equation. Then, each slice's resulting Lexile measure is applied to the Rasch psychometric model to determine the Lexile measure for the entire text.

For example, books like "Arthur and the Recess Rookie" (370L), "Arthur Goes to Camp" (380L) and "Arthur, Clean Your Room!" (370L) fall within the Lexile Range of a typical second grader. These books have shorter sentences and words appear frequently. Conversely, books in the "Harry Potter" series (which measure between 880L and 950L), "Little Women" (1300L) and "Don Quixote" (1410L) contain longer sentences and more complex words.

From the GaDOE website:

http://public.doe.k12.ga.us/lexile.aspx

The Lexile Framework is an educational tool that links text and readers under a common metric known as the Lexile.The Georgia Department of Education has worked with MetaMetrics, the developers of the Lexile Framework, to customize a "map" that provides a graphic representation of texts and titles matched to appropriate levels of reading ability.

To see the map, visit:
http://public.doe.k12.ga.us/DMGetDocument.aspx/GA-LexileMap-FINAL_2-27-06.pdf?p=39EF345AE192D900F620BFDE9C014CE65F48E7E4CC653240EBCFB76B1C2F47368CDB53B3758CEB56&Type=D

What Is a Quantile Measure?

The Quantile® Framework for Mathematics helps you personalize math learning for students by linking assessment to instruction.

What Student Quantile Measures Tell You

A growing number of math programs and state assessments report Quantile measures. The student Quantile measure is a number followed by the letter "Q." Quantile measures range from below 0Q to above 1400Q and span the skills and concepts taught in kindergarten through high school. For example, a student's Quantile measure should be at 1350Q by high school graduation to handle the math needed in college and most careers.

A student Quantile measure helps you to know:

- Which skills and concepts students are ready to learn
- The level of success students are expected to have with an upcoming skill or concept
- How students are growing in mathematics on a single scale across grade levels

How Quantile Measures Help Student Math Achievement

Quantile measures are more than a math score because they help you identify the math concepts your students know and match them with the concepts they are ready to learn.

Quantile measures are linked to specific math concepts. For example, maybe one of your students needs to review one or two concepts before a a current classroom unit on quadrilaterals. Or, they might be able to complete enrichment activities to move ahead even more quickly.

Quantile measures help match students to their "optimal" challenge. When they work on materials that they're ready to learn, they experience more success and less frustration.

Matching Student Quantile Measures to Skills and Concepts

Mathematical skills and concepts build upon one another. All math learners need to progress through a complex web of skills and concepts that fit together The Quantile Framework for Mathematics has defined almost 500 mathematics skills and/or concepts. Each of these concepts has a measure, and each measure shows how difficult one skill is in relation to the others.

The description of each skill and its Quantile measure is called a Quantile Skill and Concept (QSC) (https://www.quantiles.com/educators/understanding-quantile-measures/skill-and-concept-measures/). As the difficulty or demand of the skill increases, so does the Quantile measure.

The difference between the Quantile measure of the Quantile Skill and Concept (QSC) and a student's Quantile measure gauges how difficult that skill or concept may be for a student to learn.

For optimal learning and growth, a student should practice mathematics within a Quantile range of 50Q above and 50Q below his or her Quantile measure.

Match Students to the Right Materials

Learn how mathematical skills and concepts are connected to Quantile measures.

WHAT TO DO WITH A QUANTILE MEASURE (HTTPS://WWW.QUANTILES.COM/EDUCATORS/UNDERSTANDIN

PARENTS & STUDENTS
(HTTPS://WWW.QUANTILES.COM/PARENTS-STUDENTS/)

EDUCATORS
(HTTPS://WWW.QUANTILES.COM/EDUCATORS/)

Understanding Quantile® Measures
(https://www.quantiles.com/educators/understanding-quantile-measures/)

Math Resources for the Classroom
(https://www.quantiles.com/educators/math-resources-for-the-classroom/)

Fostering Math at Home (https://www.quantiles.com/educators/fostering-math-at-home/)

Measuring Growth with Quantile Measures
(https://www.quantiles.com/educators/measuring-growth-with-quantile-measures/)

EDUCATION COMPANIES
(HTTPS://WWW.QUANTILES.COM/EDUCATION-COMPANIES/)

DEPARTMENTS OF EDUCATION
(HTTPS://WWW.QUANTILES.COM/DEPARTMENTS-OF-EDUCATION/)

MetaMetrics® (http://metametricsinc.com)

Blog (https://metametricsinc.com/about-us/blog/)

News (https://metametricsinc.com/about-us/news/)

Contact Us (https://www.quantiles.com/contact/)

Trademark Statement (https://metametricsinc.com/trademark-statement/)

Privacy Policy (https://metametricsinc.com/privacy-policy/)

(https://www.linkedin.com/company/metametrics-
(https:[illegible]cs)

Use the Growth Planner to Help Prepare for College & Careers

The **Quantile® Growth Planner** (https://www.quantiles.com/parents-students/measuring-growth-with-quantile-measures/forecasting-growth-with-quantile-measures/) offers a snapshot of a student's progress toward college and career readiness. With the Lexile® and Quantile® Growth Planners, you can see the Lexile level and/or Quantile level associated with entry-level reading and/or math demands of hundreds of careers to inform goal setting.

MetaMetrics studied the difficulty of lessons in mathematics textbooks commonly used in the United States to help understand the mathematics demand that students will likely encounter in their elementary through high school mathematics courses. Results are shown in the table below. In a related study, MetaMetrics found that the mathematics ability needed for college and career readiness ranged from approximately 1220Q to 1440Q, and the median mathematics demand for college and career readiness was 1350Q.

Quantile Lesson Measures to Guide Mathematics Instruction for College and Career Readiness

Grade	Lessons Complexity Measures Beginning of Year	Lessons Complexity Measures End of Year
1	EM50Q*	80Q
2	40Q	300Q
3	240Q	490Q
4	390Q	680Q
5	560Q	810Q
6	680Q	890Q
7	800Q	950Q
8	840Q	1050Q
9	900Q	1150Q
10	1070Q	1230Q
11	1100Q	1350Q

*When a Quantile measure is below 0Q, an EM (Emerging Mathematician) code is reported with the measure.

Read our research briefs describing this work: A Quantitative Task Continuum for K-12 Mathematics (https://metametricsinc.com/research-publications/quantitative-task-continuum-k-12-mathematics/) and The Quantile Framework for Mathematics Quantifies the Mathematics Ability Needed for College and Career Readiness (https://metametricsinc.com/research-publications/quantile-framework-mathematics-quantifies-mathematics-ability-needed-college-career-readiness/).

Looking for More Research?

MetaMetrics has gathered years of research as well as conducted its own research on better ways to measure student math ability and report growth.

(http://metametricsinc.com)

MetaMetrics®

Trademark Statement (https://metametricsinc.com/trademark-statement/)

Privacy Policy (https://metametricsinc.com/privacy-policy/)

(https://www.linkedin.com/company/metametrics-

(https:[illegible]cs)

Skill and Concept Measures

Mathematical skills and concepts build upon one another. A child's readiness to learn a specific skill or concept depends on having learned other skills or concepts first.

The Quantile® Framework for Mathematics has defined almost 500 mathematics skills and/or concepts. Each of these concepts has a measure, and each measure shows how difficult one skill is in relation to the others.

The description of a skill and its Quantile® measure is called a Quantile Skill and Concept (QSC). The table below shows a few of these skills and their measures. As difficulty, or demand, of the skill increases, so does the Quantile measure.

The QSCs connect to each other forming an enormous web of mathematics skills and concepts related through their content and their measures. This web of content spans content from kindergarten through secondary school mathematics.

Examples of QSCs

Description of Skill or Concept	Quantile Measure
Solve quadratic inequalities graphically or algebraically.	1250Q
Use properties of circles to solve problems involving arcs formed by central angles or inscribed angles.	1140Q
Solve linear inequalities using the properties of inequality.	980Q
Divide two fractions or a fraction and a whole number.	870Q

Solve problems involving elapsed time.	450Q
Identify and name: hexagon, trapezoid, parallelogram, and rhombus.	250Q

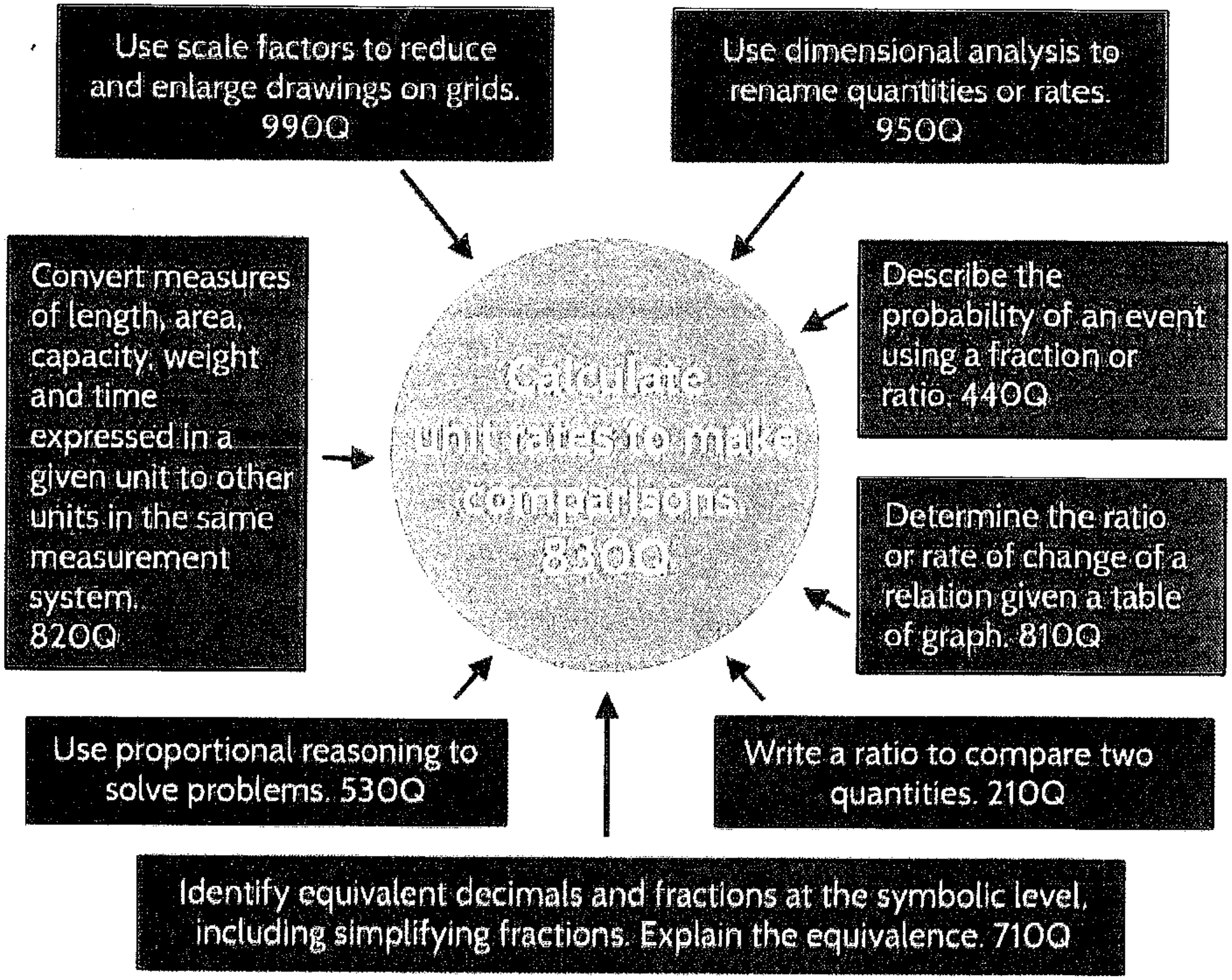
Quantile Knowledge Clusters

Learning math isn't a linear process. Some concepts can't be learned until two or three others are mastered. And, when certain concepts are mastered, there are many directions you can go next.

That's how we've structured QSCs. Each QSC concepts relates to other QSCs that are prerequisite concepts that students need to understand to progress in their study of mathematics.

These relationships form a Knowledge Cluster. Knowledge Clusters show the connections between mathematical skills and give their relative difficulty to one another using the Quantile scale. For example, below is the Knowledge Cluster for the math skill unit rate. This QSC has a Quantile measure of 830Q, and there are many QSCs that relate to it.

- **Prerequisites** are shown in green. They generally have lower Quantile measures than the skill the student is trying to learn.
- **Supplemental skills**, in orange, are related to learning about unit rate but are not dependent on learning or having learned about it.
- **Impending skills**, in purple, are those which depend on understanding unit rate, and can be learned after the student has mastered unit rates.



The QSCs connect to each other forming an enormous web of mathematics skills and concepts related through their content and their measures. This web of content spans content from kindergarten through secondary school mathematics.

The Quantile Map

The Quantile map (https://www.quantiles.com/educators/understanding-quantile-measures/fact-sheets-faqs/) is a visual representation of how QSCs, Knowledge Clusters, and student Quantile measures work together. It provides examples of how The Quantile Framework for Mathematics works at various difficulty levels on the Quantile scale from basic mathematics concepts at EM400Q to more advanced mathematics skills at 1600Q.

Learn More

Discover more math resources to use in the classroom.

MATH RESOURCES (HTTPS://WWW.QUANTILES.COM/EDUCATORS/MATH-RESOURCES-FOR-THE-CLASSROC

PARENTS & STUDENTS (HTTPS://WWW.QUANTILES.COM/PARENTS-STUDENTS/)

EDUCATORS (HTTPS://WWW.QUANTILES.COM/EDUCATORS/)

Understanding Quantile® Measures (https://www.quantiles.com/educators/understanding-quantile-measures/)

Math Resources for the Classroom (https://www.quantiles.com/educators/math-resources-for-the-classroom/)

Fostering Math at Home (https://www.quantiles.com/educators/fostering-math-at-home/)

Measuring Growth with Quantile Measures (https://www.quantiles.com/educators/measuring-growth-with-quantile-measures/)

EDUCATION COMPANIES (HTTPS://WWW.QUANTILES.COM/EDUCATION-COMPANIES/)

DEPARTMENTS OF EDUCATION
(HTTPS://WWW.QUANTILES.COM/DEPARTMENTS-OF-EDUCATION/)

(http://metametricsinc.com)

MetaMetrics®

Blog (https://metametricsinc.com/about-us/blog/)

News (https://metametricsinc.com/about-us/news/)

Contact Us (https://www.quantiles.com/contact/)

Trademark Statement (https://metametricsinc.com/trademark-statement/)

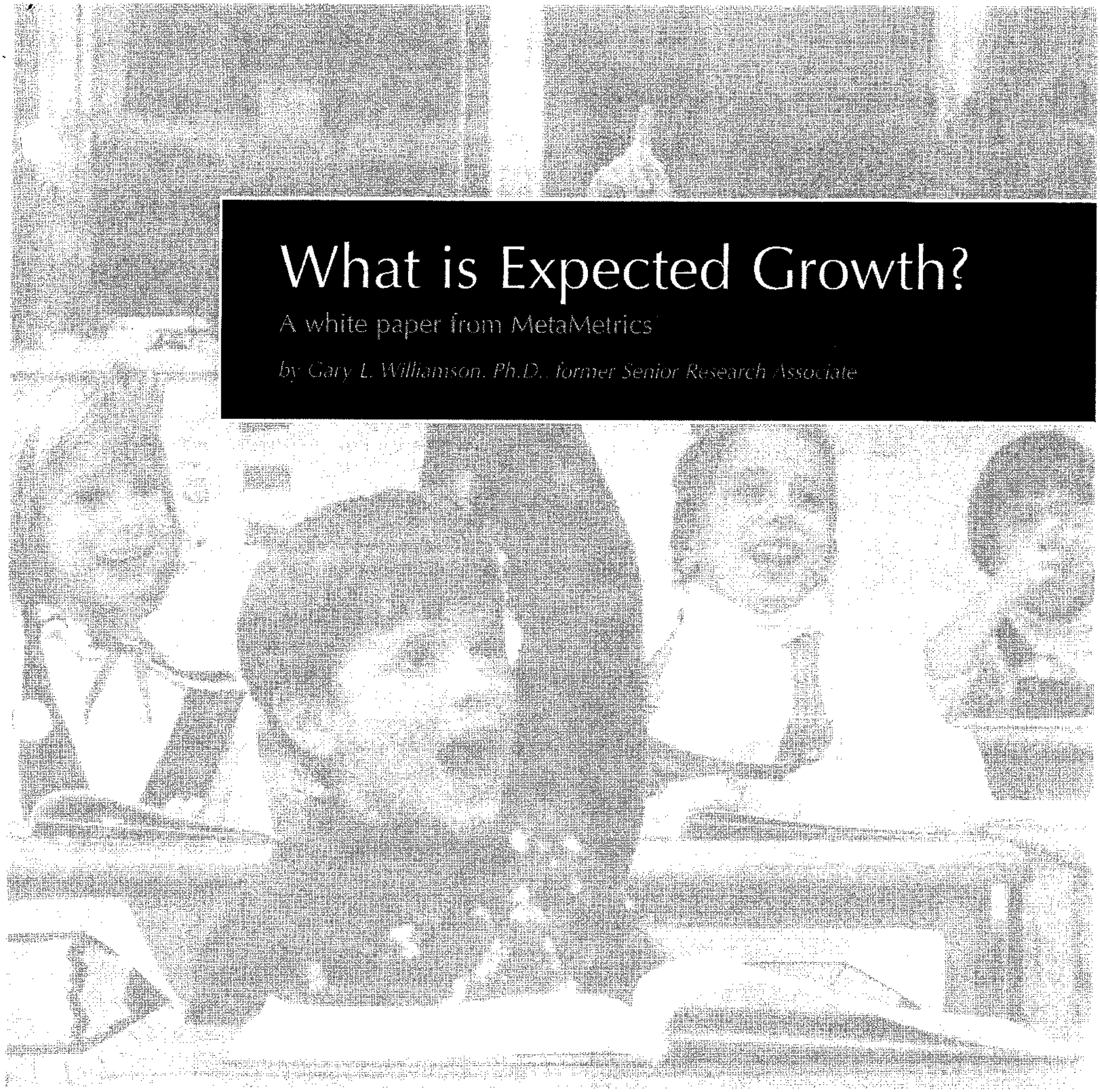
Privacy Policy (https://metametricsinc.com/privacy-policy/)

(https://www.linkedin.com/company/metametrics-
(https://[illegible])cs)

What is Expected Growth?

A white paper from MetaMetrics®

by Gary L. Williamson, Ph.D., former Senior Research Associate



LEXILE® QUANTILE®

Table of Contents

Overview

We are all familiar with children, either through knowing our own or through acquaintance with those of other people. Perhaps no other thing in life is as obvious as the dramatic way that human beings develop and grow. Our key social and political institutions devote a significant part of their resources to ensuring that children grow and learn to function as productive citizens. Growth and learning are central to the mission of our country's public schools.

In January 2002 the President signed into law a major reauthorization of the Elementary and Secondary Education Act (ESEA) that has become known as the No Child Left Behind (NCLB) Act of 2001. The law established sweeping new requirements for educational measurement and accountability for all schools. Not surprisingly, the focus is on the academic achievement and progress of students. These terms (achievement, progress) and related ones (learning, growth, development, performance, proficiency, etc.) occur over 1,660 times in the text of the 670-page law. Setting goals for student performance and monitoring the progress that students make toward those goals are at the heart of the new federal accountability requirements.

NCLB prescribed one way of setting goals and monitoring student progress. States have worked diligently since its enactment to comply with the law and to integrate their efforts within already existing accountability frameworks. In 2005, the U.S. Secretary of Education, Margaret Spellings, created an opportunity for some flexibility when she invited states to propose growth models as part of their strategy to address the requirements of NCLB.

Because there are a number of alternative ways to conceptualize student growth and to measure it, states face a challenge to design and implement accountability systems that address a variety of information needs and still comply with state and federal laws. In this context, there are naturally many viewpoints about how best to conceptualize and measure student growth and to set appropriate goals for growth. This makes it especially important for students, parents and educators to better understand student growth, how it is measured, and how growth expectations may be set in different contexts for different purposes.

What is growth?

In the simplest terms, growth is change over time. To study growth, we measure a thing repeatedly on successive occasions and draw conclusions about how it has changed. People may speak of growth in the context of a system (e.g., a population) or in terms of an organism (i.e., an individual). In the former, we may be concerned with how many individuals comprise the population, how they are dispersed and how rapidly their number increases. In the latter instance, we are generally concerned with how attributes of the organism (e.g., height, weight, reading ability) change over time. Although both notions of growth are interesting, in this paper we are mainly concerned with the second idea because it most closely relates to the concern we have for how individual students develop physically and cognitively.

Most people are familiar with physical growth and some of the ways in which it is measured. For example, one of the things doctors do with new babies is to weigh them and measure their length. Height and weight measurements are continued as the child matures. On any given occasion, specific measures of height (length when very young) and weight are obtained in terms of inches and pounds. Each year (or more often when very young) the measurements can be repeated and a history of development can be gathered for the individual. The change in these measurements over time tells us about the growth in height and weight of the individual, which in turn gives us clues about the child's general health and well-being. Similarly, when children become students in our public schools, their academic performance is measured, for example, in reading. On any given occasion a specific measure of their reading ability is obtained in some metric. Each year (or perhaps more often in some situations) the measurements can be repeated and a history of the student's reading achievement is possible. The change in these measurements over time tells us about the growth of the student's reading ability, which in turn gives us clues about the cognitive health and well-being of the child.

In the preceding paragraph it sounds as though the measurement of physical growth and the measurement of cognitive growth are very similar. In some respects they are, but there is actually a huge difference in practice. You may have noticed that in the preceding comments about height and weight, the measurements were in terms of inches or pounds. In contrast (and this is significant!) the measurement of reading ability was in "some metric." The difference is that whenever we measure height and weight we always use inches and pounds. (In Europe it would perhaps be centimeters and kilograms, but this is not a fundamental difference because there is a direct universal relationship between inches and centimeters and between pounds and kilograms.) In sharp contrast, for the majority of the last century there was no universally accepted metric for the measurement of reading achievement. For the most part, each reading test had its own proprietary metric and, unlike Fahrenheit and Celsius, the reading metrics were not "exchangeable," "convertible," or "translatable" from one to another.

Near the end of the twentieth century, MetaMetrics developed a common metric for reading called The Lexile® Framework for Reading, which is now the most widely used reading scale. However, other metrics still abound. This has huge implications for our understanding of academic growth, as we discuss next.

How is growth measured?

A central question to be addressed when discussing growth is "growth in what?" What are we measuring on each occasion? What is changing over time? Underlying these questions is the assumption that it is the same thing on each occasion even though its magnitude might differ over occasions. (Indeed, we expect its magnitude to change. That is why we measure it more than once.)

For example, when we measure height or weight we fully expect measurements to increase from birth to adulthood. It is this change that interests us. But even though their magnitudes increase over time, it is always height and weight that we measure on each occasion. We do not measure height and weight on one occasion and arm length and girth on the next occasion. This seems trivially obvious when we measure physical attributes, but it is not so obvious when measuring cognitive attributes, like reading ability.

Measuring reading ability is more like measuring temperature. Although we can see a person's height or weight, we cannot directly observe the temperature of an object. We can see evidence of temperature by observing the height of a column of mercury in a thermometer. Similarly, we cannot see a person's reading ability. However, we can see evidence of a person's reading ability by asking them to respond to questions about textual matter they have read. For both temperature and reading ability, we construct an instrument that gives evidence of the unseen attribute. Unseen cognitive attributes are called constructs because we infer their existence from the behavior or performance of individuals. When performance changes, we understand this reflects a change in the underlying construct. Hence, we assume that changes in these unseen constructs are the primary causes of variation in the measurements we observe.

There is a challenge to measuring constructs that is not present when measuring physical attributes such as height and weight. How can we know that the construct that we measure on the second occasion is the same one that we measured on the first occasion? For example, if we ask the same questions on subsequent occasions that we asked the first time we measured the person's reading ability, they could have remembered the answers to some of the questions. The next time we ask the same questions, the student might be able to answer them without even reading the text. In that case we would certainly not be measuring the student's reading ability!

In the example above, the construct changed. On the first occasion we may have actually measured reading ability. But the next time we may have obtained a measure of reading ability contaminated by memory of prior questions and answers. That being the case, we cannot examine the change in the two measures and conclude that the reading ability has changed. We did not measure only reading ability on both occasions!

This points out a key requirement for measuring growth. If we are to measure growth in cognitive constructs there must be a fundamental constancy or invariance in the construct over time. Its magnitude may change but its nature must remain the same. We have to measure the same thing on each occasion in order to even talk about growth.

Psychometrics is a branch of psychology dealing with the design, administration and interpretation of quantitative tests for the measurement of psychological constructs such as intelligence, aptitude (e.g., reading ability) and various personality traits. Making sure that tests really measure what they are intended to measure is one of the fundamental jobs of psychometricians. When they do this, they are ensuring the construct validity of the test. But psychologists who develop measures of cognitive growth must go even further. They must assure that tests are constructed and administered in ways that result in the same construct being measured each time the test is administered. There must be invariance of construct in studies of growth.

There is another fundamental requirement for measuring growth. We must use an appropriate equal-interval scale consistently over time.

A scale is called equal-interval whenever a unit distance at one place on the scale indicates the same amount of change in the underlying construct as a unit distance at another place on the scale. For example, a two-inch increase in height means the same thing regardless of whether the increase was from 32 to 34 inches or from 70 to 72 inches. As long as we record the numbers in terms of inches each time, we have used the same scale (inches) consistently over time.

These fundamental requirements must also apply when we measure psychological constructs. Psychometricians must develop scales that behave in an equal-interval fashion. Furthermore, when we study growth we must use the same equal-interval scale consistently over time. One famous psychometrician coined a now well-known phrase to capture this notion: "If you want to measure change, don't change the measure."

When we design studies of growth, it is important to use a valid equal-interval scale. Furthermore we must be able to persuasively demonstrate that over time there is invariance of construct and consistency of scale. These are the fundamental underpinnings for measuring growth. If these conditions are not met, the study may be interesting but it is not about growth.

The great advantage gained by employing stable constructs and consistent equal-interval scales is that we can perform mathematical operations (addition, subtraction, etc.) in sensible ways with the scale values that are recorded on each occasion of measurement. We can add the amounts of growth in consecutive time periods to establish the amount of growth over the whole time-span, for example. More importantly, we can mathematically model the growth over time and look at its functional form mathematically. For example, does the individual grow in a steady fashion with a constant rate of growth? Or does the individual grow faster when young and more slowly as he or she grows older? Do different individuals exhibit different patterns of growth? What is the most typical pattern of growth? How much variation should we expect to see across individuals? Once construct invariance and scale consistency have been demonstrated, it becomes possible to address questions like these.

What is "normal" growth?

When we ask, "What is normal?" whether it pertains to performance, height, reading ability, or growth in these attributes, we generally assume that we can make a judgment about what occurs most frequently in the general population of individuals. Usually this is accomplished by gathering information about the general population so that we have a frame of reference (data) against which to make comparisons. Such reference data are called norms.

In theory there are two types of norms for growth—cross-sectional norms and panel norms. In cross-sectional norms for growth, a sample of people representing the ages of interest are studied at a single point in time; or perhaps comparable samples of people are studied on multiple occasions, but not the same individuals each time. For panel norms, the same individuals are followed and studied on multiple occasions (as many as necessary to reflect the ages of interest.) In practice, cross-sectional norms are more common because panel norms are expensive and time-consuming to construct. Cross-sectional norms are useful for seeing how an individual compares to the general population at any given point in time. Panel norms are preferable if we want to examine the rate of growth of the individual in relation to that of the population.

Probably the most familiar cross-sectional norms for growth are the Centers for Disease Control and Prevention (CDC) Growth Charts: United States published by the National Center for Health Statistics (NCHS), one of several centers under the umbrella of the Centers for Disease Control and Prevention (CDC) of the U.S. Department of Health and Human Services. The CDC Growth Charts are used by doctors everywhere in the United States as the frame of reference for evaluating the physical development of children.

The CDC Growth Charts are based on surveys of representative samples of people of different ages at specific points in time (but not the same people each time). The NCHS examined the distribution of height (also weight and selected other physical characteristics) across all individuals of a given age in their samples. In essence, they plotted selected percentiles of the distributions for every age from 2 to 20 years and created cross-sectional "growth curves" by connecting the corresponding percentiles from the distribution at each successive age. (It was considerably more complicated than that in reality. Sophisticated curve fitting and smoothing techniques were used to assure that the curves best described the data.)

For education, test companies construct cross-sectional norms by grade, rather than age, to be more applicable to the way public schools are organized. Test companies periodically test a nationally representative sample of students in each grade and construct norms tables to show how the academic performance of students is distributed in each grade. However, these norms are usually limited to a specific point in time and to a specific edition of a test. As a result they are not really growth norms, but achievement norms. Most test companies provide such norms for reading and mathematics, and often for other subjects as well.

The CDC Growth Charts show how the sizes (heights and weights) of individuals in the population vary at different ages. However, this is different from showing how the size of any specific individual changes as he or she ages over time. To do that you must follow the same individual over time and make measurements on the same individual at each successive occasion. If you do that with a representative population of individuals, then you have the basis for panel norms.

The CDC makes this distinction in their report 2000 CDC Growth Charts for the United States: Methods and Development. In the report, they use terms like "growth (or size)" and "size-attained" and "growth progress" to characterize the information about physical growth obtained from their cross-sectional norms. In contrast, they use the term "growth velocity" to distinguish the kind of information about growth that could be provided by panel norms. They say:

> There is a difference between growth (or size) charts and growth velocity charts. The 2000 CDC Growth Charts for the United States are based primarily on cross-sectional national survey data that were statistically smoothed to create percentile curves. ... Therefore, these charts more appropriately may be considered size charts. When serial values for an individual are plotted, assessments can be made of that individual's growth progress over time.
>
> Growth velocity charts are constructed from incremental data obtained from longitudinal observations. (p. 14)

In education, the publishers of achievement tests do not publish growth curves comparable to those displayed in the CDC Growth Charts. However, cross-sectional norms tables are published independently for each grade and academic subject. In the past, some educators have used such tables for successive grades to see if students maintain the same percentile rank in the norms. Maintaining the same percentile rank was considered "normal growth." (Earlier evaluations of federally funded Title I programs were based on such a model.)

This approach provided for academic achievement an analog to the "growth size" information available from the CDC Growth Charts for height or weight. But as discussed above, this information is cross-sectional in nature. Such an approach does not produce "growth velocity" information of the kind that could be obtained from panel norms.

Panel norms for growth would have significant advantages over cross-sectional norms for growth. Foremost, they focus on the same individual over time, so they provide a more true-to-life perspective on intra-individual change. Second, they can better detect growth. NCHS points out that "growth velocity charts are more sensitive indicators of small changes in growth status than the size-attained charts ...." Third, they are more accurate. Williamson (1988) demonstrated that a cross-sectional approach to constructing norms for a growth distribution could produce distortion, particularly at the extremes of the growth distributions.

Given these advantages, it may be surprising that panel norms for growth generally have not been constructed or widely used in practice. One can surmise several reasons. Among the possibilities, consider the following:

- It would take a long time to gather the data necessary for panel norms. Imagine that we desire to have norms for growth during the public school years. To be able to have norms for grades K–12, it would take at least 13 years to follow a group of students from Kindergarten to graduation. (That ignores the fact that some students will not actually finish in 13 years due to having to repeat one or more grades.)
- Over such long periods of time, people can move away and become unavailable for follow-up measurements.
- The attrition of students in turn compromises the validity of the norms because highly mobile students would not be included.
- There would have to be adequate measurement over the time period (invariance of construct, consistent measurement with an equal-interval scale).
- Curricular changes across grades may make it difficult or irrelevant to focus on the same construct throughout the entire timeframe.
- The cost of collecting the data for panel norms for growth would be much larger than for cross-sectional norms at a single point in time.
- Before such a study could be completed, most test companies would have published new editions of their tests. Norms for the older editions would be of less interest and not highly marketable.

Still, with the wide-spread proliferation of accountability programs in the United States, data bases are being expanded and improved. With a little work, some states may be in a good position to retrospectively create panel norms with data that already exist. Others have the opportunity to incorporate the right design features in systems that they are creating now to operate over the next 10-15 years. There would be several motivations for exploring this possibility.

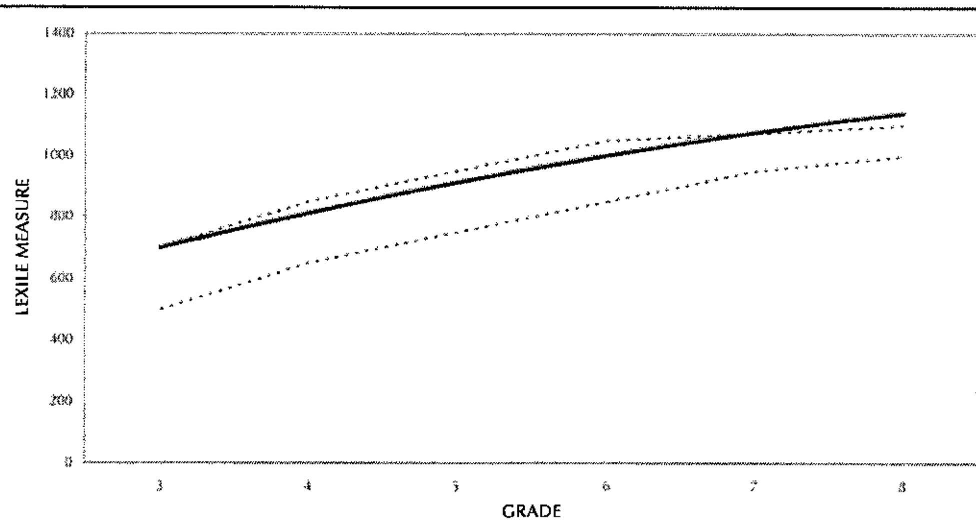
- A developmental approach to growth is possible with panel data. Analysts can mathematically model individual growth curves for students using such data.
- A focus on developmental growth curves allows one to study not only achievement, but also change in achievement over time (its rate and/or magnitude), and the functional form of growth (acceleration, deceleration, etc.).
- The relation between growth and initial achievement is more easily investigated. For example, do initially lower-performing students grow faster over time than those who start out with better performance?
- A wide variety of new analytical techniques have arisen in recent years to take better advantage of panel data.
- A longitudinal focus on student growth would be consistent with the current statutory emphases on student progress.
- New measurement scales (e.g., the Lexile scale) provide a more nearly universal measurement framework for such long-term studies.

Cross-sectional norms and panel norms provide performance-based frameworks for determining what normal growth is. There is another useful perspective based on an analysis of curricular demands that are applicable to students in school. A good example is the text demand analysis enabled by The Lexile Framework for Reading (MetaMetrics, Inc., 2006). This approach looks at typical texts used in each grade, and places them on the Lexile scale. By depicting the range of Lexile measures corresponding to texts typically used in each grade one can see what range of reading ability students must exhibit to be able to comprehend their texts. This provides a curricular-based rather than a performance-based perspective that might be used for defining "normal" growth.

An Example

The figure below provides an illustration of how more than one normative perspective may be combined. In the figure, the average growth curve for a longitudinal panel of readers is displayed in conjunction with the changing text demand for each successive grade. The growth curve represents the average growth of about 68,000 students who were third graders in the Public Schools of North Carolina in 1999, and who were followed until the end of eighth grade in 2004. Each year the students took the North Carolina End-of-Grade Test in Reading. Their scores were converted into Lexile measures to facilitate the comparison with grade-level text demand. Longitudinal data analyses were used to estimate each individual student's growth curve. The average growth is displayed in the Figure by the solid curved line, and the range of reading difficulty for typical textbooks (approximately the middle 50%) at each grade is portrayed between the two dotted lines. The chart shows that on average these North Carolina students were reading near the upper difficulty levels of typical grade level texts.

1999-2004 North Carolina Average Growth Curve Relative to Grade Level Text Demand



Expected growth

In the previous sections we discussed normal growth and described three ways to characterize it (through cross-sectional norms, panel norms, or text demand). An example was given illustrating how to combine two normative perspectives (a panel norm for average growth and changing grade level text demand) into a single picture. In this section we raise a slightly different question—how much growth to expect of a given student or group of students.

There are at least two fundamental ways to address this question. We can base the decision on some reference standard regarding what growth has been in the past (norms) or we can base our decision on what we think is desirable for the future (aspirations).

If we base the decision on norms, we will look at past and present performance and ask:

- How much growth do we expect?
- How much growth is typical?

When we use past performance as our guide, then the answer to these questions is "whatever is normal" for the students given their age, grade, background, etc. For instance, we could use the average growth curve in the example from the previous section to calculate the year-to-year growth that occurred in North Carolina. This is one expression of what may be "normal" given the performance of the students in the panel.

Average Year-to-Year Growth Exhibited by the 1999-2004 Panel of North Carolina Students

Grades	Average Growth
3 to 4	113L
4 to 5	100L
5 to 6	88L
6 to 7	76L
7 to 8	64L

On the other hand, if we base our decision about what growth is expected on our aspirations for the future, we are asking questions such as:

- How much growth should the student(s) exhibit?
- How much growth is desirable?

To answer these questions, we have to look to sources other than (or in addition to) past performance. Subjective judgments must be made about how much growth is needed for some identified purpose, or consistent with some agreed upon values.

Even though norms describe what is or has been, they can still be a useful reference for what should be, or to what heights students should aspire. For example, by referring to achievement norms, educators can assess what the typical performance has been in the past and so set a baseline for future performance comparisons. Exceeding the achievement norms of the past has often been the goal of educators.

Another normative method of setting an expectation for growth is to use the curriculum as a reference and specify what students should know and be able to do at each grade. Such curricular standards are often further operationalized as proficiency standards for tests that have been developed to align to the curriculum. When a universal scale is available, such as the Lexile Framework, the texts used at each grade can be mapped on the scale to create a graph of the text demands over time. This is effectively a textual norm against which student performance and progress may be plotted, as illustrated in the example above.

Although "whatever is normal" is one answer to how much growth we should expect, it is not the only answer. It is generally not the answer in most accountability systems today, which are motivated by a desire to see students perform at higher levels than has been the case in the past. In this context, more than "normal" growth is generally desired. But how much?

One way to address this issue is to focus on a performance that may be required of students in the future. For example, using the Lexile Framework to study postsecondary text demands, Williamson

(2006) showed that a large gap exists between the difficulty level of texts used in high school (eleventh and twelfth grades) compared to texts used in the first two years of university work. Furthermore, there is a systematically increasing continuum of text difficulty that spans the most typical postsecondary activities of students—citizenship, the military, the workplace, community college and the university. Given the knowledge that students may encounter more difficult books in their postsecondary endeavors, we can ask how much growth in reading ability must occur during school to allow students to reach the desired reading ability by the end of twelfth grade. This would form the basis for a more demanding growth expectation in reading.

An example of another strategy for setting growth standards is the approach used by the Public Schools of North Carolina from 1996-97 until 2004-05. In North Carolina, as they developed the ABCs accountability program, they began by analyzing the year-to-year growth of the same students over time. Then the North Carolina State Board of Education (NCSBE) set a target of exceeding the statewide growth of previous years by 10% as one level of expectation for future growth. The decision to choose 10% was based on considerations of the potential payoff in terms of student proficiency after several years of implementation and the potential impact on schools under the accountability system of rewards and sanctions.

When setting expectations based on aspirations as North Carolina did in constructing the ABCs, a guiding philosophy is important. This is used to construct a framework for deciding how to balance the desire for challenge and rigor with a competing desire for realism and attainability. The NCSBE wanted to see improvements in proficiency and closing of achievement gaps within a short but realistic timeframe. They used input from educators across the state, analyses of data from previous years and the state's performance on other measures to provide the context for their decision. North Carolina used both norms and aspirations to set their expectations (growth goals) for the ABCs.

Sometimes aspirations are self-imposed, such as when a student decides to compete in the county spelling bee, or a state aspires to be "first in education" by some date. Sometimes aspirations are externally imposed, such as when a state establishes accountability targets for schools, or when the federal government passes laws such as NCLB to ensure that states incorporate new standards into their accountability programs.

However, rarely are expectations based purely on norms or purely on aspirations. Usually it is some combination of the two, or an intellectual interplay that involves the considerations of what is possible and what is reasonable.

Consequences of measurement error

All measurements are subject to errors that result from influences that have nothing to do with the thing being measured. For example, if we're measuring height, but the individual slouches, we'll get an inaccurate measurement of their erect stature. If we give an individual a test of reading ability, but the room where the test is administered is filled with noisy distractions, then it is unlikely that we will get an accurate indication of the person's reading ability. A more complete introduction to measurement error is presented in Williamson (2004).

Just as single measurements are subject to measurement error, repeated measurements that are used to estimate growth are each subject to measurement error. Consequently, estimates of growth produced from measurements on multiple occasions are also affected by measurement error. Knowing how dependably we have measured growth is thus complicated. There are a number of factors to consider.

One factor that affects the precision with which we can estimate growth is the precision with which we measure status on any given occasion. Two additional factors that affect the precision of growth estimates are how many replications of measurement are available and how they are distributed across time.

In general, having more data is better. It is easy to understand that we cannot measure growth if we only measure the individual on one occasion. Many studies of change have been based on two successive measures, and there has been much debate about the adequacy of this approach. Rogosa and Willett (1985,

1983) describe the limitations on what can be learned when only two measures are available. They have helped to clarify why more than two measures are needed when growth is the focus.

In addition, when the measurements are made affects the accuracy with which we can estimate growth. In essence, the more spread out the measures are in time, the more precisely we can estimate the amount (or rate) of growth. In addition, when the measurements are more spread out, we are better able to see the functional form (changes in or shape of) growth over time.

Imagine a student's growth in reading ability over six years, say from the end of second grade to the end of eighth grade. Suppose that growth is fast during the early years, but by grade five it starts to slow down and tapers off by grade eight. If we took only a few measures, what we would see would depend greatly on when we made the measurements. If they were all in the early grades, we would estimate the growth to be very fast; whereas if we measured only in the later years, we would conclude that growth was relatively slow. In order to get the whole story, we have to space the measurements out over the whole time period, with some measurements near the beginning, some in the middle and some near the end. Even when growth is uniform over the period, we can still estimate the growth more precisely when the measurements are spread out than when they are all near the middle of the timeframe.

Another thing that affects the accuracy with which we can estimate growth is the choice of growth model. Consider again the previous example. In order to correctly capture the growth trend, we must adopt a curved trajectory to describe the growth. If we chose to model the growth with a straight line, we would correctly capture the average change over the whole time frame, but we would miss the early acceleration and the slowing down in later years.

Our ability to assess the collective growth among a group of individuals is similarly affected by considerations related to precision of measurement, amount of data, and the data collection design. In addition, if we are interested in inter-individual (between persons) variation in growth, there must be real variability in growth for us to detect it. If everyone is growing in the same manner then there is no inter-individual variation to detect, and measures of change may appear to be unreliable. Similarly in such situations there may appear to be no relation between change and initial status for that group of individuals. However, these same conclusions are sometimes reached because of the way data were collected or analyzed. Rogosa & Willett (1983, 1985) provided an extensive analysis of the implications of various methodological choices that researchers have made in the measurement of change.

Although the notion of measuring growth seems simple, the details of obtaining dependable measurements and valid inferences about growth can be quite complicated. A vast literature about growth in cognitive constructs appeared during the twentieth century with an explosion of methodological advances during its final decades. Taking full advantage of these advances requires carefully constructed databases of longitudinal data, which in turn require patience and care to maintain. Among the payoffs are the capabilities to better understand growth and to set reasonable expectations for growth. These capabilities seem particularly relevant now in the context of the educational accountability requirements of the 21st Century.

References

National Center for Health Statistics. (2002). 2000 CDC growth charts for the United States: Methods and development. Vital and Health Statistics, 11(246). Washington, D.C.: National Center for Health Statistics.

MetaMetrics, Inc. (2006) Figure 1: Typical reader and text measures by grade [On-line]. Available: http://www.lexile.com/DesktopDefault.aspx?view=ed&tabindex=6&tabid=18#18

Public Law 107-110. The No Child Left Behind Act of 2001.

Rogosa, D. R. & Willett, J. B. (1985). Understanding correlates of change by modeling individual differences in growth. Psychometrika, 50(2), 203-228.

Rogosa, D. R. & Willett, J. B. (1983). Demonstrating the reliability of the difference score in the measurement of change. Journal of Educational Measurement, 20(4), 335-343.

Williamson, G. L. (2006, April). Student readiness for postsecondary endeavors. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Williamson, G. L. (2004). Why do scores change? Durham, NC: MetaMetrics Inc.

Williamson, G. L., Appelbaum, M. I., & Epanchin, A. (1988). Analyses of achievement data from Durham County (NC) schools. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED294923.)

About the Author

Gary L. Williamson, Ph.D., is a former senior research associate at MetaMetrics. With more than 30 years of experience in educational research on the academic, state and school district levels, Williamson's specialty is quantitative methodology encompassing psychometric, mathematical and statistical applications to educational data. He has written and spoken extensively on the subjects of educational assessment and accountability. Williamson earned both a doctorate of philosophy in mathematical methods for educational research and a master's of science in statistics from Stanford University. He also holds a master's of education in educational research and evaluation from The University of North Carolina at Greensboro, and a bachelor's of science in mathematics from The University of North Carolina at Chapel Hill.

MetaMetrics, an educational measurement and research organization, develops scientific measures of student achievement that link assessment with targeted instruction to improve learning. The organization's psychometric team developed the widely used Lexile Framework for Reading; El Sistema Lexile para Leer, the Spanish-language version of the reading framework; The Quantile® Framework for Mathematics; and The Lexile Framework for Writing.

MetaMetrics

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

Phone: 919-547-3400/1-888-LEXILES
Fax: 919-547-3401
Web site: www.Lexile.com

The Lexile® Framework for Reading

LEXILE Matching readers with targeted text

Making Test Scores Actionable

Today's students take many different types of assessments and receive many different scores. These scores provide important measures of student proficiency in content areas, but they are often static. The data can't be used to inform instruction or to help educators and parents select appropriate reading materials based on each child's ability.

The Lexile® Framework for Reading changes that. When an assessment is linked with the Lexile Framework, students' test scores immediately become actionable. A Lexile measure is the most widely adopted reading metric, measuring both reader ability and text difficulty on the same scale. Lexile measures are powerful tools for linking assessment with instruction across the curriculum, at home and in the library, by taking the guesswork out of selecting materials that can help to improve student reading ability.

Matching Readers With Text

Consider this: a father takes his son to the store to buy shoes. The salesperson asks them, "What kind of shoes do you need?" The father replies, "He needs basketball shoes." As the salesperson leads them to the basketball shoes, he asks, "How old is your son?" The father answers, "He is 12." And so the salesperson points to five pairs of shoes on the wall and says, "There are our age-12 basketball shoes."

Not likely, right? We do not buy shoes by age, we buy them by size. However, for years we have matched students with books and other learning materials based on age or grade level. If a student likes science fiction books and is nine years old or in fourth grade, he is given "fourth-grade science fiction" to read. But what if that fourth grader's reading ability is far higher than the "average" student his age, or he is not quite reading on a fourth-grade level? Like the boy's age-12 basketball shoes, the text simply does not fit the student.

When shoe shopping, the scale used to measure the boy's foot for his sneakers tells you his shoe size, not his age. Similarly, the Lexile Framework determines a student's reading ability, not his grade level. With Lexile measures, assessment results are used to ensure a "good fit" because the measure is used to select reading material that meets and challenges each student's ability.

Differentiating Instruction Across the Curriculum

Lexile measures provide more than a way to pick the right book for a student to read. They are powerful tools for targeting instruction and improving achievement across grade levels and content areas. Using Lexile measures, educators connect students with instructional resources that match their reading abilities. The number of fiction and nonfiction books, textbooks, periodicals and Web sites with Lexile measures grows every day.

Consider an educator who is teaching a unit on the Battle of Gettysburg. Typically, a fifth-grade class of 30 students will have a wide range of reading abilities, with only half of those students reading well enough to comprehend the content in the textbook. When the teacher uses the students' Lexile measures to connect them with ability-appropriate content from periodicals and Web sites, they stay on track for meeting state performance standards in social studies and continue to strengthen their reading skills.

More Information, Not More Testing

Best of all, the Lexile Framework adds value to state or classroom assessments—adding more information, not more testing time. It is not another test or a reading intervention program. Lexile measures provide a thermometer for measuring students' reading abilities that talented educators, involved parents and motivated students use to improve learning. Lexile measures tie day-to-day work in the classroom to critical high-stakes tests. The Lexile Framework offers a "big picture" view of growth of student reading ability from preschool through graduate school.

The Lexile Framework for Reading

Measuring Student Growth on a Common Scale

Many reading comprehension assessments and programs are linked with the Lexile Framework. This offers educators and parents a common scale for monitoring student progress throughout the school year and their entire education, regardless of what assessments they take. A Lexile measure can provide the same continuity for reading growth that families have when they mark a child's height on a wall with a pencil. Like the wall, the scale never changes, so progress is easy to see even if the student changes grades or takes a different assessment. Lexile measures range from below 200L for beginning readers and text to above 1600L for advanced readers and text.

Connecting Students With Lexile Measures

Lexile measures may already be available for your students. Tens of millions of students receive a Lexile measure each year from a reading assessment or program. Visit www.Lexile.com for a list of standardized assessments and programs that are linked with the Lexile Framework. You may already have access to this actionable tool for linking assessment with instruction.

Building Reading Abilities Year Round

Lexile measures also provide parents with a powerful tool for connecting children with reading materials at home. Unlike other test results that may simply get posted on the refrigerator or cause parents sleepless nights, Lexile measures offer a way to take action. During the summer, after school or on the weekends, families can visit the library or bookstore and use Lexile measures to select leisure-reading materials. In fact, an increasing number of automated library card catalogs include Lexile measures for books and other materials. Or families can use "Find a Book" to build book lists based on a child's Lexile measure and interests. The book search and other online resources are available at no cost.

Experts agree that the best way to build reading ability is practice. Lexile measures provide a way to make that practice meaningful all year round. By reading material at their Lexile level, students can strengthen literacy skills and develop a lifelong love of reading.

The Science Behind The Lexile Framework for Reading

The Lexile Framework for Reading was developed after more than 20 years of research by the renowned psychometric team at MetaMetrics®, an educational measurement and research organization. The organization's research was initially funded with grants from the National Institute of Child Health and Human Development, part of the National Institutes of Health. Today, MetaMetrics continues to pioneer scientific measures of student achievement that link assessment with targeted instruction to improve teaching and learning.

MetaMetrics, an educational measurement and research organization, develops scientific measures of student achievement that link assessment with targeted instruction to improve learning. The organization's psychometric team developed the widely used Lexile Framework for Reading; El Sistema Lexile para Leer, the Spanish-language version of the reading framework; The Quantile® Framework for Mathematics; and The Lexile Framework for Writing. In addition to licensing Lexile and Quantile measures to state departments of education, testing and instructional companies, and publishers, MetaMetrics offers professional development, resource measurement and customized consulting services.

For more information on using Lexile measures to match readers with targeted text, visit **www.Lexile.com.**

MetaMetrics.
LINKING ASSESSMENT WITH INSTRUCTION

1000 Park Forty Plaza Drive, Suite 120
Durham, North Carolina 27713
919-547-3400/1-888-LEXILES
www.Lexile.com

What does the Lexile® measure mean?

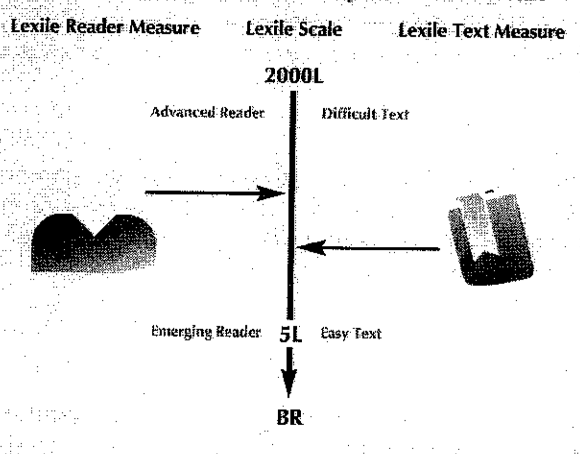
The Lexile® Framework for Reading

The Lexile® Framework for Reading is a scientific approach to reading and text measurement. There are two Lexile® measures: the Lexile reader measure and the Lexile text measure. A Lexile reader measure represents a person's reading ability on the Lexile® scale. A Lexile text measure represents a text's difficulty level on the Lexile scale. When used together, they can help a reader choose a book or other reading material that is at an appropriate difficulty level. The Lexile reader measure can also be used to monitor a reader's growth in reading ability over time. Figure 1 shows the relationship between the Lexile reader measure and the Lexile text measure on the Lexile scale.

A higher Lexile reader measure represents a higher level of reading ability on the Lexile scale. A Lexile reader measure is usually obtained by having the reader take a test of reading comprehension. There are about two dozen tests that can report Lexile reader measures including the Scholastic Reading Inventory™, The Iowa Test of Basic Skills® and many end-of-grade state assessments. A reader's score on the test is reported as a Lexile measure. The reader measures can range from below 0L to above 2000L. When a reader scores below 0L, a BR (Beginning Reader) code is displayed on his or her report. Some test reports may report scores below 0L with a more specific measure. These Lexile measures are shown as BR followed by a number and L (e.g. BR150L). 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. In May 2014, MetaMetrics® rolled out this enhanced way of reporting beginning reader measures. While some of MetaMetrics' partners are offering this new format, others plan to make the change at a later date.

A Lexile text measure, like a Lexile reader measure, is reported on the same Lexile scale, from a low of BR to a high of 2000L. Like the reader measure, all text measures below 0L are currently reported as specific Lexile measures (e.g. BR75L). The smaller the number after the BR is, the more challenging the text is expected to be for a reader. Above 0L, the lower a book's Lexile measure, the easier it will be to comprehend. For example, a text with a Lexile measure of 850L will most likely be easier for a reader to comprehend than a text at 950L. A list of books and their Lexile measures can be found at fab.lexile.com.

FIGURE 1: Lexile reader and text measures placed on the Lexile scale



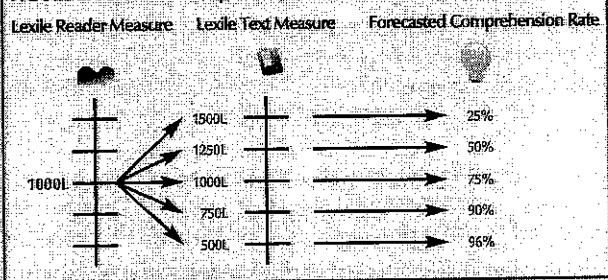
A very useful feature of Lexile reader and text measures is that they can be used together to predict how well a reader will likely comprehend a text at a specific Lexile level. For example, if a reader has a Lexile measure of 1000L, he will be forecasted to comprehend approximately 75 percent of a book with the same Lexile measure (1000L). The 75-percent comprehension rate is called "targeted" reading. This rate is based on independent reading; if the reader receives assistance, the comprehension rate will increase. The target reading rate is the point at which a reader will comprehend enough to understand the text, but will also face some reading challenge. At this challenge point, a reader is not bored by text that is too easy, but also does not experience too much difficulty in understanding. The result is a rewarding reading experience. In some cases, a reader may not want to choose a book at the 75-percent forecasted comprehension rate. For example, if a reader is highly motivated or very interested in a book's topic, the reader may want to choose a book that will be more challenging (less than 75-percent forecasted comprehension).

At other times, the reader may want to choose a book for easy independent reading (90-percent or higher forecasted comprehension). To adjust the forecasted comprehension rate, simply look for a text that has a different Lexile measure than the reader. If the Lexile text measure is higher than the Lexile reader measure, forecasted comprehension goes down. If the Lexile text measure is lower than the Lexile reader measure, forecasted comprehension goes up. For example, if a reader wants to read a book independently at a 90-percent comprehension rate, she can simply choose a book with a Lexile measure approximately 250L below her

What does the Lexile measure mean?

Lexile reader measure. A reader with a measure of 1000L would choose a book with a measure around 750L. Figure 2 shows how changing the Lexile text measure changes the forecasted comprehension rate.

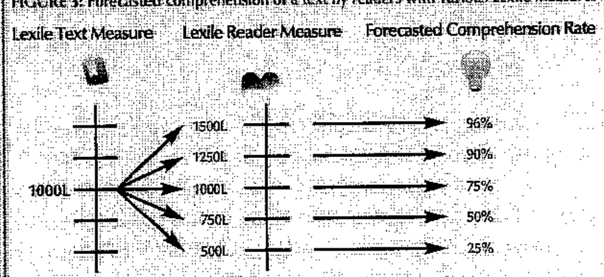
FIGURE 2: Forecasted comprehension of books with various Lexile measures



It is best to select books based on a reader's Lexile range rather than focus on one reader measure. At MetaMetrics, we refer to a "Lexile range" as the suggested range of Lexile measures that a reader should be reading—50L above to 100L below his or her Lexile measure. A reader with a Lexile measure of 1000L would have a Lexile range of 900L–1050L. If a student attempts material above his or her Lexile range, the level of challenge may be too great for the student to be able to construct very much meaning from the text when reading independently. Likewise, material below the reader's Lexile range may provide that student with little comprehension challenge. It should be noted that material above or below the reader's Lexile range may be used for specific instructional purposes.

It is important to note that the Lexile measure of a book refers only to its text difficulty. A Lexile measure does not address the content or quality of the book. A text receives a Lexile measure by running it through the Lexile Analyzer® which utilizes a linguistic algorithm that examines the semantic and syntactic features of the text. Many other factors affect the relationship between a reader and a book, including its content, the age and interests of the reader, and the design of the actual book. The Lexile measure is a good starting point in the book selection process, but these other factors should be considered when making a decision about which book to choose.

FIGURE 3: Forecasted comprehension of a text by readers with various Lexile measures



In some classroom situations, a textbook is used as the main source of reading material for all students. The reading ability of the students in the class may not be matched well to the text difficulty. This can lead to wide variation in forecasted comprehension rate among students reading the same text (see Figure 3). Reading support can be provided to students whose Lexile reader measures are well below the text's Lexile measure. More challenging readings can be provided for students whose Lexile measures are well above the text's Lexile measure.

A helpful feature of the Lexile scale is that it is a developmental scale. This means it can be used to show whether a reader's reading ability is growing (or developing) over time. For example, a student may take a test in third grade that reports a Lexile measure. If the student takes a test in a later grade that also reports a Lexile measure, he or she can see whether his or her reading ability has grown. This feature allows educators and parents to monitor a reader's growth over time. If a student's reading growth is too slow, or even stalled, it may be necessary to provide supplemental reading instruction.

Additional information about Lexile measures can be found at www.Lexile.com.

- To find the Lexile measures of books, or to search for books using a Lexile range, go to fab.lexile.com. Several search options are available, including book title, reader interest categories, book author and developmental level.
- A complete list of the tests and reading programs that report Lexile reader measures is available at www.Lexile.com.

Although Lexile measures should not be linked directly to grade levels, it is possible to describe the Lexile measures of typical students and textbooks at various grade levels. MetaMetrics collected the Lexile measures for a national sample of students and identified the Lexile measures for the students in the middle 50 percent of the Lexile range (25 percent–75 percent). The results of this study can be found at www.lexile.com/about-lexile/grade-equivalent/grade-equivalent-chart.

MetaMetrics.

KEEP STUDENTS ON TRACK FOR COLLEGE & CAREER

GUIDE STUDENTS WITH THE LEXILE GROWTH PLANNER® & QUANTILE GROWTH PLANNER™

With less than 40 percent of U.S. students scoring at college- and career-ready levels, students, educators and parents need resources to better monitor progress towards college- and career- readiness.[1]

The Growth Planner is a tool that uses Lexile® and Quantile® measures from annual statewide tests to determine if a student is on track to graduate college-and career-ready. The Growth Planner helps:

Educators & Guidance Counselors

- Forecast student growth toward college and career readiness.
- Identify gaps.
- Determine required instructional needs to promote growth and close the gaps.

Students

- Manage their preparedness for college and career.
- Determine if their reading and math skills are on track.
- Access resources that promote reading and math growth.

Parents

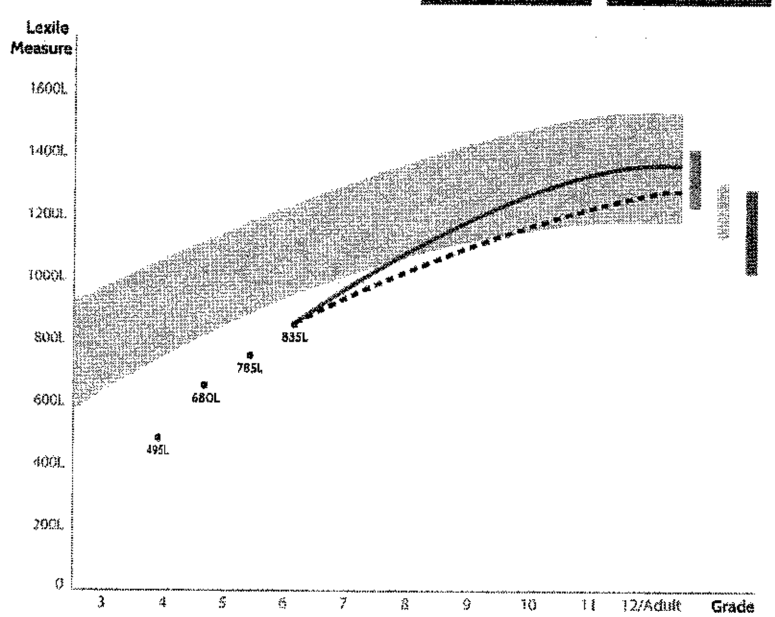
- Receive much-needed insight early in the process for their child's success in college and career so they can guide and direct as needed.

WHAT IS YOUR PROJECTED GROWTH?

The Lexile Growth Planner uses Lexile measures to display your recorded and projected growth in reading ability. To begin, click on the Add/Edit Measure button to enter your Lexile measure from statewide annual assessments.

YOUR GROWTH PLANNER RESULTS

Add/ Edit Measure | Add/ Edit Career

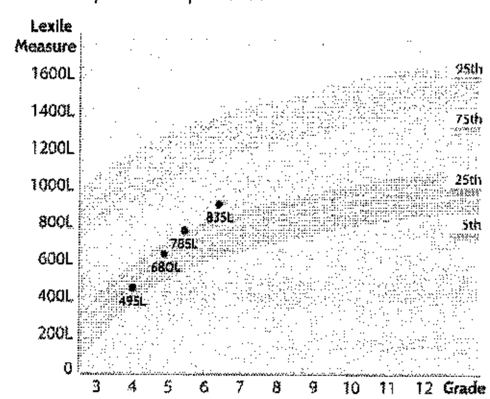


WHAT DO YOUR RESULTS MEAN?

Based on our estimated projection, you will likely need to devote extra attention to be

HOW DO YOU COMPARE NATIONALLY?

What does your score mean on a national level? Check your Lexile percentile.



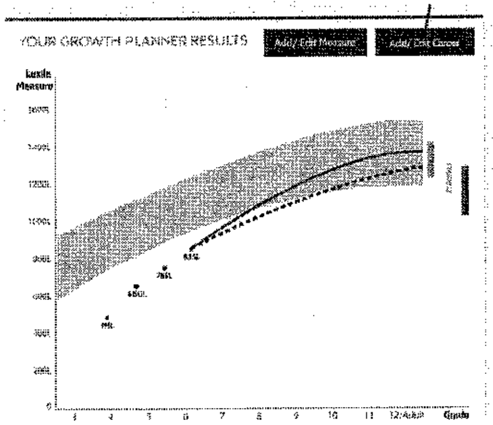
- Pathway for Reading Readiness
- Estimated Growth Path
- 600L Lexile Measure
- Recommended Growth Path
- University Readiness Range
- Community College Readiness Range
- Workplace Readiness Range

1 National Assessment of Educational Progress (NAEP). Nationsreportcard.gov. *NAEP - 2015 Mathematics & Reading Assessments*. [online] Available at: www.nationsreportcard.gov/reading_math_2015/#?grade=4 [Accessed 18 Jun. 2018].

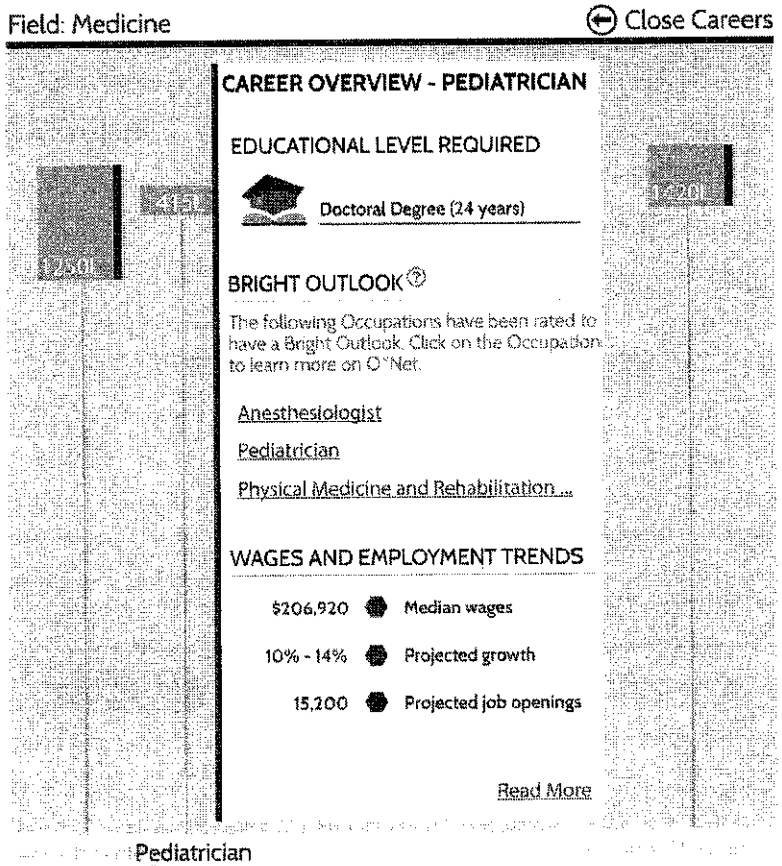
CAREER PLANNING MADE EASY

With this career planning feature, students can explore hundreds of careers and evaluate the outlook for that career, including educational requirements, salary range and job demand. Students can make job comparisons within the same field and see the reading (Lexile) and math (Quantile) levels associated with specific careers.

The CAREER PLANNING FEATURE is available in states that report Lexile and Quantile measures on annual statewide tests. It will be available soon on a subscription basis.



TAKE A CLOSER LOOK AT THE CAREER PLANNING FEATURE



HOW TO USE THE GROWTH PLANNER

It's as easy as 1, 2, 3.

GET STARTED AT CCR.LEXILE.COM | CCR.QUANTILES.COM

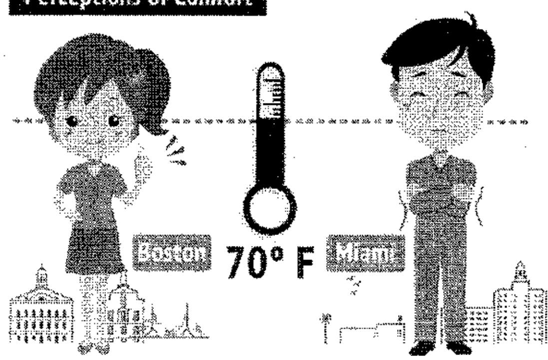
STEP 1
Enter Lexile or Quantile measures from annual statewide tests into the Growth Planner's graphing tool.

STEP 2
Forecast preparedness for college and career.

STEP 3
Encourage activities to promote reading and math growth and lay a strong foundation for college and career.

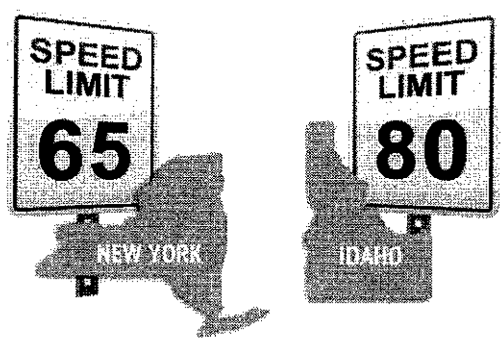
Comparability in the Measurement of Reading through The Lexile® Framework for Reading

Fahrenheit: Different Perceptions of Comfort



There are numerous examples of when it is easy to confuse underlying measurement scales with the labels we attach to various points on the measurement scales. For example, while 70° Fahrenheit is the same temperature in Boston as it is in Miami, the labels or judgments we make about the temperature are very different. Residents in Boston consider 70°F a "warm" day whereas residents in Miami consider it "cold."

Miles Per Hour: Different Maximum Speed Limits



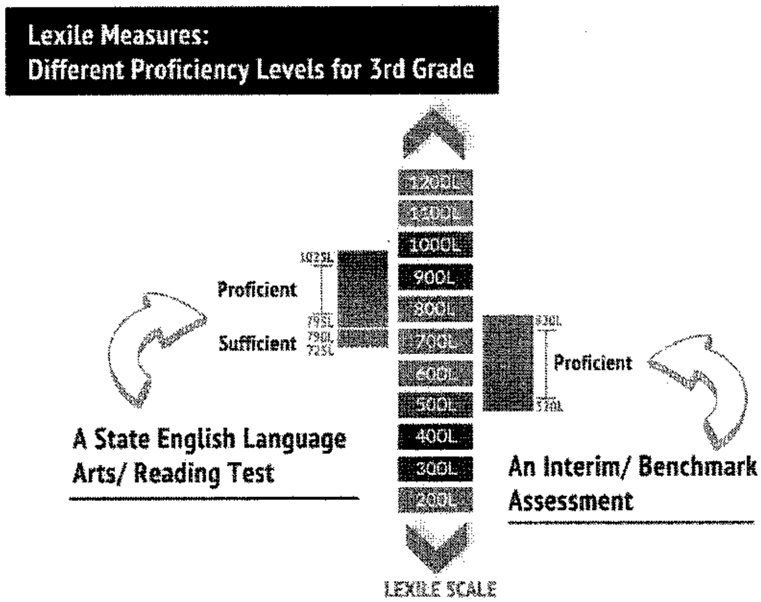
In another example, while the underlying measurement scale for reporting vehicular speed is the same across our nation, the maximum allowable speed limit across states varies from 60 mph to 80 mph. Consequently, what we consider "speeding" varies across states as well.

Inches: Different Height Requirements

When a child at a theme park asks "Can I go on this ride alone?" the answer is "If you're tall enough." But what does "tall enough" mean? This is a standard describing how tall you must be to go on the ride. In order to answer the question we need to know which theme park the child is in. "Tall enough" means different things in different theme parks. For example,

- The sign for the Space Mountain® ride at Walt Disney World® says you must be 44 tall to ride.
- The sign for the Space Mountain® ride at Disneyland® says you must be 40" tall to ride.

While there is judgment in labeling the weather cold or hot at 70° Fahrenheit, exceeding the maximum speed, or being tall enough to ride a certain ride, in each of these examples the measurement scale or metric allows for comparability. Driving 80 mph in North Carolina is the same as driving 80 mph in Montana, but the behavior has different labels and consequences. Just as we have comparability in the measurement scales in these examples, we have this same comparability in the realm of educational assessments that report Lexile measures. The good news is that **students reading at 725L on any assessment are reading at the same level.**



In the same way, states and assessment publishers have decided what scores describe proficient reading performance. Each state department of education or assessment publisher has their own definition of what it means to be "proficient" and uses that definition to define a particular range of scores to represent proficient performance. For example,

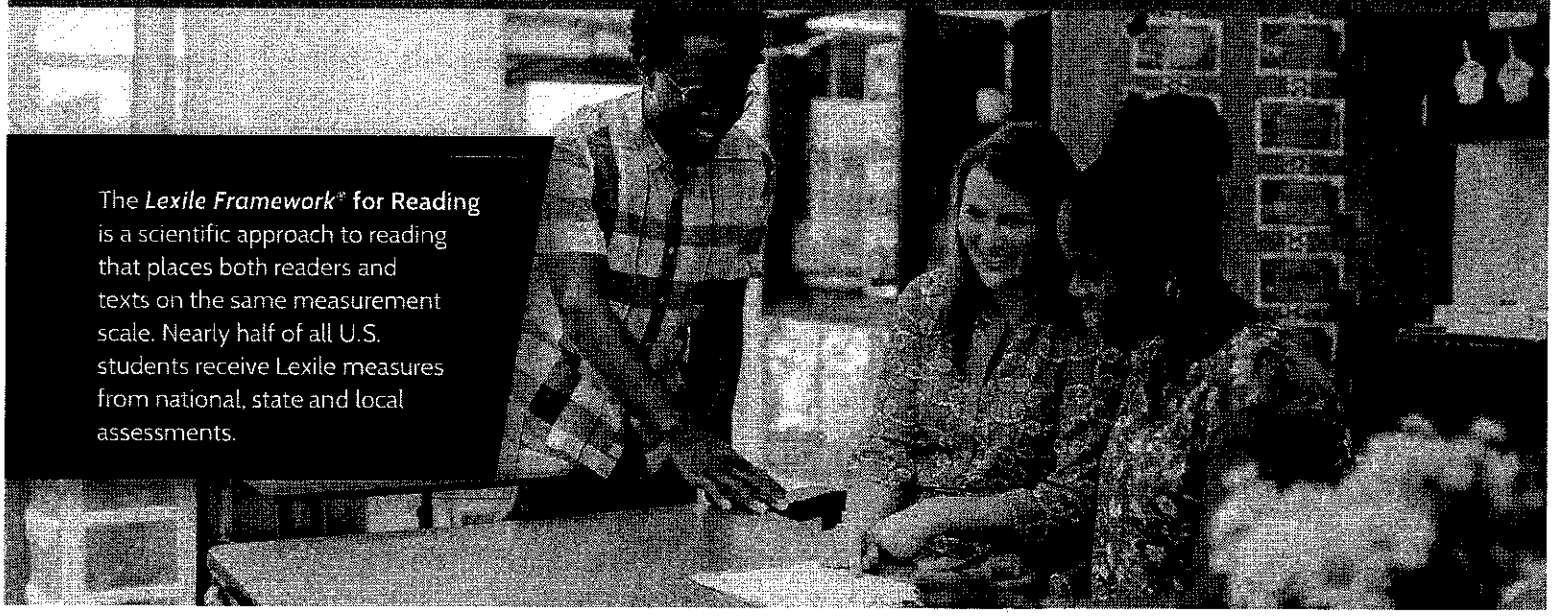
- On a state English Language Arts/Reading Test (Grade 3), the Sufficient range is from 439 to 441 (725L to 790L) and the Proficient range is from 442 to 451 (795L to 1025L).
- On one interim/benchmark assessment, the Proficient range is from 520L to 820L.
- On another interim/benchmark assessment, students reading at an appropriate level at the end of Grade 3 should be reading between 3.7 and 3.9 (495L to 537L).

Despite these different proficient standards, the underlying scale to measure a child's reading ability does not change.

Today, about twenty states provide Lexile measures as a part of reporting the state assessment results. And in each of these states, districts are using at least one of the dozens of interim and benchmark assessments that also report Lexile measures. Thus, by using a common scale such as the Lexile scale, we are providing more clarity and comparability in the measurement of reading.

LEXILE FRAMEWORK® FOR READING

EDUCATOR GUIDE



The *Lexile Framework*® **for Reading** is a scientific approach to reading that places both readers and texts on the same measurement scale. Nearly half of all U.S. students receive Lexile measures from national, state and local assessments.

What Are Lexile Measures?

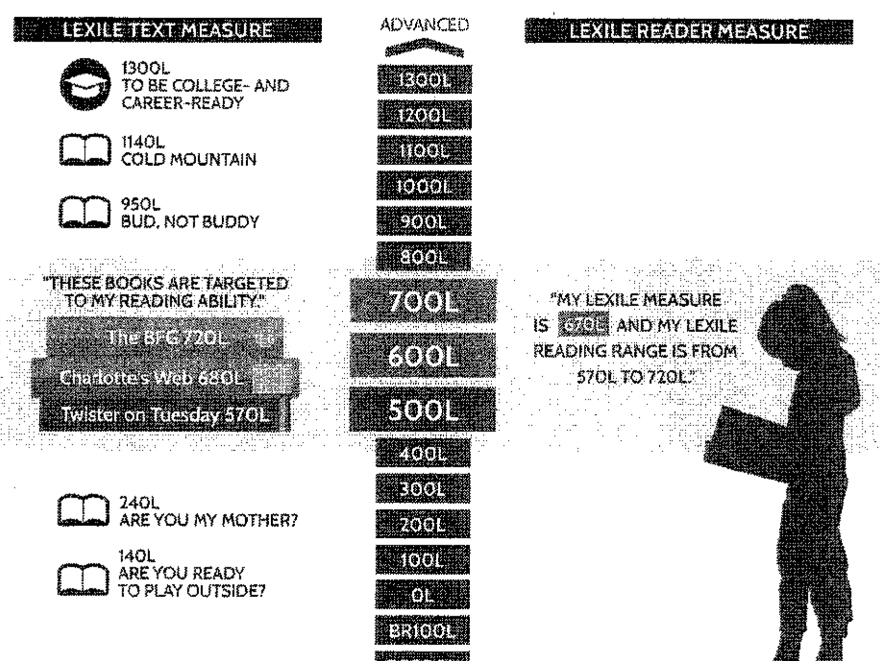
Lexile measures are represented by a number followed by an "L" (such as "800L") and range from below 0L for beginning readers and texts to above 1600L for advanced readers and texts. With Lexile measures, you can:

- **Find** just right books for independent reading.
- **Enhance** instructional planning of more challenging texts for all students.
- **Communicate and engage** with students and parents regarding reading progress.
- **Set goals and monitor** reading growth over time.

Match Readers With Texts on the Lexile Scale

A **Lexile reader measure** describes a student's reading ability. Connecting students with books in their Lexile range – 100L below to 50L above their reported Lexile measure – provides an ideal level of reading challenge.

A **Lexile text measure** tells you how challenging a text is to comprehend. Over 100 million books, articles and websites have Lexile text measures.

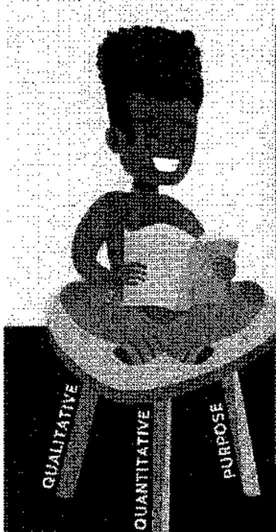


Beginning Reader (BR) is a code given to readers and texts that are below 0L on the Lexile scale. Learn more at Lexile.com/beginning-readers.

HOW EDUCATORS USE LEXILE MEASURES

Connect Students With Just Right Reading

Knowing the Lexile measure of a book and the Lexile measure of a student's reading ability can help identify reading materials at the right level of challenge. It's a good place to start, ***but nothing can substitute for your expertise as an educator*** and your knowledge of your student.



Support your student and grow their love of reading by considering...

1) QUALITATIVE FACTORS for the student (interests, motivation, age, maturity) and text (complexity of ideas and themes, style, quality, graphic supports).

2) QUANTITATIVE FACTORS such as Lexile measures.

3) PURPOSE for reading (assignment, pleasure, discovery, research, etc.).

Lexile measures should not be the only factor used when selecting a book. Expanding on college- and career-ready standards, here are our recommendations to consider.

Access Reading & Math Tools to Tailor Learning for Your Students

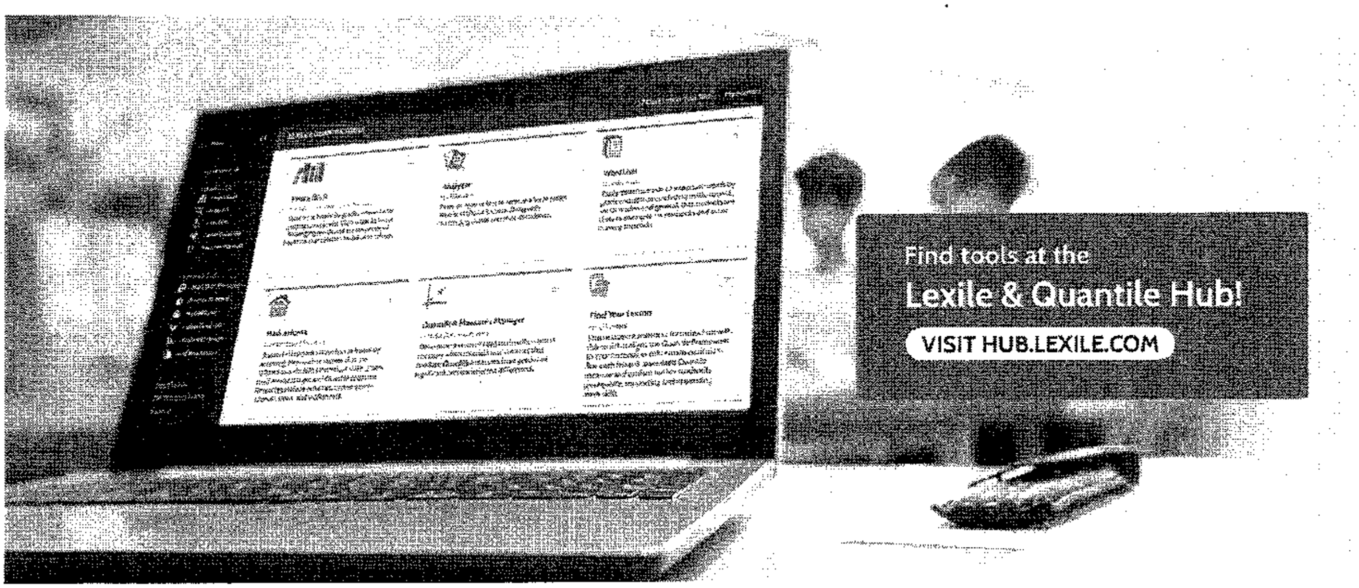
The ***Lexile® & Quantile®* Hub** is an online platform that provides easy access to reading and math tools. We listened to feedback from educators, parents and students to develop our tools. Reading tools include:

LEXILE FIND A BOOK
Identify just right books for students based on their Lexile level and interests.

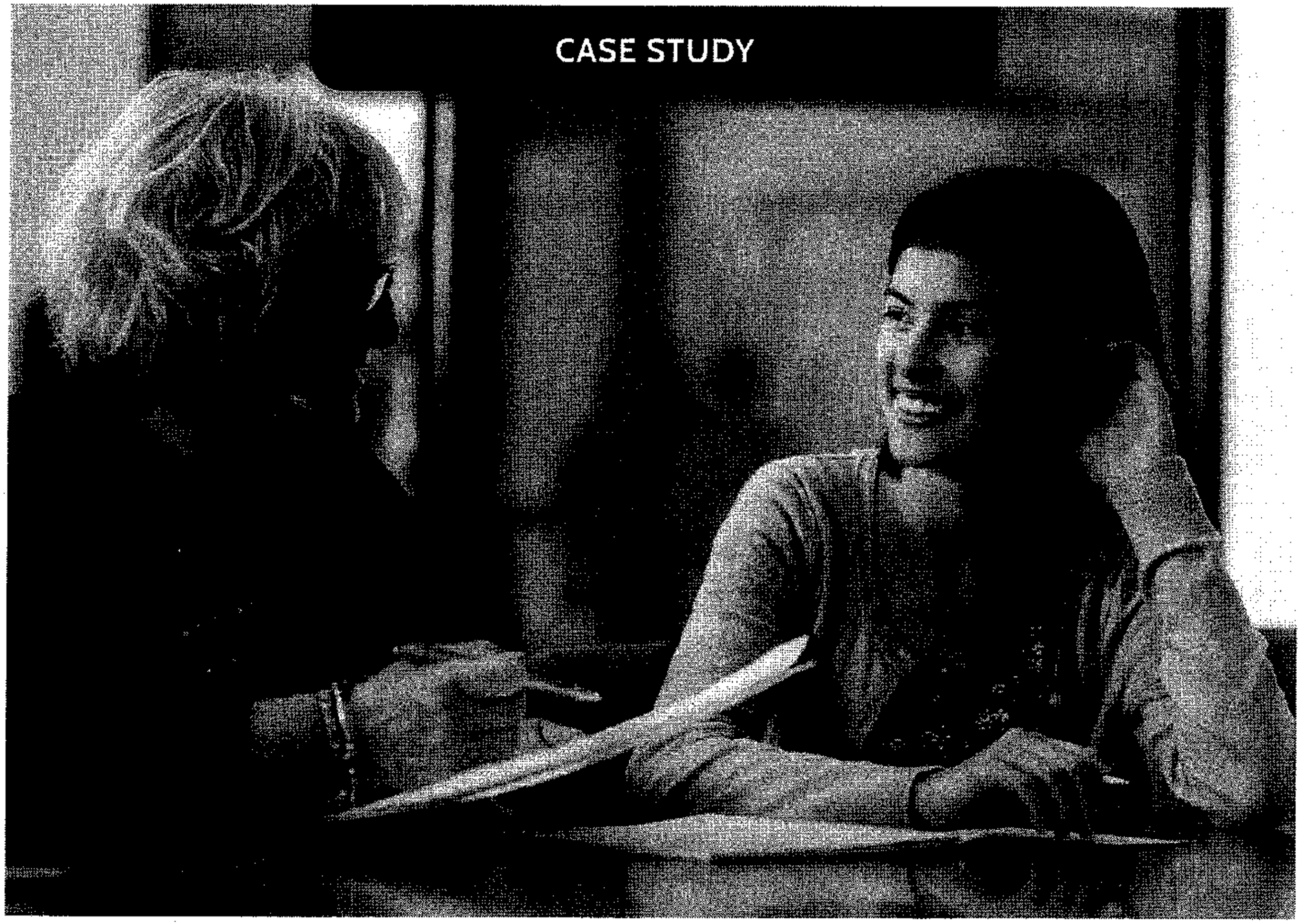
LEXILE ANALYZER
Determine the readability and challenging words of a text.

LEXILE WORDLISTS
Create custom lists of key vocabulary by grade level and domains.

LEXILE GROWTH PLANNER
Forecast growth and explore the reading demands and regional career information for hundreds of careers.



LEXILE GROWTH PLANNER QUANTILE GROWTH PLANNER



South Carolina School District Creates Personal Pathways to Prepare All Students for Future Careers

Considered the gateway to the South Carolina Lowcountry, Berkeley County is a popular relocation destination, especially for businesses. According to the U.S. Census Bureau, the county – less than an hour from Charleston – was the 17th fastest-growing in the country in 2017. Over the past two years, Berkeley County has experienced more than $1.5 billion in new economic development that comes with thousands of jobs for area residents.

This rapid growth presents both opportunities and challenges for the large district as it experiences a growth rate of nearly 1,000 students per year, said Berkeley County School District Chief Academic Officer Kevin O'Gorman, Ph.D.

"A diverse group of extremely large companies are bringing operations to the area and want to fill these new jobs from the community," he said. "However, our business partners told us that our students were lacking the

Berkeley County School District

(www.bcsdschools.net)

Moncks Corner, South Carolina

District Enrollment: PK-12
35,000 students, 43 schools

Superintendent:
Eddie Ingram, Ph.D.

Chief Academic Officer:
Kevin O'Gorman, Ph.D.

skills necessary for these jobs. We needed a strategy to put them on track to be prepared for these new employment opportunities."

With a deep commitment to Career and Technical Education (CTE), Berkeley County School District has a simple but ambitious goal: Every child will enter a career whether through college, trade program, certificate, military or straight into the workforce.

USING MEANINGFUL METRICS TO CREATE CAREER PATHWAYS

Understanding that strong reading and math abilities are at the heart of college and career readiness, district leadership turned to the Lexile Framework® for Reading and the Quantile Framework® for Mathematics, developed by MetaMetrics®, Inc. The Lexile Framework is a scientific approach to measuring reading ability and the text demand of reading materials on the same scale. Similar to the Lexile Framework, the Quantile Framework describes a student's mathematical ability and the difficulty of specific mathematical skills and concepts on the same scale.

In addition, Berkeley County is using MetaMetrics' Growth Planners, free online tools that allow students, educators and parents to forecast reading and math growth starting in third grade and compare that forecasted performance to text and mathematical demands of college and career.

A new state law, the South Carolina Read to Succeed Act, which passed in 2015, presented the district with new milestones for student literacy development. For example, third graders have to be reading on grade level by end of grade. In addition, district leaders were looking for a universal screener for IDEA and gifted and talented that was easy to use and provided teachers the information that they needed to individualize instruction. After being introduced to READ 180 and MATH 180 from Houghton Mifflin Harcourt, they began to learn more about Lexile and Quantile measures.

Beginning in the 2016-2017 school year, district leaders worked hard to implement Lexile and Quantile measures and make them a part of the academic culture at Berkeley County.

O'Gorman added, "In the meantime, our community was seeing this rapid economic growth and we had a huge push from the Charleston Chamber of Commerce for us to connect what is happening in school to preparing students for careers. Lexile and Quantile measures provide a common language for the academic and business communities to talk about student growth. Rather than talking about standardized test scores, we can look at Lexile and Quantile growth and, using the Growth Planners, talk about potential careers for our students."

Teachers are also using Lexile and Quantile measures in their classrooms to assess growth and personalize learning. "With Lexiles and Quantiles and the Growth Planners, we have tools for measuring ongoing student growth. It's not just about any one test on any given day," said O'Gorman.

PARTNERING WITH THE GROWING LOCAL BUSINESS COMMUNITY

School district leadership meets with the Charleston Metro Chamber of Commerce on a regular basis. The chamber provides an update on the community's growth and what the high-demand, high-skills jobs are going to be in the region. Then, O'Gorman and his team look at courses that will directly impact students developing entry level skills for those jobs.

"Every two years, we have to go back and rethink what programs we are offering and decide if we are meeting our local industry demands. To support this, it really helps for us to know if a student has the aptitude and interest for these local jobs," O'Gorman said.

As a result of this effort, the district revamped its "majors" and laid out student course paths for success in high school, four-year college degrees and other training that will help students develop the skills for these new jobs. New courses added include Introduction to Manufacturing, an expanded Certified Nursing Assistant program course offering that now includes summer school and other credentialing opportunities. On the horizon for the future is the possibility of adding a Dental Assistant program.

To support their progress, starting in eighth grade Berkeley County students develop an Individualized Graduation Plan (IGP) with a career goal and a pathway for achieving that goal. The plan is revisited every year until the student graduates from high school. The district has formally embedded Lexile and Quantile measures into those conversations with students which, O'Gorman said, "makes the process very clean and simple."

The district is also using a tool called STEMPremier, a social media platform for high school students that is similar to LinkedIn. With this tool, local businesses and colleges can pay to access student profiles, see their resumes and learn of their career interests.

As another demonstration of its commitment to connect with the local business community, for its 2018 summer leadership institute, district principals toured local businesses and heard directly from leaders and employees about the hard and soft skills that they are looking for in new employees. They toured businesses such as a Volvo plant, Google, a hospital, the electric utility and an automotive dealership meeting with IT professionals, technicians, doctors and nurses. The principals walked away understanding the need for even more apprenticeship opportunities, a greater understanding of the soft skills students need and more information about the entry level jobs that are available to Berkeley County students following graduation.

"We can meet students where they are – no matter where they are – and create personal pathways for each of them to make them future ready."

– Chief Academic Officer Kevin O'Gorman, Ph.D.

Learn More

To access the ***Lexile*® Growth Planner**, visit hub.lexile.com/lexile-gp.

To access the ***Quantile*® Growth Planner**, visit hub.lexile.com/quantile-gp.

PERSONALIZED PATHS TO STUDENT SUCCESS

Through its strong collaboration with the local business community and targeted strategies for supporting student growth as they plan for success in college and career, Berkeley County Schools is ensuring that all students navigate personalized learning paths.

O'Gorman said, "We can meet students where they are – no matter where they are – and create personal pathways for each of them to make them future ready. With the state providing Lexiles and Quantiles, our expanded collaboration with the business community, and tools such as the Growth Planners, this really has been the perfect storm for putting all of our students on the path to success."

MetaMetrics®

METAMETRICS RESEARCH BRIEF

Quantile Mathematical Demand for Career Accessibility

Gary L. Williamson, Senior Research Scientist; Lisa Bickel, VP, Strategic Partnerships-Quantile; Eleanor E. Sanford-Moore, SVP, Research & Development; Robert F. Baker, Director, Analytical Services; Ruth Price, Quantile Curriculum Specialist III; Heather Koons, Director, Research Services; Callie Totten, Project Coordinator; Paige Clark, Project Specialist

FEBRUARY 2018

OBJECTIVE

As evidenced by broad-based institutional initiatives (e.g., National Governors Association [NGA] Center for Best Practices, Council of Chief State School Officers [CCSSO], 2010) and United States Congressional actions (Every Student Succeeds Act [ESSA], 2015), recent educational policy discourse in the United States has promoted strategic interest in the mathematical content demands associated with college and career readiness. In previous work, Sanford-Moore et al. (2014) documented a lesson-difficulty continuum for K-12 mathematics and Williamson, Sanford-Moore and Bickel (2016) quantified mathematical content demands (and hence requisite student ability) for general college and career readiness. The purpose of this research brief is to quantify the demand of mathematical skills and concepts associated with access to specific postsecondary careers and occupations. In so doing, the student mathematics ability necessary for career *preparedness* or career *accessibility* can be inferred for specific individual careers and occupations through conjoint measurement. The current study does not focus on the mathematics demand of continuing successfully in a particular career, only the mathematics demand of career preparedness or accessibility.

Key Hypotheses: The study provides descriptive results only.

METHODS

Units of analysis:

The units of analysis for this study were mathematical skills, concepts or applications that were (a) classified as Quantile skills or concepts (QSCs) in The Quantile® Framework for Mathematics (see **Measures,** below) and furthermore (b) could be associated logically with entry into specific careers and/or occupations. Admissible mathematics content satisfied the following requirements for inclusion in the study.

- Mathematics skills/concepts selected for study were commonly required of individuals being trained for a specific career or in their first year of employment in the career. For example, mathematics skills/concepts might include (but are not limited to) those with the following characteristics:
    - Introduced in materials used for degree or certification programs required for career entry;
    - Commonly featured by professional organizations associated with the career;
    - Posted on websites of career-related professional organizations;
    - Contained in recruitment materials associated with the career; or,
    - Commonly used in materials, manuals, or references associated with on-the-job performance during the first year of employment in the career.
- Examples of the skills/concepts had to be readily available in print or editable electronic format.
- Examples of skills/concepts had to be classifiable as to source, purpose and/or use.

Career sample selection:

There were two primary sources for the vocational contexts examined in this study—Job Corps careers and O*NET Bright Outlook Occupations. Job Corps (Workforce Innovation and Opportunity Act [WIOA], 2014) is a residential education and job-training program for young adults ages 16-24 and is administered by the United States Department of Labor. The Job Corps program identifies careers by a Training and Achievement Record (TAR) Code and documents the mathematics skills and concepts associated with each career. O*NET is an online environment for vocational exploration (http://www.onetonline.org/find/career). The National Center for O*NET Development designates Bright Outlook Occupations using data collected by the U.S. Bureau of Labor Statistics (BLS). An occupation is called Bright Outlook if it (a) exhibits rapid growth, (b) has large numbers of job openings or (c) is a new and emerging occupation. O*NET identifies occupations with a Standard Occupational Classification (SOC) code.

Focal occupations for this study were selected in collaboration with Job Corps leadership and in the context of ongoing postsecondary career preparedness research conducted by MetaMetrics. Several attributes were considered favorable for selection.

- Careers that were of highest interest to Job Corps
- Bright Outlook Occupations identified by O*NET.
- Mathematics content containing calibrated Quantile skills and concepts
- Careers/Occupations previously examined with regard to text demands (e.g., Williamson & Baker, 2013)

Because The Quantile Framework for Mathematics currently does not designate or quantify QSCs associated with higher mathematics content (e.g., calculus and beyond), the study necessarily focused on careers that draw on the content typical of high school mathematics curricula or college curricula through precalculus. Consequently, careers that rely predominantly on knowledge of calculus or higher level mathematics concepts were excluded from this study. Conley (2010) pointed out that "a thorough understanding of the basic concepts, principles, and techniques of algebra" (p.37) is most important for success in higher mathematics content, thus giving credence to the relevance of the

Quantile Framework for characterizing the mathematics demand at the interface between high school mathematics and postsecondary mathematics beyond precalculus.

Some careers/occupations selected for the study belonged to both O*NET and Job Corps databases. Distinguishing such careers/occupations was facilitated by a crosswalk between Job Corps TAR Codes and O*NET SOC Codes. This research brief summarizes the mathematics demand for 103 careers/occupations analyzed as of October 2017.

Procedure:
Mathematics content was collected for selected careers/occupations according to the data collection protocols described above. Subject matter experts analyzed each content sample to identify the QSCs associated with the material. The Quantile Framework for Mathematics provided the Quantile® measure associated with each QSC. The mathematical demand of a career or occupation was then represented by a statistical summary (i.e., 25th percentile, median and 75th percentile) for its distribution of QSC measures.

In some cases, an individual QSC occurred more than once in the content sample for a specific career/occupation. The unique, non-duplicated QSCs constituted the analysis sample for each individual career/occupation. Unique, non-duplicated QSCs were used to represent a career/occupation because they best represent what happens in mathematics books on which the Quantile lesson continuum was based—that is, in most mathematics lessons the QSC of the lesson is unique and not repeated between lessons. We wished to closely follow this convention for characterizing careers, as we view the career space as an extension of the K-12 mathematics continuum.

The Quantile Framework was designed first for instructional purposes and uses a 50% matching convention by default for interpreting the relationship between person ability and mathematics content demand. With this convention, whenever a student's mathematics ability (as represented by his or her Quantile measure) matches the mathematics content demand (Quantile measure) of a particular mathematics skill or concept, the Quantile Framework forecasts 50% understanding for the student with respect to the particular skill/concept. When students demonstrate 50% understanding, it means they are ready for instruction on the skill/concept.

Yet, students who aspire to enter a career are expected to be ready to perform (as opposed to being ready for instruction). Consequently, for this study, we moderated the matching convention of the Quantile Framework so that Quantile measures related to career accessibility are based on a 75% matching convention. Thus, when the difficulty of a QSC is reported for mathematical skills or concepts required in a career context, the Quantile measure represents the mathematics ability a person must possess to have 75% understanding. All career Quantile measures reported in this research brief have been moderated accordingly.

Measures:
The primary measure used for the analysis is the Quantile measure provided by The Quantile Framework for Mathematics (MetaMetrics, 2011). The Quantile Framework is a scientific approach to measuring mathematics achievement and concept/application solvability. The Quantile Framework consists of a Quantile measure and the Quantile® scale. A Quantile measure represents the difficulty of a mathematical skill, concept or application (QSC) as well as a developing mathematician's understanding of the QSCs in the areas of Geometry; Measurement; Number Sense; Numerical Operations; Algebra and Algebraic Thinking; and Data Analysis, Statistics, and Probability. Quantile measures are expressed with numerals followed by a "Q" (for example, 850Q), and are quantitatively calibrated on the Quantile scale. The Quantile Framework spans the developmental continuum from kindergarten mathematics through the content typically taught in Algebra II, Geometry, Trigonometry, and Precalculus. Quantile measures range from below 0Q (Emerging Mathematician) to above 1600Q. At present, the Quantile Framework does not designate or calibrate QSCs for higher mathematics content (e.g., calculus and beyond).

ANALYSES

Four analyses were completed using the database of 103 careers/occupations selected for the study. First, for each Job Corps career and each Bright Outlook Occupation, selected percentiles were calculated for the distribution of associated QSC difficulties. The median (50th percentile) Quantile measure provides a one-number summary of the typical mathematics demand associated with career preparedness. The 25th percentile and 75th percentile establish the boundaries of the interquartile range (IQR), which spans the middle 50% of mathematics demand for each career/occupation and provides a simple measure of the variability in mathematical demand within a career or occupation. In the second analysis, a bivariate plot of median career math demand measures versus years of education required for career entry is presented for the 103 selected careers/occupations. In the third analysis, the data were aggregated by sixteen career clusters identified by the United States Department of Education (USED). For each career cluster, the median Quantile measure and the 25th and 75th percentiles of the distribution are presented graphically. Finally in the fourth analysis, results from a parallel study of career text complexity (Williamson & Baker, 2013) are combined simultaneously with the present results to show postsecondary reading and mathematics career demands in conjunction with student reading and mathematics growth during the public school years. Results for the four analyses are presented graphically in four figures, which are located at the end of this research brief and discussed in the next section.

RESULTS & DISCUSSION

Figure 1 summarizes the mathematical demand associated with career accessibility for the 103 careers/occupations in the study. The horizontal scale indicates Quantile measures and the vertical scale represents the continuum of careers and occupations. In Figure 1, careers/occupations are displayed in order of increasing median Quantile measure. Thus, as the reader moves upward in the list of careers and occupations, the typical mathematics demand (or median QSC difficulty) increases monotonically with each successive career/occupation. It is obvious in Figure 1 that all careers and occupations in the study exhibit substantial variability in the mathematics demand (QSC difficulty) associated with career preparedness. The typical IQR (i.e., width of box plot) is approximately 400Q[1]. As a result, there is considerable overlap in the distributions of Quantile measures for many of the careers/occupations.

Eight specific careers and occupations are identified in Figure 1 to show how career accessibility may become more challenging as mathematical demand increases. From lowest median Quantile measure to highest, the exemplars are: Tax Preparers; Tellers; Financial Managers; Structural Metal Fabricators and Fitters; Optometrists; Civil Drafters; Solar Thermal Installers and Technicians; and, Diagnostic Medical Sonographers.

In Figure 2, the 103 careers and occupations are displayed in another fashion. The median mathematics content demand of each career and occupation is plotted versus the number of years of education required for entry into that career or occupation. Based on analogous research for reading (Williamson & Baker, 2013), one may expect there to be an approximate linear relationship between the mathematical demands of career accessibility and the number of years of education required to enter the career or occupation. Although the plot is visually suggestive, the statistical relationship between median Quantile measure and years of education (required for career entry) is too weak in these data to ascribe any relationship. This may be due in part to the fact that the conditional distributions in Figure 2 are truncated at the upper end due to the absence of careers that draw on higher mathematics content for the analyses in this study.

Number of years of education required for career entry was drawn primarily from O*NET. When data were not available from O*NET, our study used the BLS Occupational Outlook Handbook (OOH) (https://www.bls.gov/ooh/). There was one career we could not find in the OOH. For that career, we consulted My Next Move (https://www.mynextmove.org/).

In Figure 3, box plots summarize the distributions of QSC difficulties by USED career cluster. In this analysis, duplicate QSCs were allowed within a cluster to reflect the fact that a unique QSC could occur in multiple occupations, which in turn were classified into a single USED cluster. The horizontal axis in Figure 3 represents the Quantile scale. USED career clusters are alphabetically ordered (from top to bottom) along the vertical axis. In Figure 3, it appears that the median mathematical demand is lowest in the Hospitality and Tourism cluster (800Q) and highest in the Science, Technology, Engineering, and Mathematics cluster (1170Q). There is substantial variability of mathematical demand within clusters, with the within-cluster IQRs ranging from 360Q (Architecture and Construction) to 580Q (Human Services).

Four distinct perspectives are depicted in Figure 4: (a) student reading growth, (b) student mathematics growth, (c) career reading demands and (d) career mathematics demands. The horizontal axis in Figure 4 denotes both grade in school and years of education required for career entry. There are two vertical axes. The left vertical axis is calibrated in terms of the Lexile® scale to accommodate the graphical depiction of student reading growth and career reading demand, while the right vertical axis is calibrated in terms of the Quantile scale to accommodate the depiction of student mathematics growth and career mathematics demand. Furthermore, reading person measures and text-complexity measures are coded in blue, while mathematics person measures and mathematical demand measures are coded in red. Thus, the blue elements in Figure 4 should be referenced to the left vertical axis (Lexile scale) while the red elements in Figure 4 should be referenced to the right vertical axis (Quantile scale). These four graphical elements have the following additional meanings.

- **Student reading growth.** The blue curve in Figure 4 represents the average reading growth of 101,610 students from the end of Grade 3 to the end of Grade 11. All students were enrolled in the public schools of a southern state whose reading performance on the National Assessment of Educational Progress (NAEP) was at or above the national average during the time frame. The curve is specified by an unconditional quadratic multilevel growth model that was statistically fitted to purely longitudinal data (i.e., serial measures on the same students across occasions) for all students who progressed from grade to grade without repeating a grade and had at least one score during the study time frame (2002-2010). The longitudinal data for reading and mathematics growth were collected within the same state.
- **Student mathematics growth.** The red curve in Figure 4 represents the average mathematics growth of 101,650 students from the end of Grade 3 to the end of Grade 11. All students were enrolled in the public schools of a southern state whose mathematics performance on the National Assessment of Educational Progress (NAEP) was above the national average during the time frame. The curve is specified by an unconditional quadratic multilevel growth model that was statistically fitted to purely longitudinal data (i.e., serial measures on the same students across occasions) for all students who progressed from grade to grade without repeating a grade and had at least one score during the study time frame (2002-2010). The longitudinal data for reading and mathematics growth were collected within the same state.
- **Reading career demand.** Among the 101 careers and occupations that comprised the sample for this study, there were 77 careers/occupations for which both reading career demands and mathematics career demands were quantified. The blue diamonds in Figure 4 represent the median reading career demands for those 77 careers/occupations.
- **Mathematics career demand.** The red circles in Figure 4 represent the median mathematics career demands for the same 77 careers/occupations for which reading career demands are represented.

It should be understood that the Lexile scale and the Quantile scale are independent scales and are not psychometrically linked. Thus, reading growth cannot be directly compared to mathematics growth because the Lexile scale unit and the Quantile scale unit are not equated. So, for

[1] For comparison, consider that the average mathematics growth of students from the end of Grade 4 to the end of Grade 9 is 399Q (Williamson, 2016).

example, we cannot conclude from Figure 4 that reading performance is higher than mathematics performance simply because the reading growth curve lies higher on the Lexile scale than the mathematics growth curve lies on the Quantile scale. Neither can we infer that the reading demand of careers and occupations is higher than the mathematics demand because the blue diamonds in Figure 4 lie mostly above the red circles.

However we can draw conclusions from Figure 4 regarding the relationship between reading elements (reading growth vs. career reading demands); similarly Figure 4 supports inferences regarding the relationship between mathematics elements (mathematics growth vs. career mathematics demands). And, we can observe in Figure 4 that the correspondence between growth and career demand is somewhat different for reading than it is for mathematics. In particular, we note that the reading growth curve seems to align well with the median reading demands of these 77 careers/occupations because the curve rises to a point that appears to align roughly with the middle of the cloud of points (diamonds) representing the median reading demands for those 77 careers/occupations. On the other hand, it appears that the mathematics growth curve represents a trajectory that will surpass the mathematics demands of most of the 77 careers/occupations. There is a plausible reason that this inference is correct—namely, the fact that the study focused on more easily accessible careers and excluded careers whose mathematics demand derived predominantly from higher mathematics content (e.g., calculus). As the Quantile Framework is extended to accommodate higher mathematics content and QSCs are identified for calculus and higher level postsecondary mathematics courses, we should be able to include careers with higher mathematics demands, which in turn will pose higher challenge for students graduating from the public schools. At the very least, we may say that the reading and mathematics growth in this particular state seem to be well-aligned with the reading demands and the mathematics demands associated with career accessibility. Such an inference is a key advantage of using the conjoint measurement scales featured in the Lexile and Quantile Frameworks.

Limitations of the research:

As with nearly all research there are some limitations of this study that should be acknowledged. These relate to sample design, representation of careers in the sample, representation of the mathematics skills and concepts in the careers and occupations, the substantive scope of the study and the contemporaneous nature of the content included for analysis. All of these ultimately have some unknown impact on the generalizability of the results reported here.

- **Sample design.** The study employed a three-stage hierarchically nested convenience sample with QSCs nested within careers and occupations, which in turn were nested within sixteen official USED career clusters. However, a given QSC could occur in multiple careers/occupations; and, a specific career/occupation could appear in multiple USED clusters. The target population of careers originally included Bright Occupations identified by O*NET as of December 2012 and occupations identified by Job Corps. However, only a portion of the careers/occupations in these two sources have been analyzed to date. The choice of which careers/occupations to examine first was based largely on availability of materials. Thus, the sample was not a probability sample and traditional sampling theory cannot be used to calculate precision, construct confidence intervals or infer statistical significance of the results.
- **Representation of careers/occupations and QSCs.** The careers and occupations included in this study were drawn selectively from O*NET and Job Corps. So it is unlikely that the particular careers and occupations included in this study are representative of the entire universe of workplace careers and occupations. For example, the Bright Outlook occupations identified in December 2012 constituted only about a third of the entire O*NET occupational database. Further, the Bright Outlook occupations were by definition different from other occupations in the database as they were characterized by fast growth or large numbers of openings, or by being new, emerging occupations. Also, the careers/occupations selected for this study are not necessarily distributed proportionally with respect to the USED career clusters (e.g., currently there is very little representation in the Hospitality and Tourism cluster compared to other clusters). Consequently, inferences regarding the career mathematics demands of USED career clusters should be viewed with some caution.
- **Substantive scope.** As mentioned at the beginning, this study focuses on career preparedness or career accessibility, not sustained success within a career. So when we refer to the mathematics demand of a career or occupation, we mean the demand associated with *entry* into the career or occupation and not the demand experienced once on the job for an extended period of time (e.g., more than a year). In addition, careers and occupations predominantly requiring higher mathematics content were not included in the study because The Quantile Framework for Mathematics at present does not extend to QSCs that characterize higher mathematics content. As a result, some of the observed distributions of career mathematics demands are likely attenuated to an unknown extent.
- **Database evolution.** For this research, we used nationally respected databases to identify careers and occupations containing mathematics content demands. However, these databases are dynamic. Already O*NET has identified additional Bright Outlook Occupations and de-identified others. In addition, data elements such as years of required education are periodically updated as new information becomes available. Consequently, the results reported in this research must be regarded as provisional and subject to revision in the future. We anticipate that our Quantile Career Database will evolve in two substantively important ways: (a) to incorporate new careers and occupations and (b) to expand the content collections used to identify relevant QSCs.

Significance of the research:

Having identified some limitations of the research and called attention to cautions that are warranted for interpretation, we also feel obliged to point out advantages of the approach we used in this research and the benefits it could bring to educational and career planning applications.

- **Conjoint measurement.** The Quantile Framework for Mathematics was designed to provide conjoint measurement of an individual's mathematical ability and the demand of mathematical skills or concepts. It is thus possible to place both the person and the content demand on a common quantitative scale. This facilitates instructional applications by allowing teachers to match individuals with content of an appropriate challenge level for optimal instruction. It can similarly allow students, parents, and educators to judge a student's readiness to perform on skills and concepts required at job entry. Although mathematical ability may not be *sufficient* for career accessibility, it is unquestionably a *necessary* requirement for careers or occupations that demand competence with mathematics content.

- **Establishing a baseline.** In this research brief we have presented a first attempt to systematically quantify the mathematics content demands of specific careers and occupations. The results encompass (a) analyses of the mathematics content of well-defined sources as well as (b) quantification of the demand of rigorously-classified mathematical skills and concepts. As with any first attempt, there is room for future improvement. Yet this should not diminish the importance of establishing a baseline view of mathematics career demands which may inform both public instruction and personal career planning efforts going forward.

REFERENCES

Conley, D. T. (2010). *College and career ready: Helping all students succeed beyond high school.* San Francisco, CA: Jossey-Bass.

Every Student Succeeds Act (ESSA), Pub. L. No. 114–95, § 129 Stat. 1802 (2015).

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

MetaMetrics. (2011). *The Quantile Framework® for Mathematics Technical Guide.* Durham, NC: Author.

Sanford-Moore, E. E., Williamson, G. L., Bickel, L., Koons, H., Baker, R. F. , & Price, R. (2014). *A quantitative task continuum for K-12 mathematics*. (MetaMetrics Research Brief). Durham, NC: MetaMetrics.

Williamson, G. L. (2016). Novel interpretations of academic growth. *Journal of Applied Educational and Policy Research, 2*(2), 15-35.

Williamson, G. L., & Baker, R. F. (2013). *Enriching the Concept of Career Preparedness by Examining Text Complexity Associated with Bright Outlook Occupations* (MetaMetrics Research Brief). Durham, NC: MetaMetrics.

Williamson, G. L., Sanford-Moore, E., & Bickel, L. (2016). *The Quantile® Framework for Mathematics Quantifies the Mathematics Ability Needed for College and Career Readiness* (MetaMetrics Research Brief). Durham, NC: MetaMetrics.

Workforce Innovation and Opportunity Act (WIOA), Pub. L. No. 113-128, § 128 Stat. 1425 (2014).

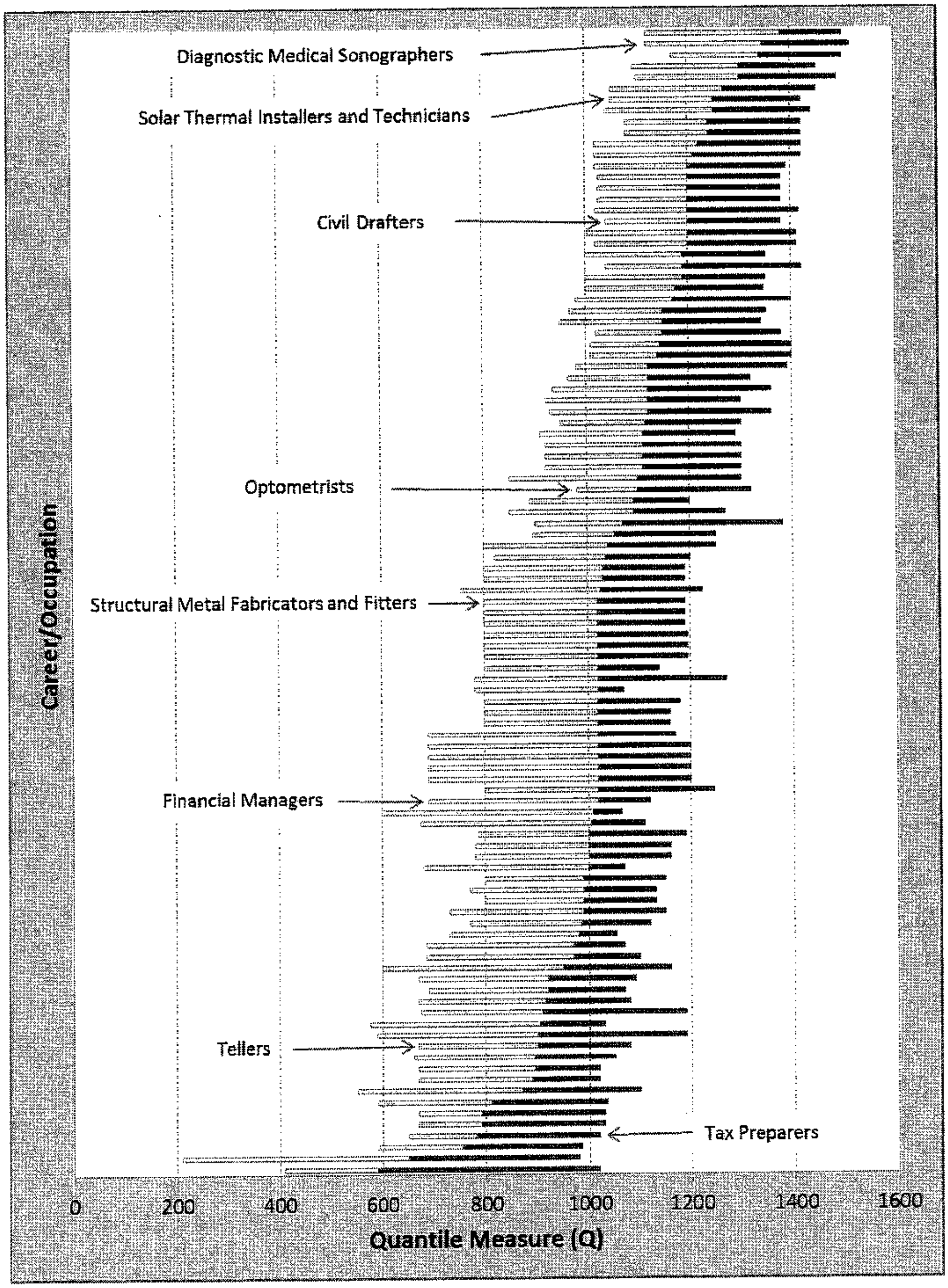


Figure 1. Mathematical content demands (difficulty of mathematical skills and concepts as measured on the Quantile scale) for 103 careers and occupations: Medians and interquartile boundaries.

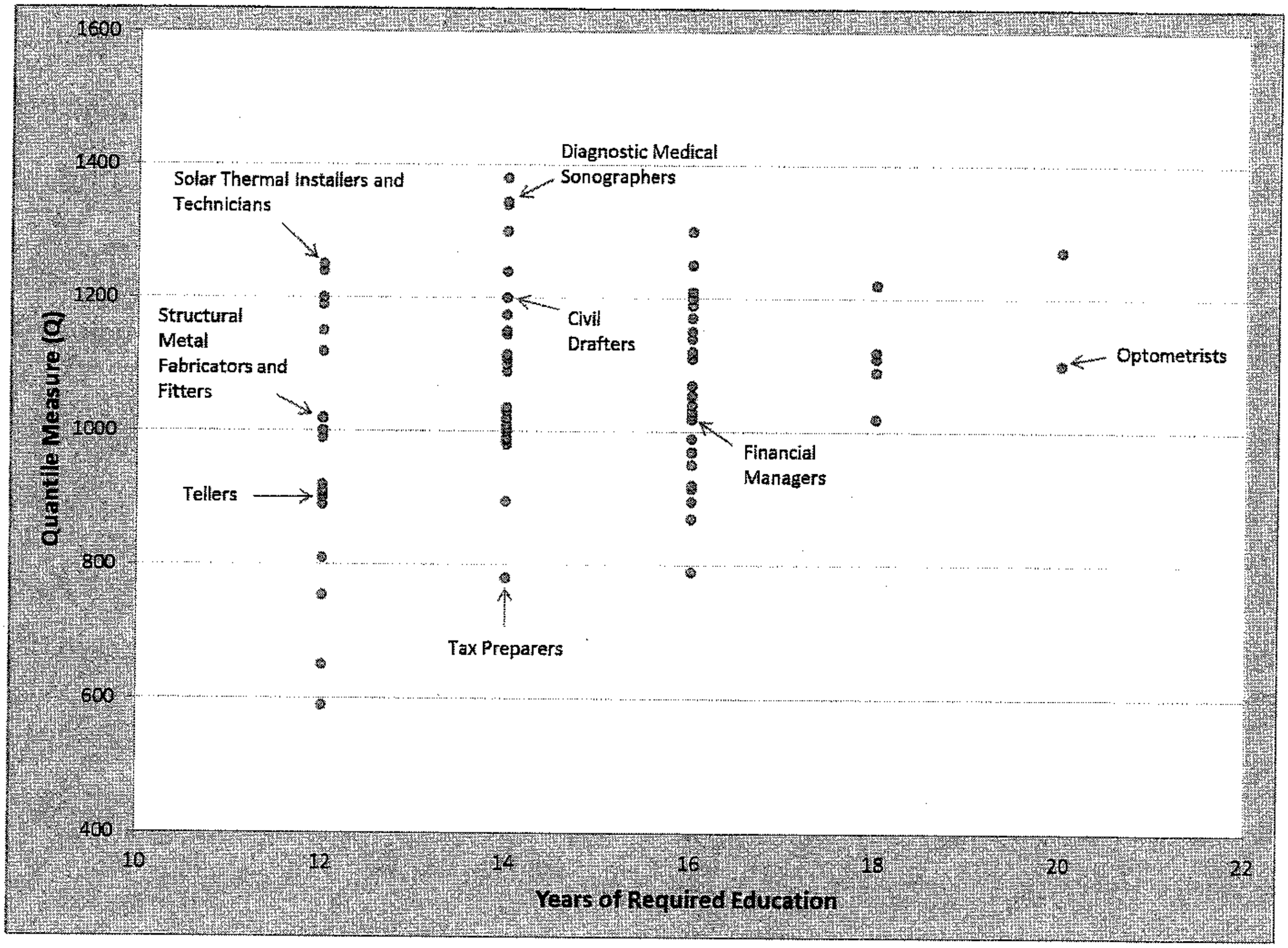


Figure 2. Median mathematics content demand (difficulty of mathematical skills and concepts as measured on the Quantile scale) versus years of education required to enter 103 careers and occupations.

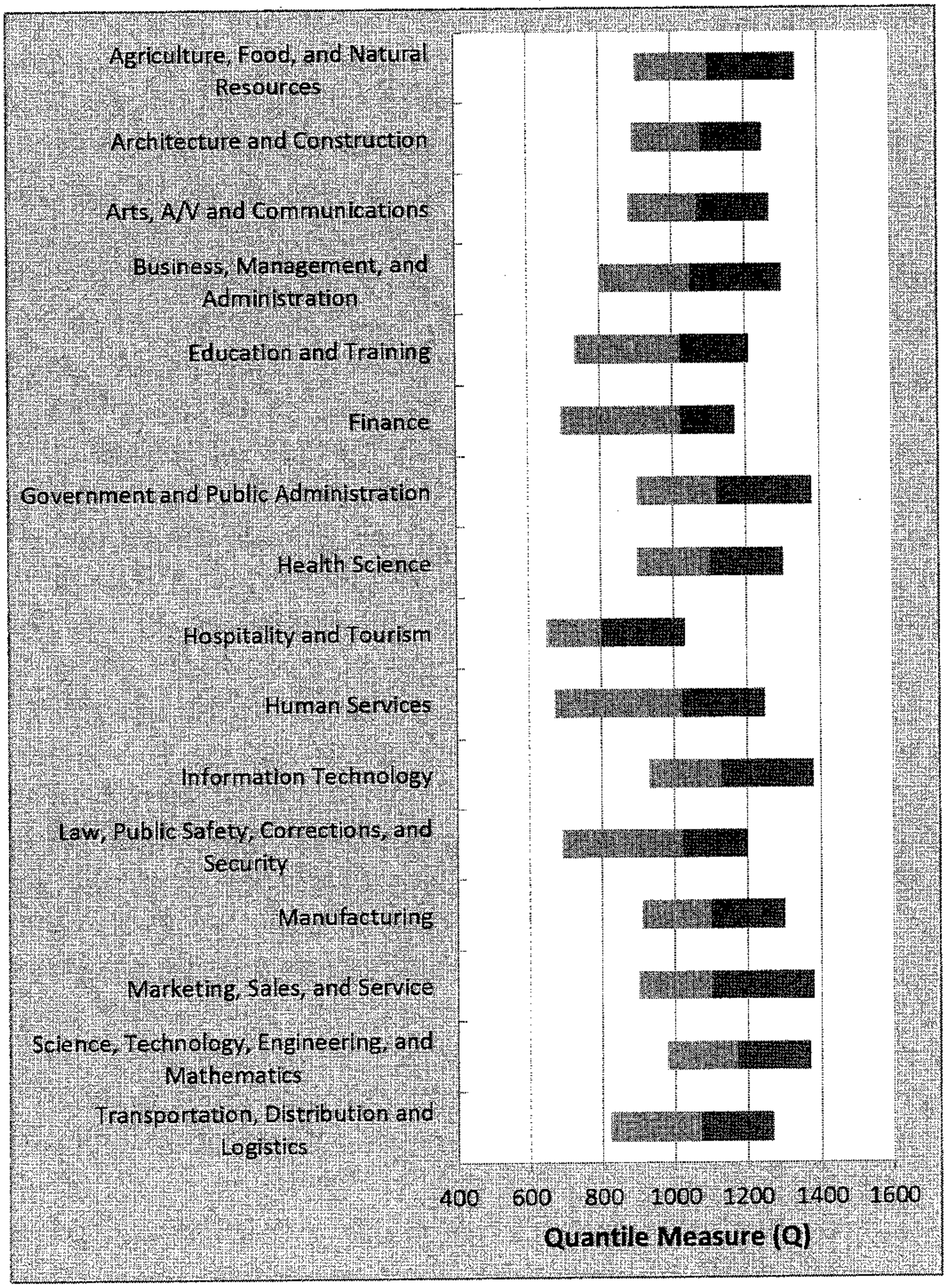


Figure 3. Mathematics content demand (difficulty of mathematical skills and concepts as measured on the Quantile scale) for 16 career clusters designated by the United States Department of Education: Medians and interquartile boundaries.

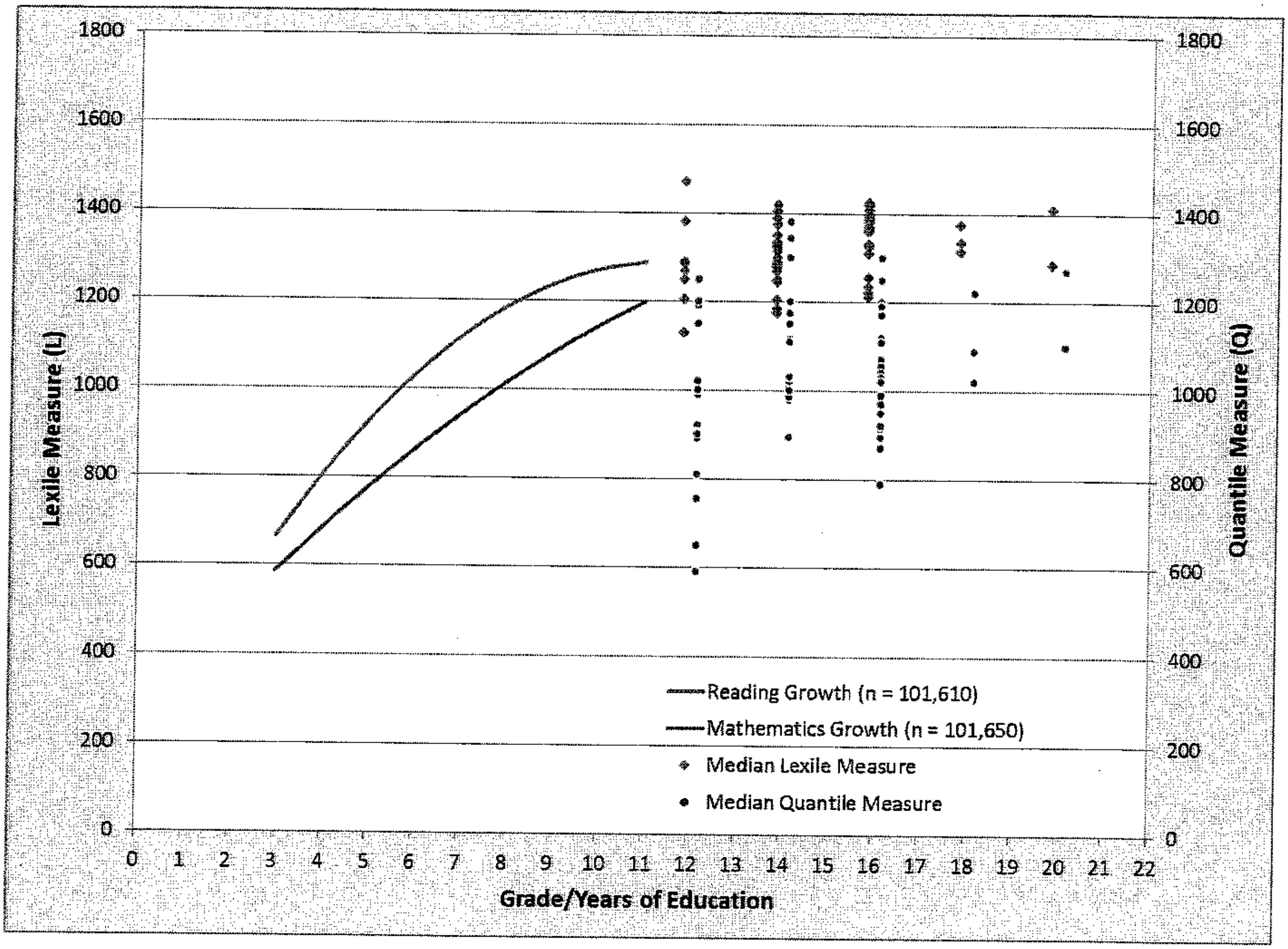


Figure 4. Average student reading and mathematics growth curves in relation to median reading demand (complexity of text as measured on the Lexile scale) and mathematics demand (difficulty of mathematical skills and concepts as measured on the Quantile scale) for 77 careers and occupations.

For more information, visit www.MetaMetricsInc.com.

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

MetaMetrics.

JAEPR, 2(2), 15-35

ISSN: 1930-9325

RESEARCH ARTICLES

Novel Interpretations of Academic Growth

Gary L. Williamson
MetaMetrics

Integrating a construct theory with Rasch measurement not only places persons and tasks on a common scale, but it also resolves the indeterminacy of scale location and unit size when the scale is anchored in an operationalized task continuum based on the construct theory. Such an approach has several advantages for understanding academic growth as evidenced in a series of empirical examples, which demonstrate how to: a) conjointly interpret student reading growth in the context of reading materials concomitantly used during instruction; b) interpret a reading growth trajectory in light of future (e.g., postsecondary) reading requirements; c) forecast individual reading comprehension rates relative to both contemporary and future text complexity requirements; and d) create growth velocity norms for average academic growth in reading or mathematics achievement.

Keywords: achievement; reading; mathematics; longitudinal; growth; norms

During the 1980's two measurement companies in the United States introduced a fundamental innovation in the scaling of student reading ability that moved the world closer to an absolute framework for the measurement of reading comprehension. The strategy entailed combining the Rasch measurement model with an operationalized reading construct theory. As with other item response theory (IRT) models, the Rasch model makes it possible to place persons and tasks (items) on a common scale, but certain scale properties (location, unit size) are arbitrary (i.e., they vary with changes in the person sample and/or items). The key innovation involved two steps that anchored the scale and defined its unit size in terms of an empirical text complexity continuum.

The first step was to define and validate a construct model that operationalized the reading difficulty of texts in terms of specific semantic and syntactic features of texts that are effective proxies for the cognitive demand experienced by readers while reading. Secondly, it was demonstrated that the empirical difficulties of a well-defined, text-based item type could be nearly perfectly predicted by the complexities of the texts associated with the items. Once students and items were measured via the Rasch model, the construct theory was used to calibrate the items to the text complexity continuum. This produced a direct correspondence between the person measures and the text measures.

The company now known as Questar Assessment, Inc. was the first to use this type of approach. They developed the Degrees of Reading Power® (DRP®) Program, which reports student reading measures from criterion-referenced tests on a proprietary DRP Scale of Text Complexity, which it uses to measure the reading difficulty of printed material (Bruning, 1985). Nelson, Perfetti, D. Liben and Liben (2012) described the scale as follows:

> DRP text difficulty is expressed in DRP units on a continuous scale with a theoretical range from 0 to 100. In practice, commonly encountered English text ranges from about 25 to 85 DRP units, with higher values representing more difficult text. (p. 11)

Questar defined a DRP *prose comprehension model* based on the application of the Bormuth (1969) readability formula to measure text complexity. Their reference item type was a text-embedded, cloze item (i.e., based on a text passage with certain words removed) administered according to a specific protocol. The unit size of the DRP scale was specified in terms of a transformation of the Bormuth text complexity measure, *R*. Research has shown that the DRP scale places both student reading ability and text complexity on a common well-defined, unidimensional scale that remains invariant over time. Thus, research supports the claim that the DRP tests "are like measures in the natural sciences." (B. L. Koslin, Zeno, & Koslin, 1987; p. 171)

At nearly the same time, a second company pursued the same fundamental idea. MetaMetrics® developed The Lexile® Framework for Reading to measure both readers and texts on a common scale. They independently developed a *construct-specification equation* to operationalize text complexity and predict item difficulties (Stenner & Smith, 1982; Stenner, Smith & Burdick, 1983). They also developed a well-defined reference item type (consisting of a text passage followed by a cloze-like, sentence-completion stem) and demonstrated that the empirical difficulties of such items could be nearly perfectly predicted by the difficulties of the associated texts (Stenner, D. R. Smith, Horabin & Smith, 1987). They coupled this construct model with a Rasch measurement model to place both a student's reading ability and a text's readability on a common invariant scale.

In order to define a logical unit for the Lexile scale, MetaMetrics chose to explicitly anchor its scale at two points on the text-complexity continuum. Based on its anchoring, a Lexile scale unit equals 1/1000 of the difference between the readability of certain specific basal primers and the readability of an online adult encyclopedia (Stenner, H. Burdick, Sanford, & Burdick, 2007). This approach provided a well-defined unit of measurement that retains its absolute size across different applications of measurement. It may be noted that this method is directly analogous to the way the meter was standardized based on the length of the meridian quadrant (i.e., the distance from the North Pole to the equator) through Paris (Legendre, 1805). It is also precisely analogous to the way that temperature scales are anchored.

Because the Lexile Framework and the DRP are based on a Rasch measurement model, they are examples of conjoint measurement. Conjoint measurement makes it possible to simultaneously scale two variables that jointly predict an outcome. For example, reader ability and text difficulty jointly predict reading comprehension; so, both the reader measure and the text difficulty measure can be placed on a common scale. Thus, both the Lexile Framework and the DRP can be utilized to generate student scores that are reported on a text difficulty continuum, giving the scores supplemental interpretability anchored in a real-world context. Since their creation, both systems have been widely implemented in the United States. The primary use of both these systems to date appears to have been the matching of students with texts of appropriate difficulty.

In 2004, MetaMetrics launched The Quantile® Framework for Mathematics, a measurement system for mathematical understanding, which uses Rasch measurement to conjointly scale both persons and items and anchors the resulting scale in a real-world task continuum. The Quantile Framework uses a quantified mathematics lesson continuum as the real-world context for anchoring the developmental scale (Sanford-Moore et al., 2014). As a companion scale to the Lexile Framework, the Quantile Framework demonstrates that the strategy of combining Rasch measurement with construct theories and anchoring the resulting scales in real-world task continua is a viable method for behavioral science measurement which generalizes to multiple constructs. As was the case with the Lexile Framework, the Quantile Framework was primarily designed to link assessment with instruction (MetaMetrics, 2009).

The purpose of this paper is to demonstrate, through several examples, that interpretations of student academic growth benefit from the use of Rasch-based measurement scales that have been anchored in a real-world task continuum by means of construct theory. These examples benefit from the fact that one state had the foresight and commitment to utilize such scales over a long period of time. The state of North Carolina (NC) began linking its reading assessment scales to the Lexile Framework for Reading starting with the first edition of its end-of-grade assessments (introduced in 1993) and continuing with subsequent editions of reading tests up to the current day. Similarly, the state began linking its mathematics assessments to the Quantile Framework starting with the third edition (introduced in 2006) of their mathematics end-of-grade tests and continuing to the present day. In addition, the state began linking its high school content area tests in 2008, providing a basis to extend the longitudinal measurement of reading and mathematics achievement on common scales beyond the elementary and middle school years.

These measurement innovations adopted by North Carolina have several advantages for the interpretation of academic growth. As demonstrated in the examples, the benefits include: a) conjointly interpreting student reading growth in the context of reading materials concomitantly used during K-12 instruction; b) interpreting a reading growth trajectory in light of future (e.g., postsecondary) reading requirements; c) forecasting individual reading comprehension rates relative to both contemporary and future text complexity requirements; and d) creating growth velocity norms for average academic growth in reading and mathematics.

Theories about the developmental velocity of physical attributes can be traced to Aristotle, who observed that height increases fastest when individuals are young; over the intervening centuries, many studies of stature have confirmed and explicated this now well-accepted fact (Tanner, 2010). However, it was not until the emergence of educational and psychological measurement in the early part of the 20th century, that studies of individual academic growth became possible.

A central question in all studies of academic growth is what mathematical function to use for modeling individual growth and the decision necessarily reflects assumptions about learning rate (i.e., growth velocity). In general there have been two traditions to address the question of functional form: a) the empirical tradition of fitting growth curves, which has been traced to Wishart (1938); and, the tradition of selecting a growth function based on an explicit theory of growth rate. In the latter approach, theories of learning rate have been adapted from chemical processes (Robertson, 1909) and the study of mortality (Gompertz, 1825; Winsor, 1932), among others. Whether one works in a purely empirical tradition or is guided by substantive theory, all potential growth models must be subjected to empirical confirmation with longitudinal data.

A key research hypothesis in longitudinal studies relates to whether growth proceeds according to a straight line (with constant velocity) or curvilinear pattern (with variations in velocity and/or acceleration). Some researchers (e.g., Catts, Bridges, Little, & Tomblin, 2008; Guglielmi, 2008; Kieffer, 2012; Sonneschein, Stapleton, & Benson, 2010) found a straight-line growth model adequate for their purposes. However, Lee (2010) reported that American students' growth in reading and mathematics achievement during the K-12 school years is curvilinear, characterized by declining velocity over time. Researchers using more extensive longitudinal research designs have confirmed this finding using the empirical approach (e.g., Schulte, Stevens, Elliott, Tindal, & Nese, 2016; Williamson, 2015) as well as the theory-driven approach to growth (Cameron, Grimm, Steele, Castro-Schilo, & Grissmer, 2015). Moreover, Andrich and Styles (1994) provided psychometric evidence to substantiate intellectual growth spurts in early adolescence.

In America, academic growth predominantly occurs in the context of schooling and growth is presumably influenced by exposure to instructional content. Accordingly, it is illuminating to note that the difficulty of reading materials (Williamson, Koons, Sandvik, & Sanford-Moore, 2012) and the difficulty of mathematical skills and concepts (Sanford-Moore, Williamson, Bickel, Koons, Baker, & Price, 2014) also proceed across Grades K-12 in a curvilinear pattern characterized by positive velocity and deceleration.

The adoption of specific, previously-determined, growth curve results for the subsequent examples carries with it a set of implicit research questions, which I here make explicit.

1. Is NC aggregate reading growth curvilinear during Grades 3-8?
   a. What is the initial status of NC average reading growth in Grades 3-8?
   b. What is the initial velocity of reading growth, Grades 3-8?
   c. What is the acceleration of reading growth, Grades 3-8?
2. Is NC aggregate reading growth curvilinear during Grades 3-11?
   a. What is the initial status of NC average reading growth in Grades 3-11?
   b. What is the initial velocity of reading growth, Grades 3-11?
   c. What is the acceleration of reading growth, Grades 3-11?
3. Is NC aggregate mathematics growth curvilinear during Grades 3-11?
   a. What is the initial status of NC average mathematics growth in Grades 3-11?
   b. What is the initial velocity of mathematics growth, Grades 3-11?
   c. What is the acceleration of mathematics growth, Grades 3-11?

Answers to these three research questions were available from previous research. The aggregate, student growth curves used in the subsequent examples all exhibit a quadratic (curvilinear) functional form with positive initial velocity accompanied by deceleration across time. Specific parameter estimates are provided in the **Examples** section.

The featured examples themselves also have associated research questions, explicitly stated below:

   d. How does average NC student reading growth compare with proposed text-complexity standards widely adopted in the US?

e. How does average NC reading growth align with the text complexity of postsecondary reading materials?
f. Given the average reading growth of NC students, what reading comprehension rates are forecasted relative to K-12 text-complexity standards?
g. Given the average reading growth of NC students, what reading comprehension rates are forecasted relative to postsecondary text complexity?
h. What incremental velocities characterize the historical, average reading growth of NC students from the end of Grade 3 to the end of Grade 11?
i. What incremental velocities characterize the historical, average mathematics growth of NC students from the end of Grade 3 to the end of Grade 11?

Answers to research questions 4 through 9 are presented and explained in the **Examples** section.

Results from the current research study can be used by educators and students alike. To illustrate, consider a student (or group of students) progressing through school. Typically students are assessed on their reading and mathematics achievement annually. As they are assessed, students can compare their individual performance and growth to the average historical growth of previous students. Similarly, teachers can compare their students' individual growth as well as the group's aggregate growth with historical growth. Additionally, students and teachers benefit from the conjoint properties of the measures in the following ways. Based on the first three examples, student reading achievement is readily compared to the text-complexity of both K-12 and postsecondary texts; and, by monitoring students' forecasted comprehension rates, teachers can individualize the match between students and texts that the teacher may assign as students improve their reading abilities. Finally, as students accumulate a history of measured performance, educators can reference the velocity of reading growth and mathematics growth to historical growth rates determined from longitudinal data. These interpretive contexts offer new insights and perspectives that can facilitate instruction as well as program monitoring and evaluation.

DATA

The "data" for the subsequent examples consist of the parameter estimates from multilevel growth models estimated for various panels of students who participated in the North Carolina assessment program. As already mentioned, the parameter estimates are adopted from previous work (e.g., MetaMetrics, 2011; Williamson, 2014). The original student-level data, which were the basis for the fitted growth models, consisted of Lexile or Quantile measures that were obtained through linking the North Carolina assessment scales to the Lexile Framework and the Quantile Framework.

North Carolina assessments have well-documented technical characteristics (Bazemore & Van Dyk, 2004; North Carolina Department of Public Instruction, 2009; Sanford, 1996) and have successfully satisfied the requirements of the Elementary and Secondary Education Act (No Child Left Behind, 2002). In general, panels were comprised of longitudinal data spanning Grades 3-8, where the assessments were administered once a year at the end of each grade. For the velocity norms examples, additional waves of data were employed through Grade 11.

EXAMPLES

The first three examples are based on a multilevel unconditional quadratic growth model, which was fit to the longitudinal data from a North Carolina panel spanning Grades 3-8 in 2000-2005. Based on data from every student who had at least one reading measure during the six-year time frame, this curve provides a historical summary of average student reading growth for 98,515 students, representing 92.8% of the Spring 2005 eighth-grade cohort that defined the panel. The estimates of the intercept, velocity and curvature parameters for the average reading growth curve were 670.2L, 119.6L/year and -6.1L/year^2, respectively. In Figure 1, I provide a visual summary of the statewide average reading growth curve, the corresponding velocity curve and the acceleration curve for reading growth based on the multilevel analysis.

Note in Figure 1 the horizontal scale is graduated by grade, where the coding refers to the end of the respective year. So for example, the numeral 3 on the grade scale refers to the end of Grade 3. Furthermore, the time scale for the growth model was centered at the end of Grade 3; thus the velocity parameter estimate refers to the velocity at the end of Grade 3. The vertical axis is denominated in Lexile scale units. The meaning of the Lexile scale unit was described earlier.

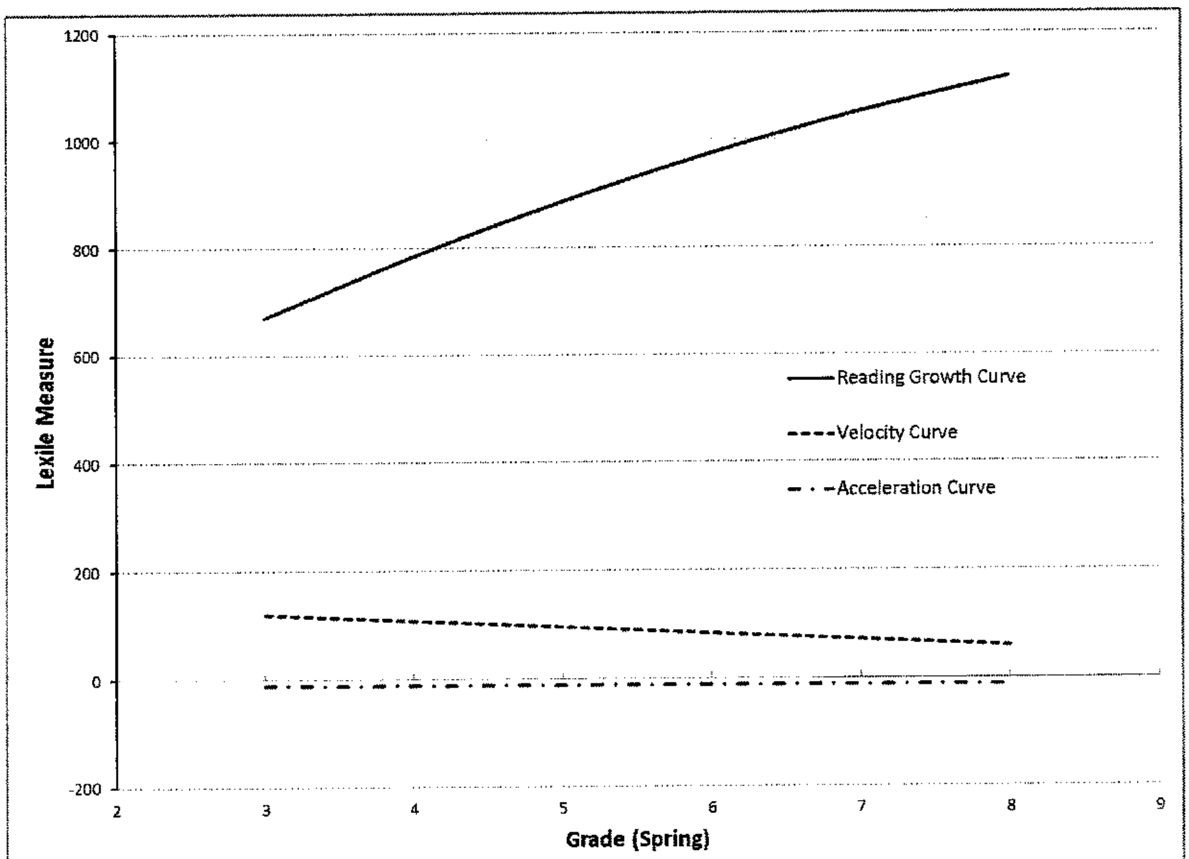


Figure 1. Average reading growth, velocity and acceleration curves for the 2000-2005 North Carolina panel (n = 98,515). The vertical axis graduates the growth curve in Lexile scale units, the velocity curve in Lexile units/year and the acceleration curve in Lexile units/year^2.

In Figure 1, notice that the growth curve begins around 670L at the end of Grade 3 and then rises quickly during the early grades; however, the curve decelerates across the Grade 3-8 time frame. The velocity curve in Figure 1 displays the fact that velocity is linearly related to time when the growth curve has a quadratic functional form. In this particular example, the velocity curve shows that velocity declines from approximately 120L/year at the end of Grade 3 to approximately 60L/year at the end of Grade 8. The slope of the velocity curve (-12.2L) is equal to the acceleration rate of the growth curve. Because the slope of the velocity curve is negative, growth is decelerating during the time frame. For a quadratic growth curve, the acceleration rate is manifested through the curvature parameter. Acceleration is constant and equal to twice the curvature parameter (i.e., -6.1L in this case). This is consistent with the constant negative elevation displayed for the acceleration curve in Figure 1. The growth, velocity and acceleration curves are relatively simple for a quadratic growth model; nevertheless, it is useful to display them in the fashion of Figure 1 because it provides a convenient and readily understandable summary of the key features of growth.

Student Growth in Reading versus the Common Core State Standards

The Common Core State Standards (CCSS) Initiative [National Governors Association Center for Best Practices (NGA Center) & the Council of Chief State School Officers (CCSSO), 2010] established quantitative text complexity standards for specific grade bands in the public schools. The standards are expressed as text complexity ranges denominated in terms of six text complexity metrics in common use in the United States. One of those metrics is the Lexile measure, which makes it possible to compare the text complexity standards of the CCSS to actual student reading achievement measured with the Lexile Framework. The CCSS College and Career Readiness Anchor Standards for Reading require that by the end of specific grades that demark the end of the CCSS grade bands (i.e., grades 3, 5, 8, 10, and 12), students must "read and comprehend literature, including stories, dramas, and poetry/poems, at the high end of the ... text complexity band independently and proficiently." (pp. 12, 37, 38) The upper end of the text complexity range for the Grade 11-12 grade band was labeled "CCR" by the CCSS to connote college and career readiness.

In Figure 2, I depict the 2000-2005 NC growth curve and the CCSS text complexity ranges for Grades 3, 5 and 8. The lower and upper boundaries of the CCSS text complexity ranges at the critical grades are represented by dots, which are connected by dashed lines to provide a visual reference as context for the growth trajectory. If student growth were commensurate with the CCSS text complexity standards, then one would expect to see the growth curve traversing a path that lies within the text complexity boundaries, rising near the upper end of the range by the specified grades, which denote the end of each grade band. In fact, the NC average growth curve approximates this behavior. Its intercept appears to be slightly above the mid-point of the text complexity range for the Grade 2-3 grade band and the curve rises nearer the upper boundary by the end of the Grade 6-8 band. If one imagines that the average growth curve is in fact the growth curve for an individual student, then it would seem that the student's growth is reasonably well aligned with the standards. Is it good enough? What does the growth curve imply about the actual reading experience that the student would have relative to the CCSS upper boundaries as he or she grows? I will come back to these questions in a subsequent example. First, I wish to introduce the idea that there are additional text requirements that characterize reading experiences which students may encounter after they graduate high school. Consequently, student growth during the K-12 years has implications for reading experiences that students will encounter later.

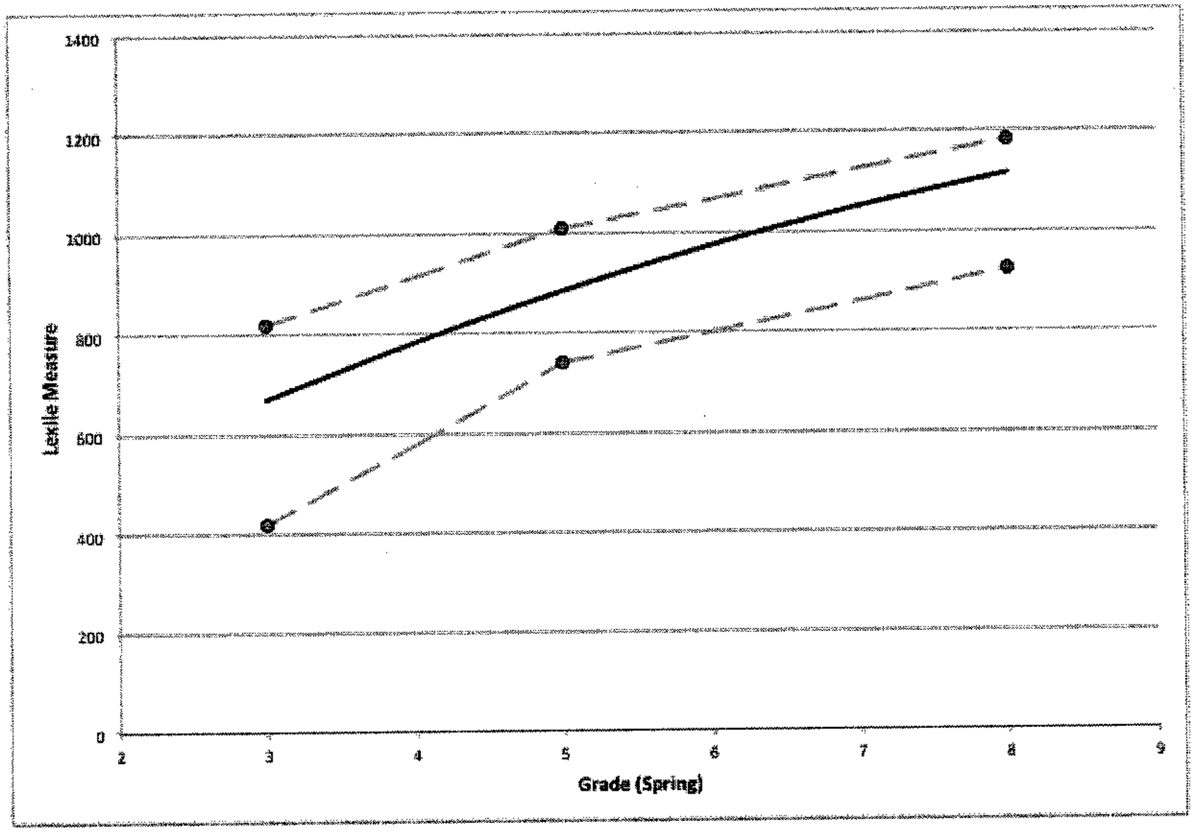


Figure 2. Reading growth relative to the Common Core State Standards (CCSS) text complexity ranges. The growth curve is the 2000-2005 North Carolina average growth curve (n = 98,515). The dots represent the CCSS Lexile range boundaries at grades 3, 5 and 8. The dashed lines provide a visual reference for the growth trajectory as it traverses the CCSS grade bands.

Reading Growth in Relation to Postsecondary Text Complexity

The objective of this example is to illustrate average student reading growth in relation to the text complexity of reading materials that students may encounter beyond high school. To accomplish this objective, I combine knowledge about the functional form of reading growth during K–12 with text complexity measures of postsecondary reading materials to construct an empirically-based model of student growth toward postsecondary performance aspirations. Such a model can be a useful first step toward understanding the possible long-term implications of growth.

Williamson (2008) elaborated a continuum of text complexity for reading materials associated with typical postsecondary endeavors (e.g., postsecondary education, the military, the workplace, citizenship). This work demonstrated substantial differences between the materials that high school students are expected to read and the materials they may encounter after high school. The latter reflect a substantially higher text demand, or correspondingly, require a higher reading ability from students in their postsecondary lives. The median Lexile measures for five postsecondary text collections summarized by Williamson are: 1395L (university), 1295L (community college), 1260L (workplace), 1230L (citizenship) and 1180L (military).

Once again, I use the statewide average reading growth curve of the 2000-2005 NC panel. Using the fixed effects estimates from the multilevel analysis, the average reading growth curve is expressed as a mathematical equation: $E(L|T) = 670.2 + 119.6\,T - 6.1\,T^2$. This equation quantifies the estimated average achievement in any grade.

The 2000-2005 panel is comprised of 98,515 North Carolina public school students who were third graders in the spring of 1999-2000 and who progressed to the end of eighth grade in the spring of 2004-05. These students progressed from Grade 3 to 8 without repeating a grade and were included in the analysis if they had at least one reading measure during the six-year time frame. Consequently, the average growth curve of these students should provide a good illustration of typical student growth toward postsecondary expectations. All of the relevant information about the growth curve is summarized in the three parameter estimates: 670.2L (initial status—end of third grade), 119.6L (initial velocity), and -6.1L (curvature).

Data were not available prior to the end of Grade 3 or after the end of Grade 8. However, with some caution, the quadratic equation that characterizes the curve through the range of observed data can be used to estimate average performance before Grade 3 and after Grade 8. Simply evaluating the growth curve at the other time points suffices.

When extrapolating, it is important to use caution for at least two reasons. First, there are no actual data to check the assumption that growth from Grades K–2 and Grades 9–12 can be described by the same quadratic equation that describes growth from Grades 3–8. Second, the nature of a quadratic polynomial is that it has a maximum point or a minimum point, after which the curve reverses direction. When the curve is concave to the time axis (as is the case for the NC average growth curve), there will be a maximum point after which the curve turns downward. It is implausible that future performance will decline back to the third-grade level and below; this would be inconsistent with normal developmental growth.

There are (at least) three ways to address these concerns. The easiest way is to analytically check the quadratic equation to determine when the maximum point occurs. If it occurs outside the range of time to which one wishes to generalize, then there is less reason to worry that the depiction of growth may be inappropriate. As it turned out, the maximum for the 2000-2005 North Carolina growth curve occurred at Grade 12.9, almost a year beyond the end of twelfth grade, which is the last occasion for which average student achievement was projected.

A more direct way to avoid non-developmental behavior in a growth model is to adopt a different mathematical model for growth—e.g., one that cannot display a reversal in direction. A linear model with a transformed time scale is one possibility, such as: $r(t) = a + b \ln t$, which increases monotonically without bound. Another alternative is to select a model that is nonlinear in the parameters, such as the negative exponential: $r(t) = a - (a - b)e^{-ct}$, which increases monotonically to an asymptote. There are many possibilities (e.g., see Singer & Willett, 2003; or, Goldstein, 1979 for a variety of specific choices). Alternative models carry with them alternate interpretations of growth, may be more complex mathematically, and may require additional data to obtain satisfactory fit. Ultimately, the choice of most appropriate model is based on multiple considerations—e.g., substantive theory, available data, empirical fit, parsimony, and perhaps other requirements.

The third way to address the risks of extrapolation is to strategically collect more data to fill in the missing time points with student achievement information. Unfortunately, this is harder than it sounds for a variety of reasons, including the costs of collecting the information and the challenge of measuring the same construct over longer and longer periods of time. For the present example, the results may be regarded as provisional, bearing in mind that extrapolations to lower and higher grades may need to be revised based on future information.

With those cautions in mind, Figure 3 shows the results of combining the information from the text analyses and the information from the NC reading growth curve. There are several important things to notice about Figure 3.

Once again, the horizontal scale represents end-of-grade in school. On this scale, zero stands for the end of the kindergarten year. Subsequent Grades (1–12) are denoted as usual. Then the numerals 13 through 14 are used to denote the next two years of postsecondary experience. The vertical scale displays the Lexile measure, which is used to quantify *both* the students' average reading achievement and the median text difficulty of each text collection.

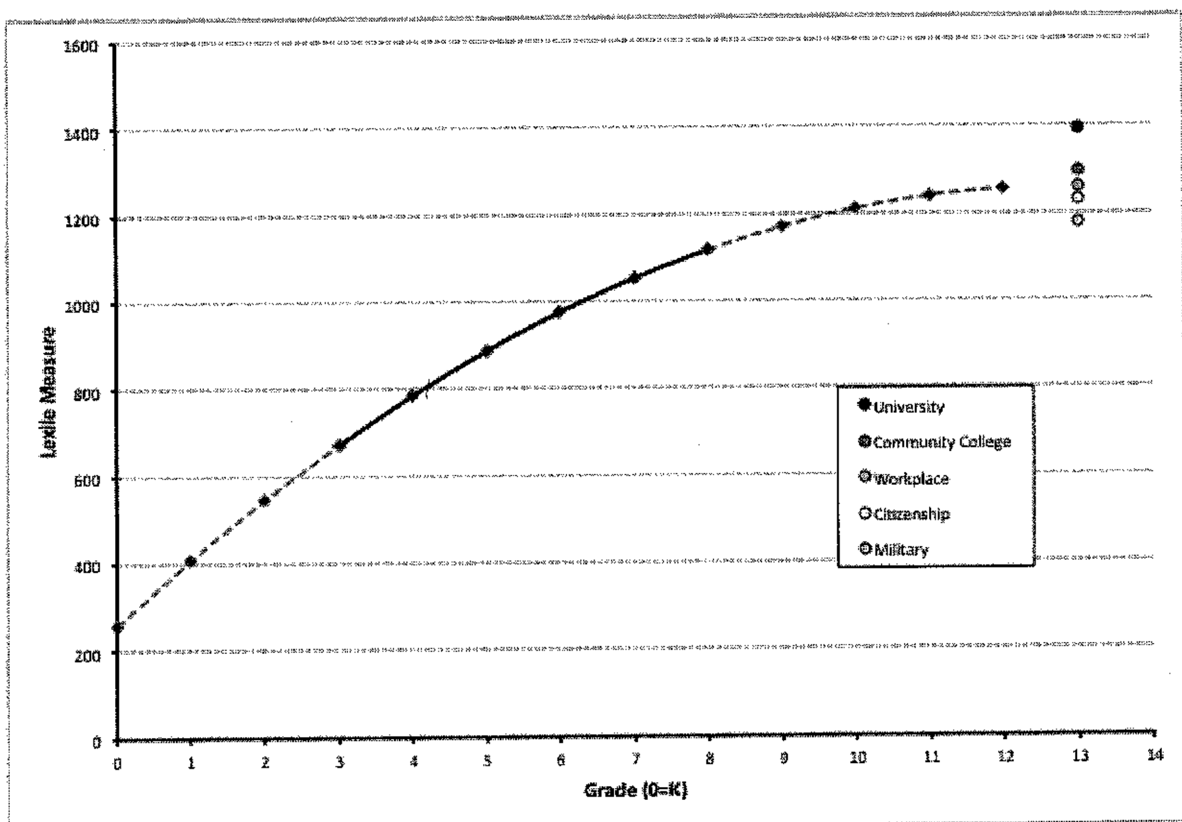


Figure 3. Average student growth in relation to postsecondary text complexity. The solid curve represents the 2000-2005 North Carolina average growth curve (n = 98,515). The dashed portions of the curve are mathematical extrapolations based on the quadratic equation for the average growth curve. The shaded dots in the upper right represent the median text complexities for the respective text collections listed in the legend (Williamson, 2008).

In the graph, diamonds are used to indicate the estimated average reading ability of students at the end of each grade. The estimates for Grades 3–8 are connected by the solid empirical growth curve to represent the fact that they are based on the available data. The estimates for Grades K–2 and 9–12 are connected with dashed curves to represent the fact that they are theoretical extrapolations determined analytically from the quadratic equation for the empirical growth curve. As such, the dashed portions of the curve are only reasonable guesses based on the observed data, subject to future revision based on more complete longitudinal records[1]. The farther one goes from the observed data (Grades 3–8), the more one has to bear in mind the provisional nature of the projections. Finally, in the figure, the median text difficulties of the postsecondary text collections are arrayed vertically at Grade 13 to indicate that students face these expectations in the year following their exit from Grade 12.

The primary feature of the chart is the alignment of the projected twelfth-grade reader measure in conjunction with the postsecondary text measures.[2] It appears that the average growth trajectory of these students, if unaltered, will carry them to a reading level (1256L) that lies near the median text requirements of the workplace (1260L). Students with higher postsecondary aspirations (e.g., the community college, the university) need to be on a higher trajectory that tracks above the average growth curve depicted in the figure.

One must remember, however, that individual growth is variable and that students vary in their individual parameters of growth. That is, students have different beginning points, different initial velocities and different degrees of deceleration. Each of these features of growth results in a different individual trajectory, which may differ from the average growth trajectory. Thus, there are many possible ways to reach a given end point. For example, one student might begin at a higher level and exhibit modest but steady growth with little deceleration over time. Another might start out lower in reading ability but progress very rapidly with some deceleration over time. Both students might reach the same twelfth-grade reading ability through different individual growth curves. Williamson, Fitzgerald and Stenner (2014) discussed alternate growth trajectories in terms of the pedagogical and educational policy implications of directly targeting key features of growth (status, velocity and acceleration). For example, early-intervention reading programs can successfully influence initial reading status; increased deliberate practice might impact velocity; and, systematic exposure to summer school could be a viable strategy to moderate deceleration.

Forecasted Comprehension Rates Based on a Growth Curve

For this example, I return to the question of what kind of reading experience students are likely to have with particular levels of text complexity—e.g., the CCSS text standards or postsecondary text requirements. Again using the 2000-2005 NC average growth curve and supposing that the curve might describe the trajectory of a particular individual, it is possible to estimate the

---

[1] Although there are assessments of U.S. students prior to the end of Grade 3 [e.g., the Early Childhood Longitudinal Study (ECLS)] and after the end of Grade 8 [e.g., National Education Longitudinal Studies (NELS)], they are generally available only for samples of individuals and reading measurements from these studies have not yet been brought onto a common scale.

[2] The median difficulty (1130L) of texts used near the end of high school (i.e., grades 11 and 12) is not shown in the figure, because it does not represent a postsecondary aspiration. High school texts are significantly easier to read on average than are citizenship materials, workplace materials, community college texts or university texts (Williamson, 2008).

individual's comprehension rate relative to texts the individual may encounter. To do this, it is necessary to have a general idea of how the Lexile Framework for Reading can be used to forecast reading comprehension given a reader of a particular reading ability and a text of a particular difficulty. Stenner, H. Burdick, Sanford and Burdick (2007) described the approach. In essence, one forecasts the comprehension rate by using the Rasch model equation, which expresses the reading outcome (comprehension) as a function of the exponentiated difference between the reader's ability and the text's difficulty. The Lexile Framework is designed so that an exact match between reader and text (i.e., reader ability equals text complexity, and so the difference between the two is zero) results in a comprehension rate of 75%. A comprehension rate of approximately 75% seems to be associated with successful reading experiences; whereas, a comprehension rate of 50% or lower results in frustration for the reader (Scholastic, Inc., 2007). MetaMetrics typically advises educators to choose texts that lie in a proximal zone ranging from 100L below the reader's ability to 50L above it when using the Lexile Framework to match readers with texts of appropriate difficulty. This proximal zone corresponds to comprehension rates that range from approximately 70% to 80%.

Consider a reader whose growth curve is equal to the 2000-2005 NC average growth curve. What happens when such a student reads a book that has text complexity equal to the upper end of the CCSS text complexity ranges? What happens when such a student reads a book that has a text complexity equal to the typical text complexity of postsecondary reading materials (1300L)? In the first scenario, the CCSS text demand changes from grade to grade as the student's reading ability (reflected by the growth curve) changes. In the second scenario, there is a fixed future, postsecondary target toward which the student is progressing. I address both situations in Table 1.

For each of the Grades 3, 5, 8, 10 and 12 (i.e., the transition grades between the CCSS grade bands), I tabulate in the first four rows of Table 1: a) the average student performance (estimated from the growth curve); b) the CCSS text complexity upper bound; c) the difference between the two; and, d) the resulting forecasted comprehension rate at the end of the grade.

TABLE 1
Forecasted Comprehension Rates Implied by the 2000-2005 North Carolina Average Reading Growth Curve Relative to a) the Common Core State Standards (CCSS) Grade Bands and Text Complexity Ranges and b) Median Postsecondary Text Complexity

	Grade				
	3	5	8	10	12
	Average Student Achievement Summarized by the Longitudinal Growth Curve				
NC (2000-2005)	670L	885L	1117L	1211L	1256L[a]
	CCSS Text Complexity Requirements				
CCSS	820L	1010L	1185L	1335L	1385L
Difference	-150L	-125L	-68L	-124L	-129L
Forecasted Comprehension	61%	63%	69%	63%	63%
	Median Postsecondary Text Complexity				
Postsecondary Texts (Median)	1300L	1300L	1300L	1300L	1300L
Difference	-630L	-415L	-183L	-89L	-44L
Forecasted Comprehension	15%	32%	57%	67%	71%

Note. A multilevel growth analysis (n = 98,515) was used to estimate the average reading achievement at the end of each respective CCSS grade band. The upper boundaries of the CCSS Lexile ranges associated with the respective grades are given in the row labeled *CCSS*. A reader who is well matched with a text at his or her Lexile measure is forecasted to have a 75% comprehension rate.
[a] The empirical data spanned Grades 3-11. The estimated average achievement at the end of Grade 12 is extrapolated from the growth curve.

In general, we expect a reader to have 75% comprehension of a well-targeted text (i.e., a text at the student's reading level). Because the CCSS text complexity standards represent a series of increasing aspirational goals, we can ask how well the average reader in our example might do relative to the changing text complexity standards as he or she grows. That is, what would be the student's comprehension rate when confronted with a text with the higher text complexity prescribed by the CCSS? Table 1 provides the answer. The forecasted comprehension rates rise from 61% (in Grade 3) to 69% (in Grade 8) during the empirical time frame for the panel. However the comprehension rate is forecasted to drop back to 63% during the high school years, if the individual continues on the same trajectory traversed during Grades 3-8. Although, the CCSS grade bands and text complexity ranges are designed to provide flexibility to accommodate readers with a wide range of abilities, this example suggests that the average student in the 2000-2005 panel may experience some challenge relative to texts at the upper ends of the CCSS text complexity ranges (because all of the forecasted comprehension rates are less than 75%).

In the bottom half of Table 1, we can see that the hypothetical average student experiences increasing rates of comprehension while growing toward the fixed postsecondary text complexity target. Although forecasted comprehension of the median (1300L) postsecondary text is understandably low (15%) when the student reads as a typical third grader, the forecasted comprehension rate steadily climbs to 71% by the end of Grade 12, based on the estimated average reading growth curve.

A nice feature of this analysis is that it can be replicated with any estimated growth curve, whether for an individual or for a group (e.g., an average growth curve). One only needs estimates of reading ability at each desired point in time, which can easily be determined from the mathematical equation for growth.

Incremental Velocity Norms for Average Reading and Mathematics Growth

Replicating or exceeding some specified previous student achievement level was the basis for educational expectations throughout most of the 20th century. Similarly, replicating or exceeding previous growth *rates* eventually emerged as a basis for student growth standards (North Carolina Department of Public Instruction, 1996). Even so, the best implementation of educational growth standards to date has been based on year-to-year gains, without the benefit of an underlying longitudinal growth curve. Growth velocity norms did not emerge even for height or weight until the work of Tanner, Whitehouse and Takaishi (1966) in the United Kingdom and later in the United States (Roche & Himes, 1980; Baumgartner, Roche & Himes, 1986). In this next example, I use two parametric models for growth (one for reading, one for mathematics) derived from NC longitudinal data (MetaMetrics, 2011). I shall use the historical results to create incremental growth velocity norms for average reading and mathematics growth. The approach yields not only estimates of year-to-year gain, but estimates of growth between any two points within the design time frame running from the end of Grade 3 to the end of Grade 11.

The starting point is the realization that an historical aggregate growth curve provides a long-term summary of observed growth for a group of students. As such, it may be regarded as a norm for growth. If this norm were treated as a growth expectation for future panels of students, the implicit policy goal would be that future students *should* grow in a manner that is similar to previous historical growth. When regarded as a set of expectations for future growth, the growth curve represents a *growth standard*. Perhaps the easiest way to operationalize such a growth standard is by generating incremental growth velocity estimates from the average growth curve. It is relatively easy to do this. One needs only the parameter estimates for the average growth curve. In this case, there are two parametric models—one based on a ten-wave analysis of reading growth and the other based on a nine-wave analysis of mathematics growth. These two growth curves are salient because they each span Grades 3-11, the grades during which accountability assessments are most often implemented in the United States and the grades most often the focus of state accountability systems.

The estimated average reading growth curve is a function of time, $r(T) = 663.8 + 148.0\,T - 8.7\,T^2$. I can use it to estimate the expected amount of growth from one time point to another. For purposes of the example, let us interpret the time scale in terms of *grade in school* with the understanding that the gains so calculated will represent the growth from one spring to another because testing took place at the end of the school year.

When I calculate the gain between adjacent grades, I have calculated the amount of change per unit of time—i.e., the incremental velocity. When I calculate the gain between any two grades more than one year apart, it produces an incremental estimate of the amount of growth that took place between those two grades.

In Table 2, I have tabulated the values of $r(k) - r(j)$ for all pairs of grades (j,k) such that $k > j$ where $j = 3, 4, \ldots, 10$ and $k = 4, 5, \ldots, 11$. The resulting values are displayed in matrix form. Quantities along the diagonal represent the expected gain for each year-to-year transition: Grade 3 to Grade 4, Grade 4 to Grade 5, and so on. These are the incremental yearly, spring-to-spring growth velocity norms based on a population of 101,610 students. The off-diagonal elements of the table display the amount of growth between every other possible pair of grades. This information is useful because it captures longer-term growth expectations, spanning multiple grades.

To illustrate the interpretation of growth using Table 2, first consider the annual yearly growth expectations displayed along the diagonal. A fourth-grade teacher might reference the entry at the intersection of the row for Grade 3 and the column for Grade 4. The entry conveys the expectation for average reading growth between the end of Grade 3 and the end of Grade 4—namely during the fourth grade year. It is 139L. Similarly, the fifth-grade teacher would reference the entry at the intersection of the row for Grade 4 and the column for Grade 5 and learn that the average growth expected of fifth graders is 122L. The principal of a middle school serving students in Grades 6-8 would be interested in the total gain expected between the end of the fifth grade and the end of the eighth grade. Referring to the intersection of the row for Grade 5 and the column for Grade 8, the principal learns that the expectation for average reading growth for students who spend all three years at the middle school is 260L.

Similarly, the average mathematics growth curve can be expressed as: $m(T) = 586.0 + 100.6\ T - 3.0\ T^2$. Having evaluated the average mathematics growth curve at all grade-pairs, I displayed the results in Table 3. The interpretation of average mathematics growth in Table 3 follows in the same manner as for reading growth (Table 2).

In both Table 2 and Table 3 it is obvious that historical growth is typically greater in earlier grades and tapers off as grade increases. This is apparent as one scans along the diagonal from upper left to lower right. This pattern reflects the deceleration of growth and quantifies it in practical terms for educators. However, the off-diagonal entries in the table reinforce the realization that long-term growth is the result of a cumulative growth process that endures across the developmental life-span.

TABLE 2

Incremental Velocity Norms for Average Reading Growth Denominated in Lexile Scale Units

	Student Achievement Estimated from the Average Reading Growth Curve								
	664L	803L	925L	1029L	1116L	1185L	1237L	1271L	1288L
End of Grade	3	4	5	6	7	8	9	10	11
3		139L	261L	365L	452L	521L	573L	607L	624L
4			122L	226L	313L	382L	434L	468L	485L
5				104L	191L	260L	312L	346L	363L
6					87L	156L	208L	242L	259L
7						69L	121L	155L	172L
8							52L	86L	103L
9								34L	51L
10									17L

Note. The table is based on an average reading growth curve (ten waves of measurement) for North Carolina students (n = 101,610), spanning grades 3-11 during the years 2002-2010. The fitted model is summarized by the equation: $E(L|T) = 663.8 + 148.0\,T - 8.7\,T^2$ where the time scale is centered at Grade 3 (i.e., T = Grade -3). Velocity increments for adjacent grades (i.e., spring-to-spring gains) are shown in the shaded diagonal.

TABLE 3

Incremental Velocity Norms for Average Mathematics Growth Denominated in Quantile Scale Units

	Student Achievement Estimated from the Average Mathematics Growth Curve								
	586Q	684Q	775Q	861Q	941Q	1014Q	1082Q	1144Q	1200Q
End of Grade	3	4	5	6	7	8	9	10	11
3		98Q	189Q	275Q	355Q	428Q	496Q	558Q	614Q
4			92Q	177Q	257Q	331Q	399Q	460Q	516Q
5				86Q	165Q	239Q	307Q	369Q	425Q
6					80Q	153Q	221Q	283Q	339Q
7						74Q	142Q	203Q	259Q
8							68Q	130Q	185Q
9								62Q	118Q
10									56Q

Note. The table is based on an average mathematics growth curve (nine waves of measurement) for North Carolina students (n = 101,650), spanning grades 3-11 during the years 2002-2010. The fitted model is summarized by the equation: $E(Q|T) = 586.0 + 100.6\,T - 3.0\,T^2$ where the time scale is centered at Grade 3 (i.e., T = Grade -3). Velocity increments for adjacent grades (i.e., spring-to-spring gains) are shown in the shaded diagonal.

CONCLUSION

In this paper, I have proposed novel interpretations of student academic growth based on conjoint measurement and longitudinal data analyses. In three examples, I illustrated how to interpret student reading achievement and growth in light of the text complexity associated with reading materials that students may encounter during schooling or in the postsecondary world. In the final example, I implemented a strategy to create incremental velocity norms for average academic growth and provided examples of velocity norms for reading growth and for mathematics growth, each based on over 100,000 students.

The first three examples highlighted the power of conjoint measurement when combined with the longitudinal perspective of student growth curves. We first saw how to compare student growth to changing text complexity requirements such as those expressed in the CCSS. Then, we saw a student growth curve juxtaposed with postsecondary text requirements and I suggested that alignment between the two is desirable. Next, we saw how the first two examples lead us to forecasted comprehension rates for readers who are themselves growing in their reading ability. Although these three examples featured reading ability relative to text complexity requirements, it is possible to provide similar examples for growth in mathematics ability relative to the complexity of mathematical skills and concepts.

Finally, we saw how parametric growth curves can strengthen the basis for setting growth standards based on longitudinal panel data, rather than the usual practice of setting year-to-year growth standards based on non-developmental (e.g., status projection) or short-term growth (e.g., gain score) formulations. Incremental velocity norms such as those presented here are an indispensable complement to traditional cross-sectional norms for interpreting student achievement because velocity norms a) base year-to-year gains on a longitudinal growth curve and b) make it possible to construct expectations of growth between any pair of grades.

Although, the growth velocity norms provided in this paper are for statewide average growth, they are easily extended to sub-populations. To briefly elaborate, one possibility for expanding growth standards is to disaggregate an historical average growth curve into multiple growth curves conditioned on initial status. For example, by grouping students into deciles based on initial performance, average growth curves can be estimated for each of the ten deciles. Once decile growth curves have been determined, incremental velocity norms can be established for each decile group simply by replicating Table 2 (or 3) for each group's aggregate growth curve. Conditioning growth standards on initial performance is a feature that has been desired in some accountability systems.

Furthermore, if common scales were universally used for educational constructs and longitudinal data were routinely collected and analyzed, then growth velocity standards could have even greater generalizability. Individual state norms, national norms, perhaps even international norms for academic growth velocity would become possibilities.

In the present study, the measurement of growth was constrained to Grades 3-11. An important policy challenge for educators is extending the measurement of reading and mathematics abilities beyond traditionally assessed grades. This entails devising ways to measure the same constructs over longer portions of the lifespan using a common scale so that we can accurately chart the academic growth of students from emerging readers and mathematicians, throughout formal instruction and schooling, and into adulthood. Our current educational measurement capabilities are focused on a fraction of the developmental lifespan and miss much of the growth that we might otherwise observe. Notably, we miss critical transitions such as the entry into K-12 education and transitions into various postsecondary endeavors (e.g., higher education, the workplace). Similarly, we know little about the effects of aging on academic growth trajectories because we have not fully developed our capacity to measure reading and mathematics abilities across the life course using a common scale.

Improving the measurement and study of academic growth is more than a research agenda or a matter for the research and measurement community. Educational leaders and policy makers should commit resources to support the intellectual endeavor because it enriches and sustains the educational enterprise, possibly with residual benefits for long-term human intellectual capacity and quality of life. The advantages become palpable when conjoint measurement is brought to bear as a means to link assessment with instruction. As we have seen in these examples, when conjoint measurement is combined with longitudinal analyses of academic growth, unique insights and perspectives emerge to inform educational practice.

REFERENCES

Andrich, D., & Styles, I. (1994). Psychometric evidence of intellectual growth spurts in early adolescence. *Journal of Early Adolescence, 14*(3), 328-344.

Baumgartner, F. N., Roche, A. G., & Himes, J. H. (1986). Incremental growth tables: Supplementary to previously published charts. *The American Journal of Clinical Nutrition, 43*, 711-722.

Bazemore, M., & Van Dyk, P. B. (2004). *North Carolina reading comprehension tests: Technical report* (Ed.2). ("Citable Draft"). Raleigh, NC: North Carolina Department of Public Instruction.

Bormuth, J. R. (1969). *Development of readability analyses.* Final Report, Project No. 7-0052, Contract No. OEG-3-7-070052-0326, Office of Education, Bureau of Research, U.S. Department of Health, Education and Welfare.

Bruning, R. (1985). Review of Degrees of Reading Power. *The Ninth Mental Measurements Yearbook, Vol. I*, 443-444.

Cameron, C. E., Grimm, K. J., Steele, J. S., Castro-Schilo, L., & Grissmer, D. W. (2015). Nonlinear Gompertz curve models of achievement gaps in mathematics and reading. *Journal of Educational Psychology, 107*(3), 789-804.

Catts, H. W., Bridges, M. S., Little, T. D., & Tomblin, J. B. (2008). Reading achievement growth in children with language impairments. *Journal of Speech, Language, and Hearing Research, 51*, 1569-1579.

Goldstein, H. (1979). *The design and analysis of longitudinal studies.* New York: Academic Press.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London, 115*, 513-583.

Guglielmi, R. S. (2008). Native language proficiency, English literacy, academic achievement, and occupational attainment in limited-English-proficient students: A latent growth modeling perspective. *Journal of Educational Psychology, 100*, 322-342.

Kieffer, M. J. (2012). Early oral language and later reading development in Spanish-speaking English language learners: Evidence from a nine-year longitudinal study. *Journal of Applied Developmental Psychology, 33*, 146-157.

Koslin, B. L., Zeno, S., & Koslin, S. (1987). *The DRP: An effectiveness measure in reading.* New York: The College Entrance Examination Board.

Lee, J. (2010). Tripartite growth trajectories of reading and math achievement: Tracking national academic progress at primary, middle, and high school levels. *American Educational Research Journal, 47*(4), 800-832.

Legendre, A. M. (1805). *Nouvelles méthodes pour la détermination des orbites des comètes.* Paris: Courcier.

MetaMetrics. (2009). *The Quantile Framework for Mathematics: Linking assessment with mathematics instruction.* Durham, NC: Author. Retrieved from https://s3.amazonaws.com/quantile-resources/resources/downloads/static/BriefingDocument.pdf

MetaMetrics. (2011). *NC EOC Feasibility Studies Research & Development Report.* Durham, NC: Author.

National Governors Association Center for Best Practices (NGA Center) & the Council of Chief State School Officers (CCSSO). (2010). *Common Core State Standards for English Language Arts and Literacy in*

History/Social Studies, Science and Technical Subjects. Washington, DC: Author. Retrieved from http://www.corestandards.org/wp-content/uploads/ELA_Standards.pdf

Nelson, J., Perfetti, C., Liben, D., & Liben, M. (2012). *Measures of text difficulty: Testing their predictive value for grade levels and student performance.* Washington, DC: Council of Chief State School Officers (CCSSO). Retrieved on June 11, 2012 from: http://www.ccsso.org/Documents/2012/Measures%20ofText%20Difficulty_final.2012.pdf

No Child Left Behind Act of 2001, PL 107-110 (2002).

North Carolina Department of Public Instruction. (1996, September). Setting annual growth standards: "The formula" (Accountability Brief, Vol. 1, No. 1). Raleigh, NC: Author.

North Carolina Department of Public Instruction. (2009). *North Carolina Reading Comprehension Tests: Technical report* (Ed. 3). Raleigh, NC: Author. Retrieved from http://www.ncpublicschools.org/docs/accountability/testing/reports/eogreadingtechman3.pdf

Robertson, T. B. (1909). A biochemical conception of the phenomena of memory and sensation. *The Monist,19*(3), 367-386.

Roche, A. F., & Himes, J. H. (1980). Incremental growth charts. *The American Journal of Clinical Nutrition, 33*, 2041-2052.

Sanford, E. E. (1996). *North Carolina end-of-grade tests technical report # 1: Reading comprehension, mathematics.* Raleigh, NC: Department of Public Instruction.

Sanford-Moore, E. E., Williamson, G. L., Bickel, L., Koons, H., Baker, R. F., & Price, R. (2014). *A quantitative task continuum for K-12 mathematics* (MetaMetrics Research Brief). Retrieved from https://s3.amazonaws.com/quantile-resources/resources/downloads/static/Quantile+Framework+Math+Continuum+Research+Brief+2014.pdf

Scholastic, Inc. (2007). *Scholastic Reading Inventory™ technical guide.* New York: Author.

Schulte, A. C., Stevens, J. J., Elliott, S. N., Tindal, G., & Nese, J. F. T. (2016, January 25). Achievement gaps for students with disabilities: Stable, widening, or narrowing on a state- wide reading comprehension test? *Journal of Educational Psychology.* Advance online publication. http://dx.doi.org/10.1037/edu0000107

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York: Oxford University Press.

Sonneschein, S., Stapleton, L. M., & Benson, A. (2010). The relation between the type and amount of instruction and growth in children's reading competencies. *American Educational Research Journal, 47,* 358-389.

Stenner, A. J., Burdick, H., Sanford, E., & Burdick, D. S. (2007). *The Lexile Framework for Reading technical report.* Durham, NC: MetaMetrics, Inc. Retrieved from http://www.lexile.com/research/9/

Stenner, A. J., Smith, D. R., Horabin, I., & Smith, M. (1987). Fit of the Lexile theory to sequenced units from eleven basal series. Durham, NC: MetaMetrics. Retrieved from https://lexile-website-media-2011091601.s3.amazonaws.com/m/resources/materials/Stenner_Smith_Horabin_Smith-_Fit_of_the_Lexile_Theory_to....pdf

Stenner, A. J., & Smith, M. (1982). Testing construct theories. *Perceptual and Motor Skills, 55,* 415-426.

Stenner, A. J., Smith, M., & Burdick, D. S. (1983). Toward a theory of construct definition. *Journal of Educational Measurement, 20,* 305-316.

Tanner, J. M. (2010). *A history of the study of human growth.* Cambridge, UK: Cambridge University Press.

Tanner, J. M., Whitehouse, R. H., & Takaishi, M. (1966). Standards from birth to maturity for height, weight, height velocity and weight velocity: British children, 1965. *Archives of Disease in Childhood, 41*, 454-471, 613-635.

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

Williamson, G. L. (2014). *Measuring and modeling individual academic growth: Methodological foundations for educational applications in the 21st century.* Manuscript in preparation.

Williamson, G. L. (2015). Measuring academic growth contextualizes text complexity. *Pensamiento Educativo: Revista de Investigación Educacional Latinoamericana, 52*(2), 98-118.

Williamson, G. L., Fitzgerald, J., & Stenner, A. J. (2014). Student reading growth illuminates the Common Core text-complexity standard: Raising both bars. *Elementary School Journal, 115*(2), 230-254.

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.

Winsor, C. P. (1932). The Gompertz curve as a growth curve. *Proceedings of the National Academy of Sciences, 18*(1), 1-8.

Wishart, J. (1938). Growth rate determinations in nutrition studies with the bacon pig, and their analysis. *Biometrika, 30,* 16-28.

MetaMetrics.

METAMETRICS RESEARCH BRIEF

Lexile Word Frequency Profiles

Jeff Elmore

AUGUST 2016

OBJECTIVE

This report describes the development of a new type of frequency measure for words that better reflects the developmental nature of word exposure. The relationship between word frequency and word knowledge has been well documented (Brysbaert, Buchmeier, Conrad, Jacobs, Bölte, & Böhl, 2011; Rudell, 1993). Indeed, word frequency is the operational measure of semantic difficulty in the equation powering the Lexile Analyzer® (Stenner, Horabin, Smith, & Smith, 1988; Stenner, Burdick, Sanford, & Burdick, 2007). However, the underlying theoretical explanation of why word frequency predicts word knowledge is exposure. Readers are exposed more often to more frequent words, and thus have greater knowledge of them (Klare, 1963). The connection between word frequency and word knowledge therefore is more meaningful if the word frequencies more accurately reflect the degree of exposure to a word for the average developing reader.

Leveraging the power of the 1.4-billion-word MetaMetrics® corpus of texts intended for readers in U.S. K-12 schools, we developed the Lexile Word Frequency Profile, a set of frequency measures describing the developmental trajectory of a word's occurrences along the Lexile scale. We demonstrate the meaningfulness and utility of Lexile Word Frequency Profiles by examining a few specific word profiles and using the profiles to predict other measures of word familiarity.

METHODS

Lexile measures (Stenner et al., 1988): a developmental scale that measures reader ability and text complexity on a common scale using semantic and syntactic features. Independent psychometric studies of the Lexile scale (Mesmer, 2007; White & Clement, 2001) indicate that it is a valid and reliable measure of reader ability and text complexity.

Data Sources

MetaMetrics Corpus: The corpus used in this study comprised 91,935 texts including textbooks, trade books, leveled readers, and other texts such as supplemental textbook material intended for K-12 students. Lexile text measures typically range from above 200L to below 1400L. The corpus contains 255,744 unique words appearing at least 10 times, totaling 1,390,320,260 running words.

Age-of-Acquisition Ratings: A database of word familiarity ratings based on the estimated age at which a person will know the meaning of a word (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). The database contains ratings from ages 1.5 to 25 for 51,625 words and is used to assess the additional predictive power of Lexile Word Frequency Profiles over a single frequency measure.

PROCEDURES AND ANALYSES

Books were digitized and edited according to the guidelines for Lexile analysis. Occurrences of each word were tallied for each Lexile Zone. A Lexile Zone contains all texts within a 100L range, for example the 200L Zone contains texts measuring from 200L to 299L. Since each text in the corpus has a Lexile measure, frequency counts can be generated for occurrences in texts within each Lexile Zone. For example, we can count the number of times the word *rabbit* appeared in 0L to 99L texts, 100L to 199L texts, etc. A Lexile Word Frequency Profile consists of a set of frequency measures from below 0L to 2200L. In addition to a raw count of each word in each zone, several other kinds of counts were calculated and are described below.

Word Family Frequency Counts

To account for the derivational nature of the English language, frequency counts were generated for several levels of morphological word family relationships. For example, the word *uncommonly*, which may be relatively infrequent, is likely a more familiar word because it is composed of the relatively frequent word *common* and two frequent affixes *un-*, and *-ly*. Because not all derivations are equally transparent, four levels of morphological relationships were considered and word frequency counts were calculated at each level (Elmore, Fitzgerald, Graves, & Bowen, 2015):

- Level 1—every word form is counted uniquely.
- Level 2—base words and their inflected forms and derived forms with the suffixes such as *-ed* (past), *-en* (past participle), and *-ing* (present participle) are counted together.
- Level 3—all the forms in Levels 1 and 2 plus the 10 most frequent prefixes and suffixes such as *-ly* and *un-* are counted together.
- Level 4—all the forms in Levels 1 through 3 plus 107 prefixes and 108 suffixes listed in the English Lexicon Project (ELP) database (Balota et al., 2007) are all counted together. For example, the word *pseudoscientific* would be considered a part of the *science* word family.

Raw, Relative, Cumulative, and Reverse Cumulative Counts

For each word, four types of counts were calculated at each Lexile Zone and for each word family level: raw, relative, cumulative, and reverse cumulative. Relative counts are calculated as the frequency of a particular word in a particular zone divided by the total number running words in that zone. Cumulative counts tally the number of occurrences in a particular Lexile Zone and all previous zones. Reverse cumulative counts tally the number of occurrences in a Lexile Zone and all subsequent zones (e.g., the number of occurrences in all texts above 1200L). Finally, confidence intervals were calculated for all of the counts (Brown, Cai, DasGupta, 2001).

All possible combinations of Lexile Zones, word family levels, and types of counts were calculated. For example, for the word *jump*, one could access a count of the number of times either *jump, jumps, jumper,* or *jumping* occurred in all texts 600L and below, or the relative frequency of just the word 'jumper' in only texts from the 200L Lexile Zone, with an estimated 95% confidence interval.

After Lexile Word Frequency Profiles were created for all words, several pairs of words with approximately the same overall frequency were compared to illustrate the additional information contained in the profile above and beyond a single frequency measure.

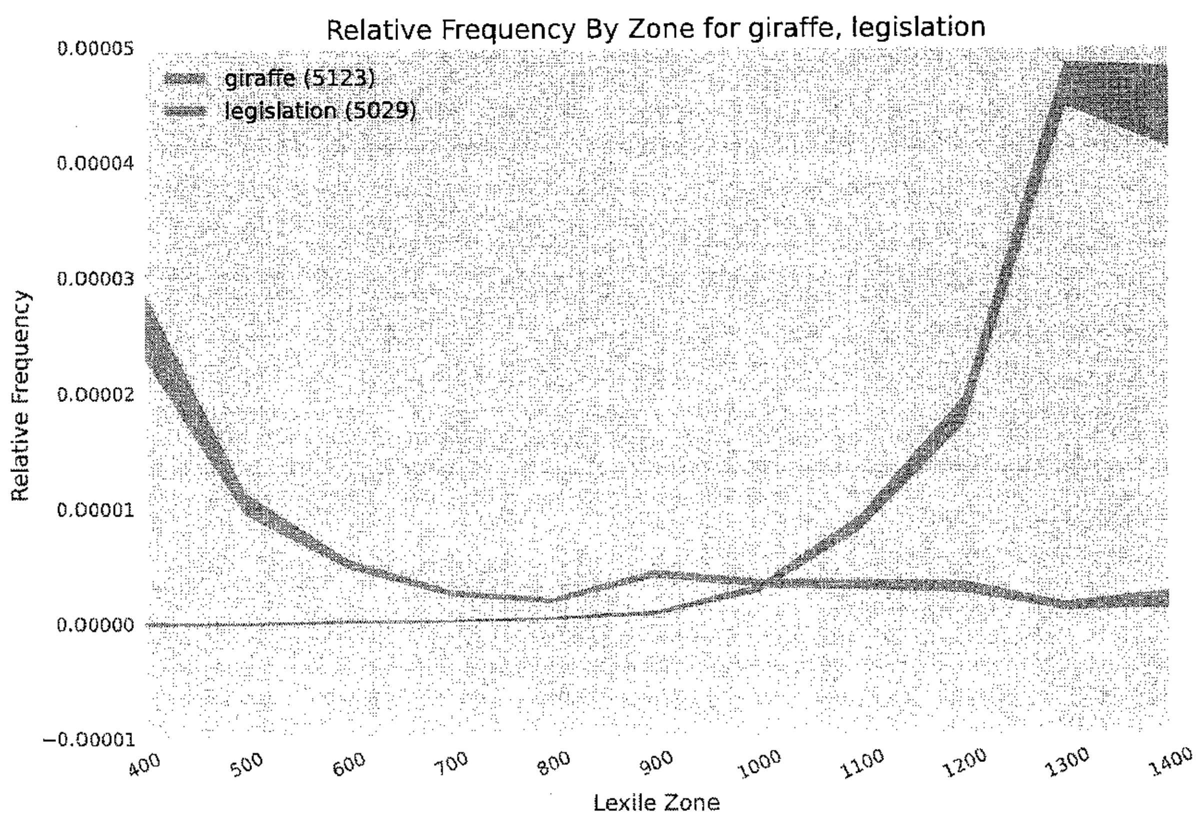
Finally, the power of Lexile Word Frequency Profiles to predict another measure of word familiarity was compared with the baseline of a prediction using a single summative word frequency measure. To evaluate the additional value of Lexile Word Frequency Profiles, two machine learning models called "Random Forest Regression" models (Breiman, 2001) were developed to predict age-of-acquisition ratings (Kuperman, et al., 2012): (1) a model using only the overall frequency, and (2) a model using all 624 Lexile Word Frequency Profile measures. Accuracy of the models was evaluated using the out-of-bag R^2 measures.

RESULTS

In total, 624 individual frequency measures were calculated for 255,744 words. Next, we selected several pairs of words with similar overall frequencies to demonstrate the additional information that Lexile Word Frequency Profiles provide. Although two words may occur about the same number of times overall, they can often have significant differences in the patterns of their use as reflected in their profiles.

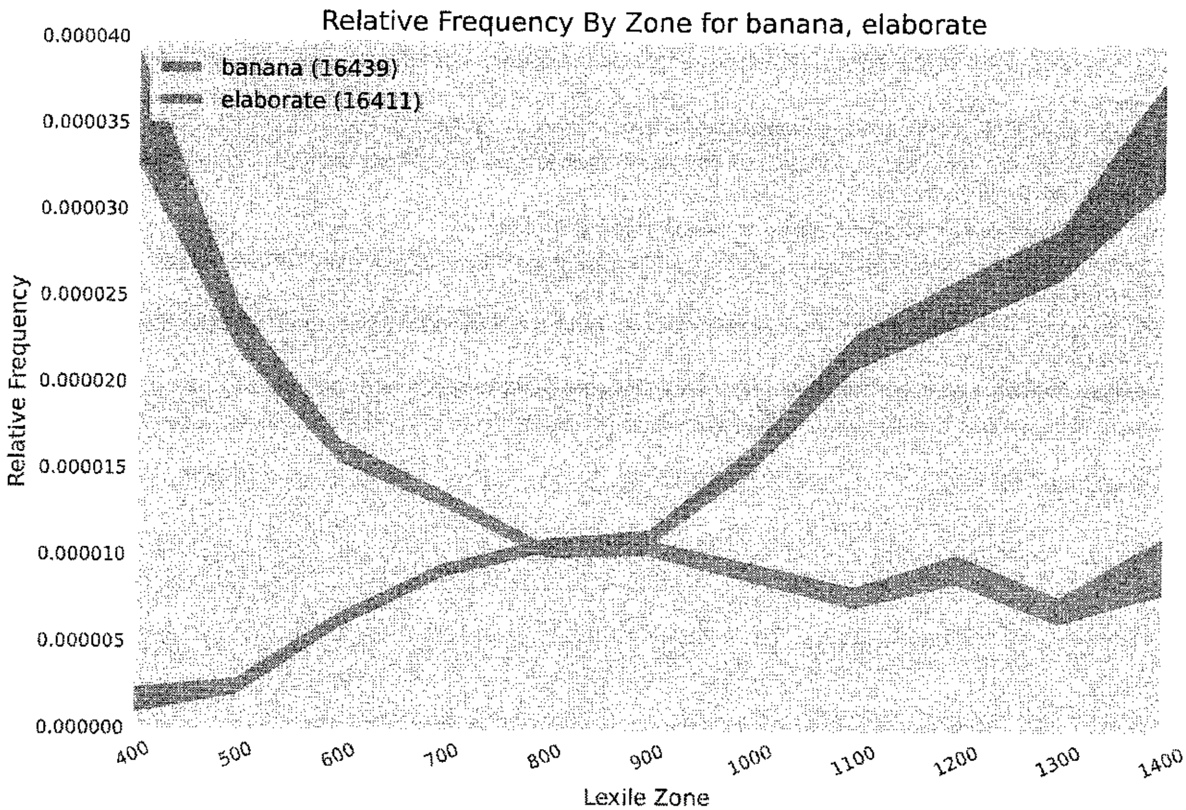
It was hypothesized that so-called academic words that tend to appear in school reading materials but not in stories or in conversations, like *analyze* or *consider*, would likely be more concentrated in higher Lexile texts; whereas, words representing familiar concrete objects such as animals and foods would likely be more concentrated in lower Lexile texts. For example, Figure 1 shows the Lexile Word Frequency Profiles for the words *giraffe* and *legislation* which both appear approximately 5,000 times overall, but have radically different profiles with *giraffe* occurring most frequently in lower Lexile texts and *legislation* occurring most frequently in higher Lexile texts above 1100L.

Figure 1.



Similarly, in Figure 2, *banana* and *elaborate* both occur approximately 16,000 times overall, but have a similar pattern of difference in their Lexile Word Frequency Profiles to the words in Figure 1.

Figure 2.



These examples suggest that Lexile Word Frequency Profiles do indeed capture meaningful differences in word usage for important categories of words such as academic words and concrete nouns like animals and foods.

However, anecdotal evidence from a few examples is less compelling than examining differences over tens of thousands of words. To assess more broadly the additional value of Lexile Word Frequency Profiles, we assessed the power of the profiles to predict another word familiarity measure. Two different random forest models were fit to predict age-of-acquisition ratings for approximately 50,000 words. The first model was a baseline model using a single summative measure of word frequency. The second model was a Lexile Word Frequency Profile model using all 624 measures. The baseline model accounted for 25% of the variance in age-of-acquisition measures while the Lexile Word Frequency Profile model accounted for 75% of the variance in age-of-acquisition measures.

CONCLUSION

Considering evidence from both a visual inspection of pairs of similarly frequent words and quantitative analyses related to predicting other word familiarity metrics, Lexile Word Frequency Profiles appear to provide a rich source of information about the familiarity of words. The profiles may offer insight into how different kinds of words exhibit different patterns of usage in texts of varying complexity level corresponding with different levels of likely exposure for the average reader. More accurate information about word exposure is potentially of value to publishers, curriculum developers, educators, and researchers.

REFERENCES

Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39*, 445–459.

Breiman, L. (2001). Random Forests. *Machine Learning, 45*(1), 5-32.

Brown, L.D., Cai, T.T., & DasGupta, A. (2001). Interval Estimation for a Binomial Proportion. *Statistical Science, 16*(2), 101–133.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A.M., Bölte, J., & Böhl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology, 58*(5), 412–424.

Elmore, J., Fitzgerald, J., Graves, M., & Bowen, K. (2015). The Vocabulary of Elementary Grades Disciplinary Textbooks (MetaMetrics Research Brief). Durham, NC: MetaMetrics, Inc..

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

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978-990

Rudell, A.P. (1993). Frequency of word usage and perceived word difficulty: Ratings of Kucera and Francis words. *Behavior Research Methods, Instruments, & Computers, 25*(4), 455-463.

Stenner, A.J., Horabin, I., Smith, D.R., & Smith, M. (1988). The Lexile Framework. Durham, NC: MetaMetrics, Inc.

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

Facilitator Guide for Analyzing Student Work

Arranging the Room

- Seat participants so that everyone can see each other around the table.
- Although you are the leader of the group, you are still a member and should sit around the table with the others.
- Eliminate as many distractions as possible: turn off all phones, ask the office not to make announcements in the room, post a sign on the door alerting others that a meeting is in progress.

Planning for the Analysis

- Review the standard(s) and the strategy around which the lesson was built.
- Define the expectation for student performance.

Low Performance	*Expected Performance*	*High Performance*
What will it look like if students have not mastered the expected knowledge or skills?	What will it look like if students are appropriately meeting the knowledge and skill expectations?	What will it look like if students have some key knowledge and skills beyond the expectation?

- Ideally, these expectations will be set when the lesson is being developed—before it is delivered to students. Students would benefit from having these explained as part of the instruction they receive.

Analyzing the Student Work: The Three-Stack Method

- Have the teacher providing the student work state the objective of the lesson and briefly describe what occurred during implementation of the lesson and strategy.
- Remind teachers that they are not evaluating individual student performance. They are looking for instructional implications apparent in the work. Review the standard for mastery that was agreed upon during the planning.
- Mix together all student work samples and ask the teachers to sort the work into three piles/stacks: work demonstrating low performance, work demonstrating the expected performance, and work demonstrating high performance.
- Have the teachers describe what is common about or what characterizes the work within each gradation of performance. Look for patterns. Record the information.

Low Performance	*Expected Performance*	*High Performance*
___% of class	___% of class	___% of class

- Using the information about what characterizes work in each stack, consider the instructional implications by asking questions such as:

Deborah K. Reed, Ph.D.

deborah-reed@uiowa.edu

- What does this tell us about our planning and/or implementation?
- How does this influence the way we think about the standard(s)?
- Is there evidence that students had sufficient background knowledge
- Is there evidence that students had misconceptions of the concepts?
- Are there patterns in students' strategy use?
- What does this tell us about students' literacy abilities?
- How does this align with standards-based assessment data?
- What will it take to move students to the next level of performance?
- How does this influence the way we will plan support for students with low performance?
    - What can be incorporated into the upcoming lessons?
    - What needs to be addressed in a targeted lesson?

Facilitating Participation

- Make sure that all members contribute to the discussion. Those who do not volunteer, should be asked the following questions:
    - What do you notice about the student work?
    - Was there anything the teacher described about the implementation of the lesson that could have contributed to the results?
    - What do you think this student work is telling us about our instruction?
    - Do you have any suggestions for future planning?
- Remind members (both at the start of the sessions and periodically throughout the session) that all discussion is to be focused on analyzing the students work—not on complaints, management issues, or particular concerns with one child. This session is about the instructional strategy the teacher chose to implement.
- If members offer opinions, ask them to show the student work samples that provide the data or examples of their statements.
- If the teacher felt the lesson did not go entirely smoothly, offer support that prevents the teachers from wanting to abandon the strategy. Focus instead on how problems the teacher encountered can be improved or corrected for the next lesson. Ask the other members to make suggestions when one person points out a difficulty.

Concluding the Meeting

- Establish the plan for adjusting the lesson (or future lessons), including what will be done, for which students, when, and by whom. Have someone record and share these plans.
- As a group, reflect on how the process of analyzing student work "felt."
    - Did it occur in a safe professional environment?
    - Did it enlighten anyone about student learning across classrooms?
    - Did it change the way you normally interact with your colleagues?
    - Will it impact your teaching practice?
- Make appointments to meet with or visit teachers who have lingering concerns.
- Establish the date, time, and location of your next meeting.

Set the Expectation

Lesson: ______________________________

Evaluate Student Performance

Lesson: ______________________________

_____% of class	_____% of class	_____% of class

ACTION PLAN FOR EXAMINING STUDENT WORK TEMPLATE

Objective for Examining Student Work:

TASK/ACTIVITY	BY WHOM?	WHEN?	RESOURCES NEEDED

Team Members

Name	Position	Date

Summary of Collaborative Data Review Team's Examination of Student Work:

bringing student work samples

12/12
after IEP

Set the Expectation

Lesson: Students will write math answers in complete sentences.

* only number answer, not work/thinking or * shows work and answer, but is not written in a complete sentence or * answer in sentence does not include unit(s)	* shows work/thinking and * writes answer in a complete sentence and * includes unit(s) in answer	All 3 expectations AND * student explains thinking/answer in words

Team:

Team:

Deborah K. Reed, Ph.D.

deborah-reed@uiowa.edu

- bringing student work samples

12/12 after 1st

Set the Expectation

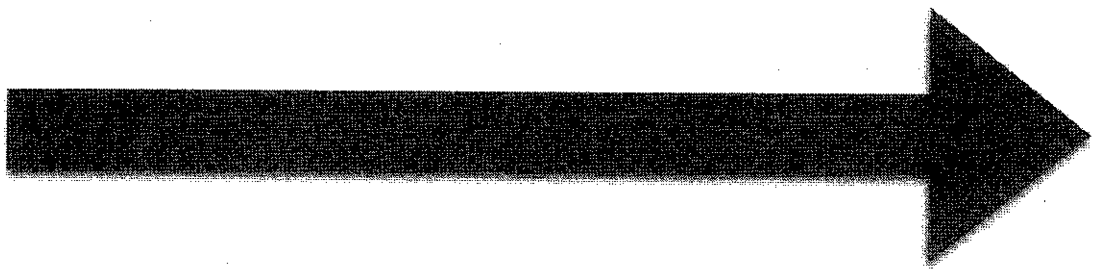
Lesson: Students will produce clear, relevant claims.

* claim is missing or is more than 2 sentences or * Claim is unclear, includes 1st person pronouns, or is a run-on sentence or * Claim does not include a relevant or complete analysis or * Claim is a summary	* claim is written using 1 or 2 sentences. and * claim is clear & assertive (no "I" statements; concise) and * claim includes an analysis (not a summary)	* Claim includes all 3 expectations AND * no spelling or punctuation errors ? and/or * thorough, in-depth analysis that connects to the text

Team: Team:

COLLABORATIVE DATA TEAMS CHECK LIST

CONTINUUM OF DATA USE:
adapted by Gurzon & Guckenburg (2015) from Means, Padilla, & Gallagher (2010)



STAGE 1
Focus of work on school improvement planning using accountability data and placement decisions

STAGE 2
Focus of work on curriculum development, student grouping, and attention by teachers on **what to teach** (e.g., re-teach groups)

STAGE 3
Focus of work on comparative analysis of student growth, instructional changes based on **how to teach**—entry into continuous improvement

STAGE 4
Focus on linking assessment and data literacy at the classroom level to address **how students learn**, with immediate feedback based on data

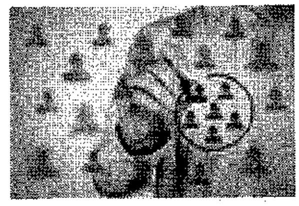
COMPONENTS OF COLLABORATIVE DATA TEAMS:
adapted by Reed (2015)

1. Multiple forms of data are collected:
   - Instructional
   - Screening
   - Diagnostic
   - Progress monitoring
   - Outcome/summative

2. Data teams have a basic structure for operating:
   - Are formed at each grade level or within each department
   - Meet monthly
   - Start and stop on time
   - Have data accessible
   - Assigned roles for members

3. Data teams have established processes:
   - Members contribute equally
   - Members know how to analyze data
   - Members identify problems in data
   - Members develop written plans for solving problems

COLLABORATIVE DATA TEAMS



MEMBERS OF THE DATA ANALYSIS TEAM:

- Team Leader
- Data Specialist
- Recorder
- Time Keeper
- Focus Monitor
- Engaged Participant(s)

MINIMUM REQUIREMENTS FOR A DATA TEAM

- Collegiality
- Data Analysis: What you are looking for, how you will look for it, and how you know whether you have found it.
- Problem Solving: Explicit decision rules for assessing student, class, grade, school-level progress.
- Action Plan

DATA TEAM STRUCTURES

- Meet regularly—once a month, set date and time
- Be punctual
- Have an agenda
- Have meeting minutes, notes—for review, for absent members
- Have data available
- Members should know roles

SAMPLE PROTOCOL

- Stage 1-Facilitator: Introduction—review goal (0-3 minutes)
- Stage 2-Data Discussion (10-15 minutes)
    - Review data
    - Discuss what has shown growth
- Stage 3-Instructional Strategies (5 minutes)
- Stage 4-Action Plan-(10 minutes) Identify **3** strategies to put in place **<u>NOW</u>**
- Stage 5-Observer Comments

DATA COORDINATION

- School-wide assessment schedule
- Training for administering assessments
- Designated individual(s) for managing data
- System for pulling together data
    - Historical at grade transitions
    - Immediately within year
- Training for interpreting data
- Status reports for stakeholders

DRAFT #2-8/11/15-Sue Abel

Collaborative Data Teams Needs Assessment

School: ____________________

Evaluator: ____________________ Date: ____________________

1. There are data teams formed at each grade level or within each department.
   ▯ YES ▯ NO (skip questions 2- 4 and answer questions 5 - 15)
2. Data teams are meeting on a monthly basis.
   ▯ YES ▯ NO
3. The data teams have basic structures for operating:
   a. Start and stop on time ▯ YES ▯ NO
   b. Have data accessible ▯ YES ▯ NO
   c. Assigned roles for members ▯ YES ▯ NO
4. The data teams have established processes:
   a. Members contribute equally ▯ YES ▯ NO
   b. Members know how to analyze data ▯ YES ▯ NO
   c. Members identify problems in data ▯ YES ▯ NO
   d. Members develop written plans for solving problems ▯ YES ▯ NO
5. Multiple forms of data are gathered:
   a. Instructional quality ▯ YES ▯ NO
   b. Screening ▯ YES ▯ NO
   c. Diagnostic ▯ YES ▯ NO
   d. Progress monitoring ▯ YES ▯ NO
   e. Outcome/summative ▯ YES ▯ NO
6. The school has an assessment schedule that all teachers follow.
   ▯ YES ▯ NO
7. Teachers receive expert training on how to administer assessments.
   ▯ YES ▯ NO
8. Teachers receive expert training on how to interpret data.
   ▯ YES ▯ NO
9. There is a designated individual for managing data at the school.
   ▯ YES ▯ NOT

10. Data are made available in a timely and efficient manner.
☐ YES ☐ NO

11. There is an electronic system for aggregating data from multiple sources.
☐ YES ☐ NO

12. There is an electronic system for displaying data within and across years.
☐ YES ☐ NO

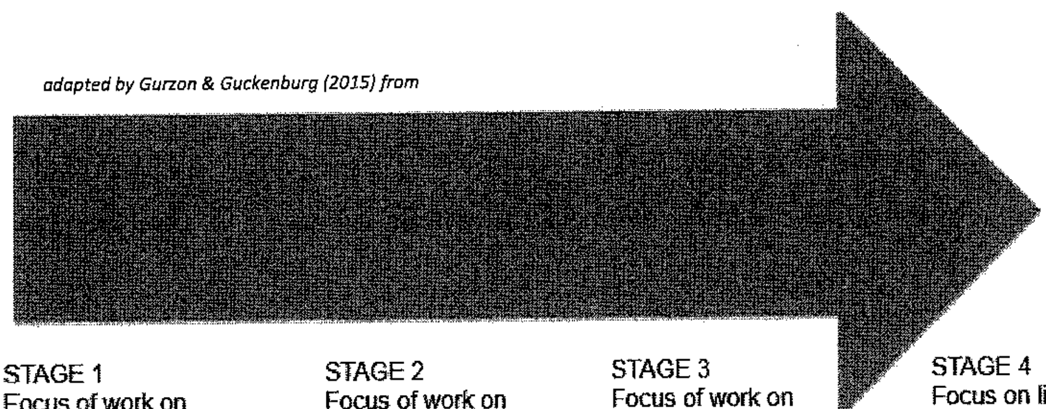
13. There are formal plans for communicating data to stakeholders.
☐ YES ☐ NO

14. There is agreement among faculty and administrators about the data to use and how to use it.
☐ YES ☐ NO

15. The principal is a model of data use on the campus.
☐ YES ☐ NO

Where is the school on the continuum of data use?

adapted by Gurzon & Guckenburg (2015) from



STAGE 1
Focus of work on school improvement planning using accountability data and placement decisions

STAGE 2
Focus of work on curriculum development, student grouping, and attention by teachers on **what to teach** (e.g., re-teach groups)

STAGE 3
Focus of work on comparative analysis of student growth, instructional changes based on **how to teach** – entry into continuous improvement

STAGE 4
Focus on linking assessment and data literacy at the classroom level to address **how students learn**, with immediate feedback based on data

Handout 4: Barriers to a culture of data use, by framework element

Participate in the flow of information for evidence use

- Data management systems are cumbersome, inefficient, and frustrating for teachers to use.
- Technical limitations of data systems suppress data use.
- Educators go to great lengths to compensate for a lack of integration with data systems and might even have built their own systems to address the lack of a district system.
- Teacher leaders take on the task of coordinating and preparing data for teacher team meetings, usually during their own time.
- Educators can become frustrated with the amount of time it takes to access and analyze data, most frequently because of a lack of system integration.

Provide resources and assistance to make meaning from data

- Data are used, but there is little evidence of collegiality.
- Protocols for data use are followed but only at the most procedural and basic level. Team dialogue does not focus on dialogue about instructional change.
- Structures and protocols for collaboration are not used or understood.
- Teachers who believe in data-use practices do it on their own time—before school, after school, or at lunch.
- Ineffective access to data prevents collaboration and dialogue about instructional practices.

Communicate professional expectations for data use

- Data are mistrusted and are seen as a compliance tool rather than an instructional support.
- Teachers have competing time demands for data use and analysis.
- Formal district policies around data use do not exist.
- Educators appear to be negative about data use, but deeper questioning reveals that it is not the idea of data use in itself that concerns educators. Rather, it is the perceived difficulties that arise with data use—the amount of time required, the lack of access to data, and so on—that prompt the negative responses.
- Different data-use expectations are in play across district and schools.

Provide professional development on data-use knowledge and skills

- Professional development occurs in large group settings and does not address skill development using teachers' own data.
- Professional development takes place in the early stages of data use, but as teachers shift to more challenging practices (for example, interpreting evidence to use during instruction), professional learning is no longer focused on data use.
- Evidence about teachers' current practices and learning needs is not collected, and professional learning is not aligned to teacher learning needs.
- Different messages from competing professional development providers hinder coherence and application of new practices.

Provide leadership to nurture a culture of data use

- Time is not provided for collaborative review of data, or the time that is available is barely monitored for effective practices.
- The culture is one in which teachers do not feel safe revealing where they need to improve practice, and leaders (inadvertently or not) punish teachers for sharing areas of weakness or concern.
- Decisions at the district or school level are made "from the gut" and do not model effective data-use practices.

Gerzon, N., and Guckenburg, S. (2015). Toolkit for a workshop on building a culture of data use (REL 2015–063). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from http://ies.ed.gov/ncee/edlabs.

Handout 5: Examples of policy and guidance to support a culture of data use, by framework element

Participate in the flow of information for evidence use

- Written expectations for data use that show which practices align with improved practice and document the shift away from an accountability-based data-use approach.
- Written clarification identifying which users of data are meant to answer which questions.
- Calendars and timelines of district data-use expectations by grade.
- Description of district data systems to clarify functions and uses.
- Description of district data systems that describe how they support and align with the everyday work of educators.

Provide resources and assistance to make meaning from data

- A school calendar documenting scheduled time to analyze evidence.
- A common location (computer drive) with written protocols for using evidence during team meetings.
- Written role expectations for data teams with specific team-member functions outlined.
- A schedule with opportunities for calibration.
- Agreed upon norms for data use posted and reviewed at all meetings.
- Written documentation of how coaches or instructional specialists will support collaborative data-use practices.

Communicate professional expectations for data use

- Written guidance about the focus of data use as designed to support all students.
- Written communication highlighting how the focus on all students will raise the achievement of struggling learners through increased differentiation and personalized student supports.
- Written expectations about annual "products" that include student evidence, including written guidance for use of evidence at parent-teacher meetings, portfolios, and information that is documented across years.

Provide professional development on data-use knowledge and skills

- An annual schedule of professional learning for data use, including formal, informal, large-scale, team-based, and daily learning, focused on common learning goals schoolwide, and including individual (or team) areas of focus for teacher learning.
- Written structure to document teacher learning goals regarding using data.
- Protocol for principals to review teachers' practices with various types of data use (including both schoolwide and classroom uses) and to outline next steps in their individual learning.

Provide leadership to nurture a culture of data use

- Job descriptions that capture the role of data-use leaders throughout the district: principals, teacher leaders, data team leaders, district leaders.
- A hiring protocol outlining expectations or activities that show facility with data use.
- Documentation related to how leaders learn data use over time, with increased expectations outlined over time.

Gerzon, N., and Guckenburg, S. (2015). Toolkit for a workshop on building a culture of data use (REL 2015–063). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from http://ies.ed.gov/ncee/edlabs.

Changing expectations of data use – teacher activities on the continuum

Material adapted from Means, Padilla and Gallagher, 2010

STAGE 1
Focus of work on school improvement planning using accountability data, and placement decisions

STAGE 2
Focus of work on curriculum development, student grouping, and attention by teachers on **what to teach** (ex. re-teach groups)

STAGE 3
Focus of work on comparative analysis of student growth, instructional changes based on **how to teach** – Entry into continuous improvement

STAGE 4
Focus on linking assessment and data literacy at the classroom level to address **how students learn**, with immediate feedback based on data

REL Northeast & Islands
Regional Educational Laboratory at EDC
relnei.org

EDC Learning transforms lives.

Major Partners: AIR WestEd

38

Table 1. Conditions for data use: Guiding questions

Conditions for data use	Guiding questions
Quality	• What data do we have that can help answer the questions we are currently asking about student learning? • What improvements to our data quality would expand our ability to ask and answer these and other questions?
Capacity	• What are the organizational structures and systems that enhance (or impede) our ability to use data effectively? • Do all members of our school or district have the data they need to make effective decisions? • Do all members of our school or district have the knowledge and skills necessary to make use of the data available to them?
Culture	• Are we basing the decisions we need to make on data and evidence? • Are we using data to communicate our decisions in ways that foster engagement by all stakeholders in improvement efforts?

References

Balfanz, R., & Byrnes, V. (2006). Closing the mathematics achievement gap in high poverty middle schools: Enablers and constraints. *Journal of Education for Students Placed at Risk (JESPAR), 11*(2), 143–159.

Team data analysis protocol: Five-phase data review

Stage 1: Predict (5–8 minutes)	Stage 2: Go visual (5–8 minutes)	Stage 3: Observe and analyze (10–15 minutes)	Stage 4: Infer, question and clarify (15 minutes)
The goal during this time is to think about the team's collective knowledge and experience of this standard prior to collective data review. Key questions: What are we thinking about this standard? Are there noticeable misconceptions or challenges students had learning this specific standard? Conversation starters include: · I predict… · I assume… · When I taught this standard I was wondering…	Each team member reviews the data quietly for a few minutes and then shares key facts from the data. Key questions: What do you notice about these data? What "data statements" can be made in reference to these data? What important points seem to "pop out"? Conversation hints: · Share only factual statements · Resist saying "because"	During this collective data analysis, team members look for patterns and trends in the data. Key questions: What patterns and themes are emerging from the data? How does this information compare to the predictions the team made earlier? What do you notice about trends or outlier information from this data set? Conversation starters include: · I notice that… · I see that…	Team members move from looking at causes to determining instructional actions. Key questions: In what ways do these data offer suggestions for next steps to extend student learning? In what ways do these data offer suggestions for next steps to provide interventions for students who require it? Which instructional strategies might we apply? Conversation starters include: · A clear student-learning need is… · The data suggest that… is a specific area of weakness…
Notes:	Notes:	Notes:	Notes: Areas of learning to revisit: Areas of learning needs to extend:

Stage 5: Action planning
Instructional actions for specific identified needs for "revisiting" key standards:
Instructional actions related to next steps in learning for students who are meeting key standards:

Gerzon, N., and Guckenburg, S. (2015). Toolkit for a workshop on building a culture of data use (REL 2015–063). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from http://ies.ed.gov/ncee/edlabs.

Common interim assessment tuning protocol

How to use this tool:

- Select a recently implemented common interim assessment task.
- Complete this reflection tool individually.
- Bring your completed reflection tool, the assessment used, and your planning documents to your team meeting. Discuss your reflections with colleagues.
- Clarify what worked well with this common interim assessment and what you might like to change.

Upon reflection, to what extent did this common interim assessment:

Reflections prior to group review	Very well				Not at all	Group discussion notes
1. Clearly align with the learning goal?	4	3	2	1	0	
	Comments:					
2. Address the knowledge and skills that are documented in the learning progression?	4	3	2	1	0	
	Comments:					
3. Identify and elicit the content or skills where students have misconceptions?	4	3	2	1	0	
	Comments:					
5. Provide the teacher with useful (that is, new) information about each student's level of understanding?	4	3	2	1	0	
	Comments:					
6. Provide the teacher with useful information for adjusting instruction?	4	3	2	1	0	
	Comments:					
7. Lead to a change in instruction? If so, what changed?	4	3	2	1	0	
	Comments:					
8. Lead to an exchange of feedback between student and teacher?	4	3	2	1	0	
	Comments:					

After shared review of responses, discuss the overall trends.

- If your common interim assessment did provide useful guidance to the teacher, what might you take away regarding future strategies for developing or implement-ing common interim assessment tasks?
- If your common interim assessment did not provide useful guidance to the teacher, what might you change in this assessment? What might you want to do differently as you develop or implement your next common interim assessment?

Document your feedback, and share key findings from your discussion with your team and, through team notes, with your administrators.

Gerzon, N., and Guckenburg, S. (2015). Toolkit for a workshop on building a culture of data use (REL 2015–063). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from http://ies.ed.gov/ncee/edlabs.

Grade 1 professional learning team meeting notes template

Date: ______________________________

Facilitator: ______________________________

Note taker: ______________________________

Timekeeper: ______________________________

Attendee	Please initial	Attendee	Please initial
Teacher name		Specialist name	
Teacher name		Literacy coach name	
Teacher name		Administrative liaison name	
Teacher name		Special education teacher name	

Standard: ______________________________

Stage 5: Agreements and recommendations for intervention and extension groups

Action plan: Teacher assignments and next steps for re-teach. Recommendations to response to intervention team.

Copy of agreements and recommendations sent to:

Shared drive ____________ Literacy coach ____________ Team administrative liaison ____________

Gerzon, N., and Guckenburg, S. (2015). Toolkit for a workshop on building a culture of data use (REL 2015–063). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from http://ies.ed.gov/ncee/edlabs.

National School Reform Faculty

Harmony Education Center www.nsrfharmony.org

ATLAS
Looking at Data

Learning from Data is a tool to guide groups of teachers discovering what students, educators, and the public understands and how they are thinking. The tool, developed by Eric Buchovecky, is based in part on the work of the Leadership for Urban Mathematics Project and of the Assessment Communities of Teachers Project. The tool also draws on the work of Steve Seidel and Evangeline Harris-Stefanakis of Project Zero at Harvard University. Revised November 2000 by Gene Thompson-Grove for NSRF. Revised August 2004 for Looking at Data by Dianne Leahy.

Selecting Data to Share
Data is the centerpiece of the group discussion. The following guidelines can help in selecting data or artifacts that will promote the most interesting and productive group discussions. Data or artifacts that do not lead to a single conclusion generally lead to rich conversations.

Sharing and Discussion of Data
Discussions of some forms of data sometimes make people feel "on the spot" or exposed, either for themselves, for their students or for their profession. The use of a structured dialogue format provides an effective technique for managing the discussion and maintaining its focus.

A structured dialogue format is a way of organizing a group conversation by clearly defining who should be talking when and about what. While at first it may seem rigid and artificial, a clearly defined structure frees the group to focus its attention on what is most important. In general, structured dialogue formats allot specified times for the group to discuss various aspects of the work.

1. Getting Started
- The facilitator reminds the group of the norms.
  Note: Each of the next four steps should be about 10 minutes in length. It is sometimes helpful for the facilitator to take notes.
- The educator providing the data set gives a very brief statement of the data and avoids explaining what s/he concludes about the data if the data belongs to the group rather than the presenter.

2. Describing the Data (10 Minutes)
- The facilitator asks: "What do you see?"
- During this period the group gathers as much information as possible from the data.
- Group members describe what they see in data, avoiding judgments about quality or interpretations. It is helpful to identify where the observation is being made—e.g., "On page one in the second column, third row . . . "
- If judgments or interpretations do arise, the facilitator should ask the person to describe the evidence on which they are based.
- It may be useful to list the group's observations on chart paper. If interpretations come up, they can be listed in another column for later discussion during Step 3.

3. **Interpreting the Data (10 Minutes)**

- The facilitator asks: "What does the data suggest?" Second question: "What are the assumptions we make about students and their learning?"
- During this period, the group tries to make sense of what the data says and why. The group should try to find as many different interpretations as possible and evaluate them against the kind and quality of evidence.
- From the evidence gathered in the preceding section, try to infer: what is being worked on and why?
- Think broadly and creatively. Assume that the data, no matter how confusing, makes sense to some people; your job is to see what they may see.
- As you listen to each other's interpretations, ask questions that help you better understand each other's perspectives.

4. **Implications for Classroom Practice (10 Minutes)**

- The facilitator asks: "What are the implications of this work for teaching and assessment?" This question may be modified, depending on the data.
- Based on the group's observations and interpretations, discuss any implications this work might have for teaching and assessment in the classroom. In particular, consider the following questions:
  — What steps could be taken next?
  — What strategies might be most effective?
  — What else would you like to see happen? What kinds of assignments or assessments could provide this information?
  — What does this conversation make you think about in terms of your own practice? About teaching and learning in general?
  — What are the implications for equity?

5. **Reflecting on the ATLAS-Looking at Data (10 Minutes)**

Presenter Reflection:

- What did you learn from listening to your colleagues that was interesting or surprising?
- What new perspectives did your colleagues provide?
- How can you make use of your colleagues' perspectives?

Group Reflection:

- What questions about teaching and assessment did looking at the data raise for you?
- Did questions of equity arise?
- How can you pursue these questions further?
- Are there things you would like to try in your classroom as a result of looking at this data?

6. **Debrief the Process**

- How well did the process work?
- What about the process helped you to see and learn interesting or surprising things?
- What could be improved?

40 MINUTE DATA MEETING OVERVIEW

Objectives:

- To understand that student outcomes are important for improving achievement
- To understand that looking at data collaboratively provides a method for being accountable for evaluating and modifying our instructional practices to meet student needs
- To understand that students learn better when we work collaboratively

Guiding Questions:

- What do we expect students to learn?
- How will we know what students are learning?
- How will we respond to students who are not learning?

Norms:

- Promptness
- Be prepared
- Show Respect
- Be present
- Be positive
- Assume positive intent

Roles:

- Facilitator
- Timekeeper
- Recorder/Notetaker

Protocol Summary:

- Before the meeting: each team member has a copy of the latest classroom data, has reviewed it and brings a copy to the meeting
- Introduction (2min): Begin the meeting
- Sharing (5min): Successes and Ideas
- Current Challenges (5min): Focus Areas
- Proposed Solutions (10min): Brainstorm strategies as a team
- Action Plan (10min): Agree on a strategy
- Closing the meeting (5min): Debrief and Summarize
- After the meeting: Distribute notes and summaries

Data Analysis Suggested Protocol
Adapted from Solution Tree

DATA MEETING PROTOCOL

- PRIOR TO THE MEETING
  - Data: teachers have up-to-date data and have had time to review for discussion (Classroom Data Analysis Forms are attached)
  - Tools: you will need a flip chart or whiteboard to record ideas; markers; and "parking lot" for off-agenda ideas
  - Agenda: distribute in advance
- INTRODUCTION (2 minutes)
  - Review the purpose or goal for the meeting
  - Review the norms
  - Review agenda
  - Facilitator commits to staying to the agenda: any off-topic ideas will be placed on the Parking Lot chart to be discussed at the end of the meeting or at a later date
- SHARING IDEAS (5 minutes)
  - Record these ideas where everyone can see them
  - Members share successes – you may wish to use Classroom Data Analysis form
  - Members identify areas where students were most improved
- CHALLENGES (5 minutes)
  - Record these ideas where everyone can see them
  - Determine areas of highest nee – you may wish to use Classroom Data Analysis form
  - Identify any common areas of need between classrooms
- PROPOSED SOLUTIONS (10 minutes)
  - Record these ideas where everyone can see them
  - Brainstorm possible solutions for challenges
  - State each possible solution as a concrete, doable intervention
- ACTION PLAN (10 minutes)
  - Examine successful strategies from SHARING IDEAS and ideas from PROPOSED SOLUTIONS
  - Select one strategy that everyone will work on between now and the next meeting
  - Articulate a goal for the team
  - Record the Focus Goal/SMARTE Goal where everyone can see
- CLOSING THE MEETING (5 minutes)
  - Note what went well and what was difficult during the meeting: how well did the team do based on agreed norms and goals of the meeting?
  - Complete the Meeting Summary Form I *or* II
- PARKING LOT (TBD)
  - If time permits, the team may now address the ideas in the Parking Lot
  - Any items not discussed may be placed on the agenda at a later time

MEETING SUMMARY FORM I

Meeting name: ______________________________ Date: ______________

Participants: ____

1. What was the intended goal of this meeting?

2. What were our successes?

3. What did we learn?

4. What is our next goal?

5. What is the focus of our next meeting?

6. Our next meeting will be:
    a. Date:
    b. Time:
    c. Location:
    d. Facilitator:

Data Analysis Suggested Protocol
Adapted from Solution Tree

MEETING SUMMARY FORM II

SUCCESSES	CHALLENGES
SOLUTIONS	**NEXT STEPS**
SMART GOAL:	

Data Analysis Suggested Protocol
Adapted from Solution Tree

CLASSROOM DATA ANALYSIS I

- Proficient on these assessments = ___% and higher
- Highlight each score of less than ___% on the data sheet

"# STUDENTS" = number of students who score BELOW proficient on each skill/standard:

Skill/Standard:	# Students:
Planned intervention for these students:	

Skill/Standard:	# Students:
Planned intervention for these students:	

Skill/Standard:	# Students:
Planned intervention for these students:	

Skill/Standard:	# Students:
Planned intervention for these students:	

CLASSROOM DATA ANALYSIS II

Areas where students performed AT or ABOVE benchmark:

Write the STANDARD or SKILL along with the STRATEGIES used	# students

Areas where students performed BELOW benchmark:

Write the STANDARD or SKILL along with the STRATEGIES used	# students

Ideas for changes in strategies when I teach this skill again:

PCL XL error

Subsystem: xlpaint

Error: Input Stream EOF

Operator: BezierPath

Position: 161978

SCHOOL AND FAMILY PARTNERSHIPS

1. At this school, I feel that I am:

	Strongly Agree	Agree	Disagree	Strongly Disagree	I Don't Know
welcomed	☐	☐	☐	☐	☐
very satisfied with my child's learning experiences	☐	☐	☐	☐	☐

2. At this school, my child(ren) feels/feel:

	Strongly Agree	Agree	Disagree	Strongly Disagree	I Don't Know
care for	☐	☐	☐	☐	☐
safe and secure	☐	☐	☐	☐	☐

3. My child(ren) is/are:

	Strongly Agree	Agree	Disagree	Strongly Disagree	I Don't Know
being provided extra academic help when needed	☐	☐	☐	☐	☐

4. My child's school or teacher(s)

	Strongly Agree	Agree	Disagree	Strongly Disagree	I Don't Know
use teaching strategies that support my student's learning style	☐	☐	☐	☐	☐
has an approach to student behavior that works well for my child	☐	☐	☐	☐	☐

What will make Manzanita better? ______________________________

ASOCIACIONES ESCOLARES Y FAMILIARES

1. En la escuela, yo siento que:

	Muy de Acuerdo	De Acuerdo	En desacuerdo	Muy en desacuerdo	No lo se
Bienvenido	☐	☐	☐	☐	☐
satisfecho con las experiencias de aprendizaje de mi estudiante	☐	☐	☐	☐	☐

2. En la escuela, mi estudiante(s) sienten que:

	Muy de Acuerdo	De Acuerdo	En desacuerdo	Muy en desacuerdo	No lo se
atendidos	☐	☐	☐	☐	☐
seguro y protegido	☐	☐	☐	☐	☐

3. Mi hijo(s) son/estan:

	Muy de Acuerdo	De Acuerdo	En desacuerdo	Muy en desacuerdo	No lo se
han sido proveidos con ayuda academica extra cuando la necesita (n)	☐	☐	☐	☐	☐

4. La escuela de mi estudiante(s) o maestros

	Muy de Acuerdo	De Acuerdo	En desacuerdo	Muy en desacuerdo	No lo se
usan estrategias de ensenanza que respaldan su estilo de aprendizaje	☐	☐	☐	☐	☐
cuentan con una forma de abordar el comportamiento del estudiante que funciona bien para mi hijo/a	☐	☐	☐	☐	☐

Que hara mejor a Manzanita? ______________________________

California Department of Education

English Language Proficiency Assessments for California (ELPAC)
Initial ELPAC, 2018–19
Parent and Guardian Notification Letter

Estimado padre, madre o tutor:

Cuando inscribió a su hijo en la escuela, usted indicó que el idioma nativo o lengua materna de su hijo no es inglés. En las escuelas públicas de California, todos los estudiantes que ingresan en la escuela por primera vez serán evaluados con la Prueba de Suficiencia en el Idioma Inglés Inicial de California (*Initial English Language Proficiency Assessments for California*), o "ELPAC inicial", si su lengua materna no es inglés.

La ELPAC inicial es la prueba utilizada para determinar si un estudiante está aprendiendo inglés o es proficiente en inglés. Esta prueba obligatoria ayudará a identificar a los estudiantes que necesitan ayuda para aprender inglés. Esto es importante para poder brindarles la ayuda necesaria para que se desempeñen bien en todas las materias escolares.

Dados los resultados de la encuesta de lengua materna, **su hijo será evaluado con la ELPAC inicial.**

Usted es una parte importante de la educación de su hijo. Para ayudar a su hijo a prepararse para esta prueba, puede:

- Leerle a su hijo o hacer que su hijo le lea a usted en forma periódica.
- Usar dibujos y pedirle a su hijo que le diga lo que ve en o lo que está ocurriendo en cada dibujo.
- Dar a su hijo oportunidades de usar el idioma fuera de la escuela.
- Hablar con el maestro de su hijo sobre sus destrezas de comprensión auditiva, expresión oral, lectura y escritura, para poder ayudar a su progreso.

Para obtener más información sobre las ELPAC, visite la página web de las "Guías de padres para comprender" *(Parent Guides to Understanding)* del Departamento de Educación de California: https://www.cde.ca.gov/ta/tg/ca/parentguidetounderstand.asp.

También puede ver preguntas de muestra en las pruebas de práctica, que se encuentran en el sitio web de las ELPAC: https://www.elpac.org/resources/practicetests/.

Si tiene alguna pregunta sobre las pruebas ELPAC que va a tomar su hijo, comuníquese con **Patricia Lopez** al **510-222-3500.**

Atentamente,

Chantel Caldwell, Principal
Manzanita Charter Middle School

English Language Proficiency Assessments for California
Initial ELPAC, 2019-2020
Parent and Guardian Notification Letter

Dear Parent/Guardian:

You indicated that your child's primary or home language is not English when registering your child for school. In California public schools, all students entering school for the first time will be assessed with the Initial English Language Proficiency Assessments for California, or "Initial ELPAC," if their home language is not English.

The Initial ELPAC is the test used to determine if a student is an English learner or is fluent in English. This required test will help identify students who need help learning English. This is important so they can get the support they need to do well in all school subjects.

Based on the home language survey results, **your child will be assessed with the Initial ELPAC**.

You are an important part of your child's education. To help your child get ready for the test, you can:

- Read to your child, or have them read to you on a regular basis.
- Use pictures and ask your child to tell you what they see, or what is happening in each picture.
- Provide your child with opportunities to use language outside of school.
- Talk with your child's teacher about your child's listening, speaking, reading, and writing skills to help support their progress.

To learn more about the ELPAC, go to the California Department of Education Parent Guide to Understanding the ELPAC Web page at https://www.cde.ca.gov/ta/tg/ep/elpacparentguide.asp.

You also can look at sample test questions on the ELPAC practice tests, which can be found on the ELPAC Web site at https://www.elpac.org/resources/practicetests/.

If you have any questions about your child taking the ELPAC, please contact **Patricia Lopez** at **510-222-3500**

Sincerely,

Chantel Caldwell, Principal

Manzanita Charter Middle School

English Language Proficiency Assessments for California
Summative Assessment
Parent and Guardian Score Report Letter Template

Dear Parent or Guardian:

Last spring, your child took a test called the Summative English Language Proficiency Assessments for California (ELPAC). This test is part of the California assessment system. The ELPAC helps teachers across the state see how well students listen, speak, read, and write in English.

Your child's 2018–19 Student Score Reports is attached to this letter.

This report shows an overall score and performance level, an oral language (speaking, listening) score and level, and a written language (reading, writing) score and level. It also shows a performance level for each domain: listening, speaking, reading, and writing.

To learn more about your child's scores, go to the new parent web page called *Starting Smarter,* available at https://elpac.startingsmarter.org/.

This site includes:

- resources to help understand results on the student score reports
- access to sample test questions and practice tests
- no-cost resources to support learning
- a guide for parent-teacher conferences

In our district, the test results are just one way to look at how well our students are doing. We use the results to find areas where students are doing well and areas in which they need help. It is also important to know that the test results are not used to determine whether a student moves to the next grade. If you have questions or concerns about your child's progress, please call the school office at 510-222-3500 to arrange a conference with your child's teacher.

Sincerely,

Chantel Caldwell, Principal

Manzanita Charter Middle School

English Language Proficiency Assessments for California (ELPAC) Summative ELPAC Parent and Guardian Notification Letter

Dear Parent/Guardian:

Identifying students who need help learning English is important so they can get the support they need to do well in English language arts/literacy, mathematics, science, and other subject areas in school. The Summative English Language Proficiency Assessments for California, or "Summative ELPAC," is the test used to measure how well students understand English when it is not the language they speak at home. Information from the ELPAC tells your child's teacher about the areas in which your child needs extra support.

This spring, your child will take the Summative ELPAC.

Students in kindergarten through grade twelve who are classified as English learners will take the Summative ELPAC every year until they are reclassified as proficient in English. Students are tested on their skills in listening, speaking, reading, and writing.

ELPAC Testing for the 2019-2020 school year will take place on: April 16th and 23rd.

You are an important part of your child's education. To help your child get ready for the test, you can:

- Read to your child, or have them read to you on a regular basis.
- Use pictures and ask your child to tell you what they see, or what is happening in each picture.
- Provide your child with opportunities to use language outside of school.
- Talk with your child's teacher about your child's listening, speaking, reading and writing skills to help support their progress.

To learn more about the ELPAC, go to the California Department of Education Parent Guides to Understanding Web page at https://www.cde.ca.gov/ta/tg/ca/parentguidetounderstand.asp.

You also can look at sample test questions on the practice tests, which can be found on the ELPAC Web site at https://www.elpac.org/resources/practicetests/.

If you have any questions about your child taking the ELPAC, please contact **Patricia Lopez** at **510-222-3500**

Sincerely,

Chantel Caldwell, Principal
Manzanita Charter Middle School

Guía del Tablero de Información Escolar de California para padres

¿Qué es el Tablero de Información Escolar de California?

El Tablero de Información Escolar de California (o simplemente denominado Tablero) es una **herramienta en línea** que muestra a los padres y a las comunidades qué tan bien las escuelas y los distritos están satisfaciendo las necesidades de los estudiantes. Informa el desempeño tanto en las medidas estatales como las locales. Estas múltiples medidas de éxito reflejan el nuevo sistema de rendición de cuentas de California, que tiene como base las diez áreas de prioridad de la Formula de financiamiento con control local (conocida en inglés como Local Control Funding Formula o LCFF), la cual se describe en la página web del CDE en http://www.cde.ca.gov/fg/aa/lc/lcffoverview.asp.

Medidas estatales

Estas medidas están basadas en la información recopilada a nivel estatal y permiten a los padres comparar las escuelas y los distritos de todo el estado de California.

Seis Medidas estatales

- Absentismo crónico (para los grados del kindergarten al octavo)
- Tasa de suspensión (para los grados del kindergarten al duodécimo)
- Progreso de los aprendices del inglés (para los grados del primero al duodécimo)
- Tasa de graduación (para preparatoria solamente)
- Universidad/carrera profesional (para preparatoria solamente)
- Académica (para los grados del tercero al octavo y el undécimo)
    - Artes del lenguaje inglés y la lectoescritura (ELA, por sus siglas en inglés)
    - Matemáticas

Puede obtener más información sobre cada uno de estos indicadores en el Apéndice A: ¿Qué más debemos saber sobre las medidas estatales?

Medidas locales

Mientras que las medidas estatales se basan en la información recopilada a nivel estatal, las medidas locales se basan en la información recopilada por los distritos. Estas medidas incluyen las condiciones básicas (acreditaciones de los maestros, edificios limpios y seguros y libros de texto para todos los estudiantes); la implementación de los estándares académicos; las encuestas sobre el clima escolar; y la participación y compromiso de los padres.

¿Cómo se mide el desempeño?

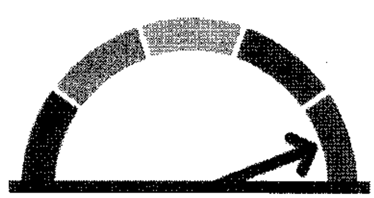
El desempeño en las medidas estatales se basa en los resultados del **año en curso** (**datos de 2018**) y del **año anterior** (**datos de 2017**).

Para cada medida estatal, las escuelas y los distritos reciben **uno de los cinco niveles de desempeño**. Cada nivel de desempeño se identifica con un **color** diferente:

Colores de desempeño	Niveles de desempeño
Azul (el desempeño más alto)	Muy alto
Verde	Alto
Amarillo	Mediano
Naranja	Bajo
Rojo (el desempeño más bajo)	Muy bajo

¿Cómo se informa el desempeño en el Tablero?

En los informes del Tablero, los niveles de desempeño se muestran como **indicadores semicirculares** que están dividido en segmentos de color: rojo, naranja, amarillo, verde y azul. Una flecha apunta al color que representa el nivel de desempeño para esa medida.



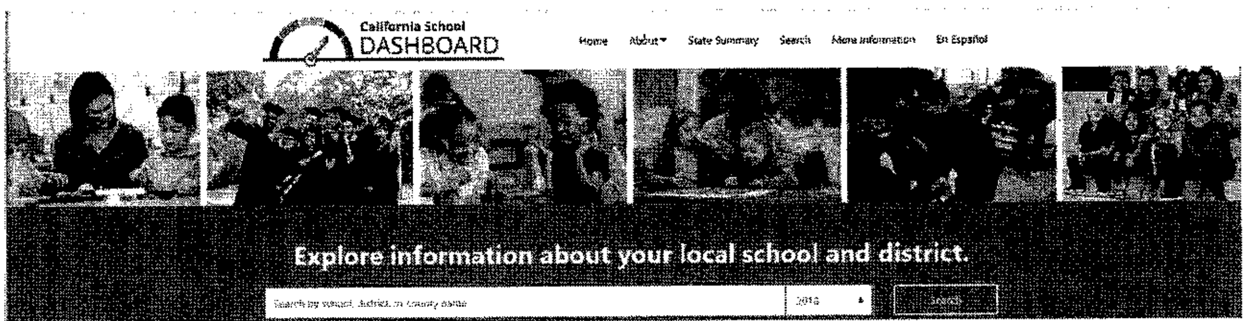
Blue

En el ejemplo anterior, la flecha está apuntando al segmento del extremo derecho que es el nivel de desempeño Azul. Este es el nivel de desempeño más alto. Para obtener la descripción completa de la imagen anterior, consulte el Apéndice B: Imagen 1.

Cómo ver el desempeño de una escuela o distrito en el Tablero de 2018

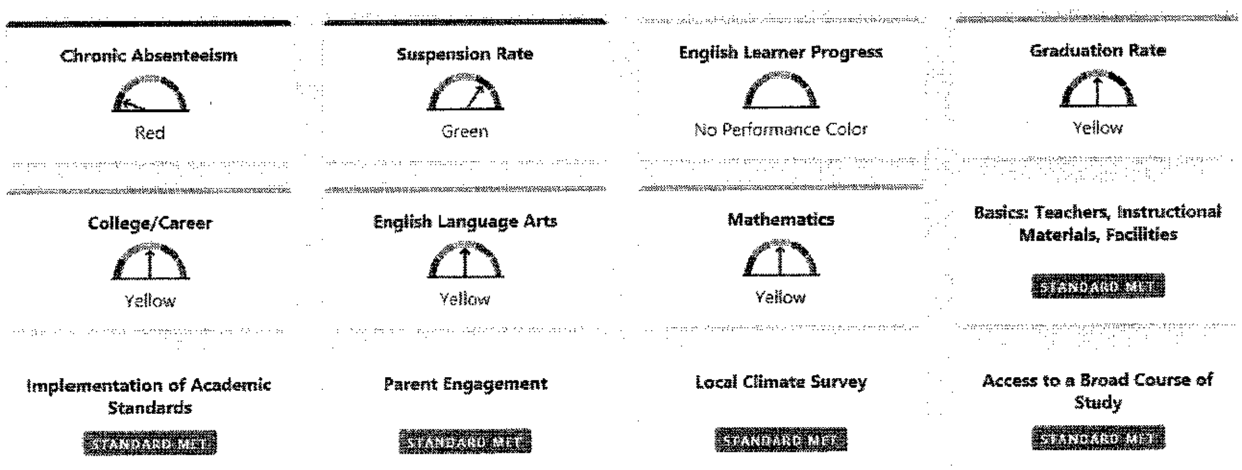
1. El Tablero se puede ver en la página web del Tablero de Información Escolar de California y el Sistema de Apoyo Web que mantiene el CDE en https://www.cde.ca.gov/dashboard.

2. Para ver los resultados de un distrito o escuela, escriba el nombre en la barra de búsqueda. Después seleccione el año.



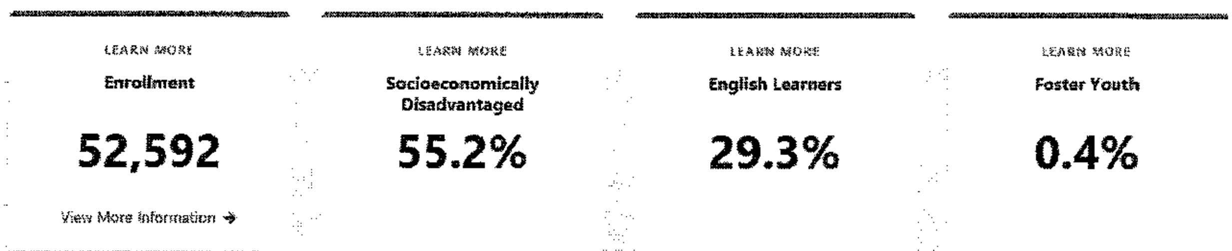
Para obtener la descripción completa de la imagen anterior, consulte el Apéndice B: Imagen 2.

3. Una vez que seleccione su distrito o escuela, podrá ver su **desempeño en las medidas estatales y locales.**



En el ejemplo anterior, vemos que el distrito recibió un nivel de desempeño Rojo para Absentismo crónico y un nivel de desempeño Verde para Tasa de suspensión. El distrito recibió un nivel de desempeño Amarillo para las cuatro medidas estatales restantes: Tasa de graduación, Universidad/carrera profesional, ELA y matemáticas. (Tenga en cuenta que las escuelas y los distritos no recibirán un color para Progreso de los aprendices del inglés en 2018). El distrito también cumplió con el estándar en los cinco indicadores locales. Para obtener la descripción completa de la imagen anterior, consulte el Apéndice B: Imagen 3.

4. Abajo de la vista global del desempeño de la escuela o distrito, encontrará información acerca de la **población estudiantil.**

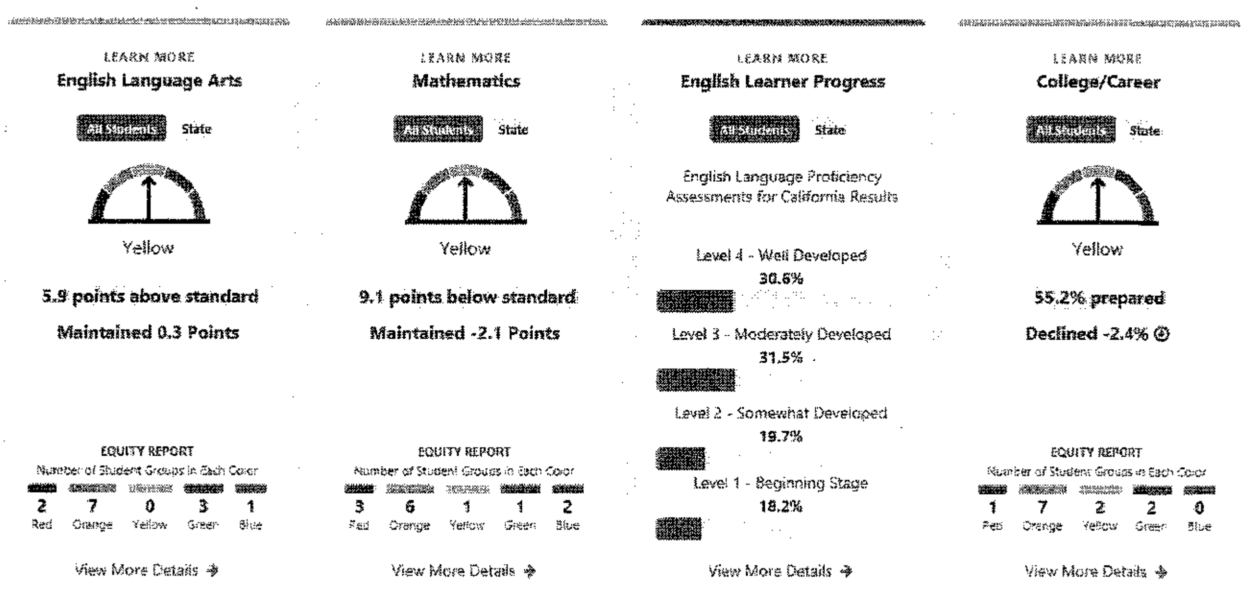


Para obtener la descripción completa de la imagen anterior, consulte el Apéndice B: Imagen 4.

5. Si continúa desplazándose hacia abajo en la página, encontrará datos adicionales sobre cada medida, como los resultados de la escuela o el distrito para el año en curso y si hubo una mejora respecto al año anterior. Las medidas locales y estatales están organizadas en tres áreas:

    - Desempeño académico
    - Participación académica
    - Condiciones y clima

    Veamos las medidas estatales que aparecen bajo Desempeño académico para nuestro distrito de ejemplo.

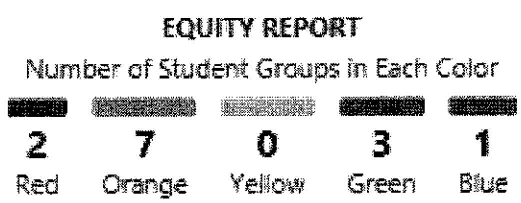


Para artes del lenguaje inglés, vemos que la escuela tuvo un promedio de 5.9 puntos por encima del estándar. Esto representa un aumento de 0.3 puntos respecto al año anterior y se considera que se mantuvo. Para obtener la descripción completa de la imagen anterior, consulte el Apéndice B: Imagen 5.

Para universidad/carrera profesional, vemos que el 55.2 por ciento de los estudiantes de preparatoria del distrito fueron considerados como que estaban preparados para la universidad o una carrera profesional después de su graduación. Esto representa una disminución (-2.4 por ciento) respecto al año anterior.

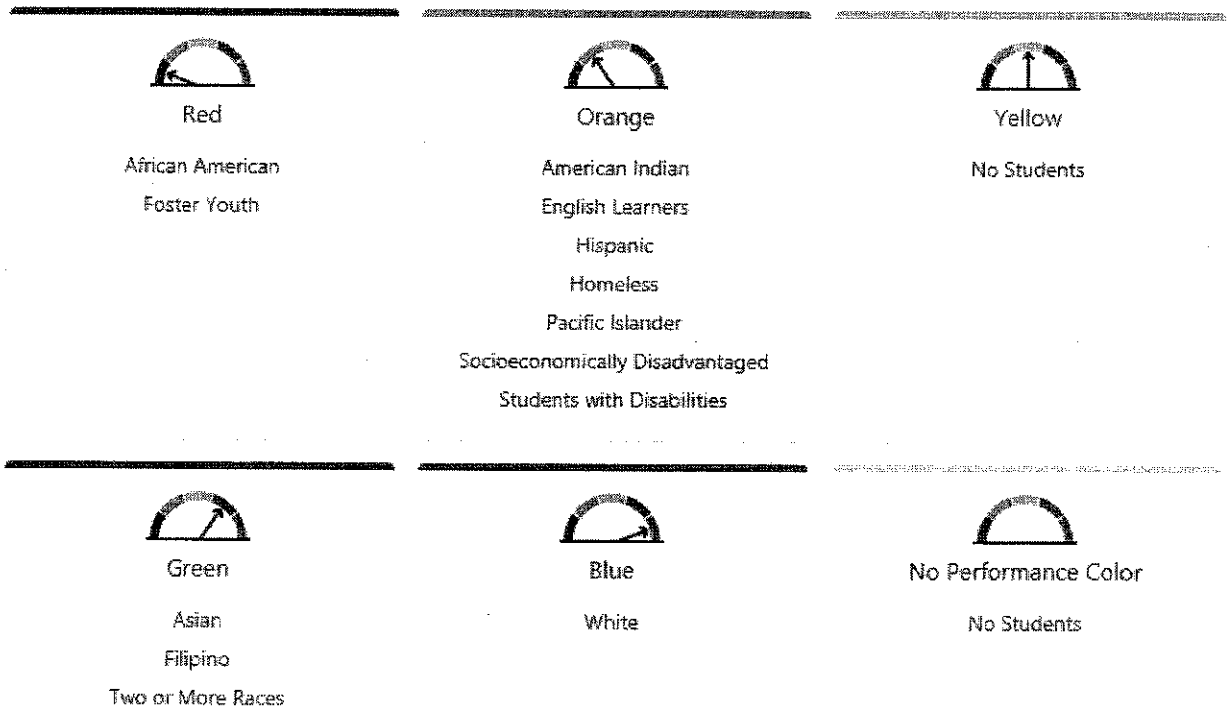
Tenga en cuenta que puede alternar entre "Todos los estudiantes" y "Estado", lo que le permite comparar los resultados de los estudiantes de su escuela o distrito con los resultados de todo el estado.

6. Un **Informe de equidad** se muestra en la parte inferior de la tarjeta para cada medida estatal. Éste muestra el número de grupos de estudiantes asignados a cada nivel de desempeño (color) para esta medida. El siguiente es un Informe de equidad para matemáticas de un distrito de ejemplo.



Al observar los datos anteriores, vemos que dos grupos de estudiantes recibieron un color Rojo (el nivel de desempeño más bajo) para ELA y un grupo de estudiantes recibió un color Azul (el nivel de desempeño más alto). Para obtener la descripción completa de la imagen anterior, consulte el Apéndice B: Imagen 6.

7. Hacer clic en el enlace "**Ver más detalles**" abajo del Informe de equidad, lo llevará a una página que brinda información detallada sobre el desempeño de los **grupos de estudiantes** en ese indicador. Aquí están los **Detalles de los grupos de estudiantes** para ELA:



Para obtener la descripción completa de la imagen anterior, consulte el Apéndice B: Imagen 7.

Estos son sólo algunos de los datos que encontrará en el Tablero de 2018.

Apéndice A: ¿Qué más debemos saber sobre las medidas estatales?

Ahora veamos cada una de las medidas estatales con más detalle.

1. Absentismo crónico

La medida Absentismo crónico muestra cuántos estudiantes estuvieron ausentes durante el 10 por ciento o más del total de días de instrucción escolar. Por ejemplo, la mayoría de las escuelas tiene 180 días de instrucción; si un estudiante está ausente 18 o más de esos días, se le considerará como ausente de manera crónica. Los estudiantes que están ausentes de manera crónica pierden importante instrucción académica. La captura de esta información en el Tablero permite a los padres y educadores ver si el absentismo crónico es un problema en una escuela o en todo el distrito.

2. Tasa de suspensión

La medida Tasa de suspensión muestra el porcentaje de estudiantes que fueron suspendidos en cualquier momento durante el año escolar. Los estudiantes que son suspendidos pierden importante instrucción académica. La captura de esta información en el Tablero permite a los padres y educadores ver si hay un problema de suspensiones en una escuela o en todo el distrito y si ciertos grupos de estudiantes son suspendidos más que otros. (Nota: Un estudiante se cuenta una sola vez para esta medida, incluso si tiene varias suspensiones a lo largo del año).

3. Progreso de los aprendices del inglés

El dominio del idioma inglés es un primer paso para que los estudiantes tengan un buen desempeño en otras áreas, como la lectura, la escritura, las matemáticas y las ciencias. La medida Progreso de los aprendices del inglés examina el progreso que los aprendices de inglés (por ejemplo, los estudiantes que no hablan inglés como primer idioma) están logrando hacia el dominio del idioma inglés. Cada primavera, los aprendices del inglés toman los Exámenes del dominio del idioma inglés para California (ELPAC, por sus siglas en inglés), que mide qué tan bien saben y entienden el inglés. Los resultados de los ELPAC se usan para calcular la medida Progreso de los aprendices del inglés. Debido a que los ELPAC son un nuevo examen, los resultados se comunicarán en el Tablero de 2018, pero no habrá ningún nivel de desempeño (o color) disponible.

4. Tasa de graduación

La medida Tasa de graduación es aplicable sólo a las escuelas y distritos que atienden a estudiantes de preparatoria. Para las escuelas tradicionales, esta medida se basa en el número de estudiantes que se gradúan con un diploma de la escuela preparatoria regular en cuatro años. (Nota: Los estudiantes que obtienen un Certificado de conclusión de la educación especial (conocido en inglés como Special Education Certificate of Completion) o un certificado de equivalencia general no son contados como graduados). Independientemente de si los estudiantes asisten a la universidad, un diploma de la escuela preparatoria es el requisito mínimo para la mayoría de los puestos a nivel de nuevo ingreso en la economía actual. Este representa el dominio de las habilidades fundamentales en matemáticas, lectura y escritura.

5. Preparación para la universidad/carrera profesional

La medida Universidad/carrera profesional muestra cuántos estudiantes se gradúan de la escuela preparatoria mejor preparados para la universidad o una carrera profesional. Utiliza muchas medidas diferentes de la preparación para la universidad y la carrera profesional, examinando los cursos que los estudiantes tomaron en la escuela preparatoria o los exámenes que aprobaron. Para obtener más información sobre cómo las escuelas ayudan a los estudiantes a prepararse, consulte el folleto informativo de la medida Universidad/carrera profesional en https://www.cde.ca.gov/ta/ac/cm/documents/ccidashboardflyer.pdf.

6. Académica: Artes del lenguaje inglés/lectoescritura (ELA) y matemáticas

La medida Académica muestra qué tan bien los estudiantes están cumpliendo con los estándares de ELA y matemáticas correspondientes a su grado. Se basa en el desempeño de los estudiantes en los Exámenes sumativos Smarter Balanced, que los estudiantes en los grados del tercero al octavo y el undécimo presentan cada primavera. Nota: Si menos del 95 por ciento de los estudiantes no presenta los exámenes, se ajusta el desempeño en esta medida.

Apéndice B: Texto descriptivo para las imágenes en este documento

Imagen 1:

Una imagen de los cinco indicadores analógicos que se utilizan en el Tablero de Información Escolar de California. Cada indicador es un dial semicircular que tiene cinco segmentos. Cada segmento representa un nivel de desempeño diferente. De izquierda a derecha, los colores son: Rojo, Naranja, Amarillo, Verde Y Azul. El Rojo representa el nivel de desempeño más bajo mientras que el Azul representa el nivel de desempeño más alto. Una aguja indica el nivel de desempeño para la medida.

Imagen 2:

Una captura de pantalla de la nueva página de inicio del Tablero. La parte superior de la página tiene una imagen de un indicador de desempeño con la flecha apuntando al color verde. De izquierda a derecha, las pestañas son las siguientes:

- Página de inicio
- Acerca de
- Resumen estatal
- Buscar
- Más información
- En español

Explore más información acerca de su escuela o distrito local con una pestaña de búsqueda.

Imagen 3:

Una captura de pantalla de la página de destino del Tablero de 2018 que muestra las siguientes pestañas de izquierda a derecha:

- Absentismo crónico – Color de desempeño Rojo
- Tasa de suspensión – Color de desempeño Verde
- Progreso de los aprendices del inglés – Sin color de desempeño
- Tasas de graduación – Color de desempeño Amarillo
- Universidad/carrera profesional – Color de desempeño Amarillo
- Artes del lenguaje inglés – Color de desempeño Amarillo
- Matemáticas – Color de desempeño Amarillo
- Materiales didácticos básicos para maestros, instalaciones: –Cumplieron con el indicador
- Implementación de los estándares académicos – Cumplió con el indicador
- Participación de los padres – Cumplió con el indicador
- Encuesta del clima escolar– Cumplió con el indicador

Imagen 4:

Una captura de pantalla de cómo debe verse la sección de población de estudiantes matriculados en el Tablero. Esta proporciona el total de la matrícula seguido de los porcentajes de la matrícula que representan los estudiantes de escasos recursos socioeconómicos, los aprendices del inglés y los jóvenes bajo cuidado adoptivo temporal. En este ejemplo, las pestañas se leen de izquierda a derecha en el siguiente orden:

- Matrícula 52,592
- Estudiantes de escasos recursos socioeconómicos 55.2%.
- Aprendices del inglés 29.3%
- Jóvenes bajo cuidado adoptivo temporal 0.4%

Imagen 5:

Una captura de pantalla del Tablero de 2018 para una escuela de ejemplo en la sección "Desempeño académico". Bajo esta sección están los detalles de cómo se desempeñó esta escuela en las medidas de ELA, matemáticas, el indicador del progreso de los aprendices del inglés y el indicador de la preparación para la universidad y la carrera profesional. Los detalles incluyen el color de desempeño, el desempeño del año en curso, la diferencia de desempeño respecto al año anterior y el número de grupos de estudiantes en cada color. Puede explorar más a fondo seleccionando "Ver más detalles". Para nuestra escuela de ejemplo, aquí están sus estadísticas de desempeño para cada uno de los indicadores:

Artes del lenguaje inglés – Obtuvo una calificación de desempeño Amarilla al obtener 5.9 puntos por encima del Estándar para el año en curso, lo que representa un aumento de 0.3 puntos respecto al año anterior. Hay 2 grupos de estudiantes que recibieron una calificación de desempeño Roja, 7 grupos de estudiantes que recibieron una calificación de desempeño Naranja, 3 grupos de estudiantes que recibieron una calificación de desempeño Verde y 1 grupo de estudiantes que recibió una calificación de desempeño Azul.

Matemáticas – Obtuvo una calificación de desempeño Amarilla al obtener 9.1 puntos por encima del Estándar para el año en curso, lo que representa un aumento de 2.1 puntos con respecto al año anterior. Hay 3 grupos de estudiantes que recibieron una calificación de desempeño Roja, 6 grupos de estudiantes que recibieron una calificación de desempeño Naranja, 1 grupo de estudiantes que recibió una calificación de desempeño Verde y 2 grupos de estudiantes que recibieron una calificación de desempeño Azul.

Progreso de los aprendices del inglés – 30.6% estudiantes de esta escuela tuvieron una calificación de Bien desarrollado en el ELPAC. 31.5% con calificación de Desarrollado moderadamente. 9.7% con calificación de Desarrollado ligeramente. 18.2% con calificación de Etapa inicial.

Universidad/carrera profesional – Recibió una calificación de desempeño Amarilla al tener un 55.2 por ciento de sus estudiantes en el nivel Preparado. Esto es una disminución de 2.4% respecto al año anterior. Hay 1 grupo de estudiantes que recibió una calificación de desempeño Roja, 7 grupos de estudiantes que recibieron una calificación de desempeño Naranja, 2 grupos de estudiantes que recibieron una calificación de desempeño Verde y 0 grupos de estudiantes que recibieron una calificación de desempeño Azul.

Imagen 6:

Captura de pantalla del Reporte de equidad que muestra el número de Grupos de estudiantes en cada color de desempeño En este ejemplo, hay 2 grupos de estudiantes que recibieron una calificación de desempeño Roja, 7 grupos de estudiantes que recibieron una calificación de desempeño Naranja, 0 grupos de estudiantes que recibieron una calificación de desempeño Verde y 1 grupo de estudiantes que recibió una calificación de desempeño Azul.

Imagen 7:

Una captura de pantalla de los Detalles de los grupos de estudiantes en un indicador no especificado en el Tablero de 2018. Esta sección indica cómo se ubicó cada uno los grupos de estudiantes en términos de su color de desempeño. En este ejemplo, estos son los grupos de estudiantes indicados en cada uno de los colores de desempeño:

- Bajo el nivel de desempeño Rojo
  - — Afroamericanos
  - — Jóvenes bajo cuidado adoptivo temporal
- Flecha del indicador que apunta al color Naranja
  - — Indígenas americanos
  - — Aprendices del inglés
  - — Hispanos
  - — Sin hogar
  - — Nativos de una isla del Pacífico
  - — Estudiantes de escasos recursos socioeconómicos
  - — Estudiantes con discapacidades
- Flecha del indicador que apunta al color Amarillo
  - — No estudiantes (Grupos)
- Flecha del indicador del medidor que apunta al color Verde
  - — Asiáticos
  - — Filipinos
  - — Dos o más razas
- Flecha del indicador del medidor que apunta al color Verde
  - — Blancos
- Indicador sin color de desempeño

— No estudiantes (Grupos)

Manzanita Charter School
2925 Technology Court, Richmond CA 94806 ~ (510) 222-3500 ~ Fax (510) 222-3555
www.manzy.org

Procedure for Managing EL Data (CELDT) & Reclassifying Students to (R-FEP) Status

CELDT testing data is maintained in the school database and is most easily viewed through the *Main – CELDT Data* layout. Each year's scores are entered into a new box on the right side of the layout.

Manzanita's REFP procedures are modeled on those of the WCCUSD RAP Center. To be reclassified a student needs to meet the following benchmarks:

1. An Overall CELDT Score of 4 or 5 (Early-Advanced or Advanced)
2. Scores of 3 (Intermediate) or higher in all CELDT subcategories
3. Level 3 (Standard Met) or higher on the ELA portion of the CAASPP*
4. Grades at C or above
5. All 4's or 5's on the Student Oral Language Observation Matrix (SOLOM)
6. Approval of their English teacher
7. Approval of their parent or guardian

* Past Requirements for #3:
2014-2015: At or above Mean Level RIT on the MAP ELA Assessment
Pre-2014: ELA-CST score of 325 or higher

Steps to Reclassify a Student to R-FEP Status

A data worksheet such as the one below (Testing > CELDT > RFEP > RFEP Check 1516.xls) is created as soon as the fall CELDT test results have arrived (usually December or January). The only students entered into the worksheet are those who have accomplished the first two benchmarks: a score of 4 or 5 overall AND a score of 3 or higher in each of the subcategories on the CELDT.

CAASPP ELA scores are entered into the worksheet from the previous spring testing and the students who have met the second benchmark (Level 3 on the ELA portion of the CAASPP) are highlighted.

Report cards of these students are then reviewed to determine if their grades meet the third benchmark if C or above in all academic classes.

The English teacher of any student passing the grades benchmark is then asked to complete the Student Oral Language Observation Matrix (SOLOM) form.

Letters are sent home to the families of any students receiving all 4's or 5's on the SOLOM (Manzy Database > RFEP Parent Agreement Letter). These letters notify the family that the school feels the student ready for reclassification, and asks the parents for their agreement or disagreement with this assessment.

Once these letters have been signed and returned, a R-FEP form is completed for each student being reclassified (Manzy database > RFEP Reclassification Form). The family is sent a second letter (Manzy database > RFEP Final Reclassify Ltr) confirming the reclassification of the student. Copies of the completed R-FEP forms, signed by the teacher and the Executive Director, are mailed to the family's address and filed in the CUM folder. A data upload is performed to CALPADS with the status change.

CAASPP ELA scores from the following spring testing are added to this spreadsheet to see if any additional students qualify for reclassification during the year.

Returning Students

	Grade 15-16	15-16 CELDT 4/5 Overall	15-16 CELDT 3+ all subcats	2015 CAASPP ELA		Grades C or Above	SOLOM Teacher Observation	A	B	C	D	E	R-FEP?
Student 1	7	yes	yes	1-Not	no								no
Student 2	7	yes	yes	1-Not	no								no
Student 3	7	yes	yes	1-Not	no								no
Student 4	7	yes	yes	2-Nearly	no								no
Student 5	7	yes	yes	3-Met	√	yes		5	5	4	5	4	yes
Student 6	7	yes	yes	1-Not	no								no
Student 7	7	yes	yes	3-Met	√	yes		4	5	4	4	4	yes

R-FEP Success Checklist

- ☐ Enter CELDT data into database
- ☐ Create R-FEP data worksheet
- ☐ Enter students who have passed CELDT benchmarks
- ☐ Review CAASPP scores
- ☐ Review report card data
- ☐ Distribute SOLOM forms
- ☐ Follow up on SOLOM forms
- ☐ Receive SOLOM forms back
- ☐ Enter above data on R-FEP forms
- ☐ Send parent agreement letter
- ☐ Follow up on parent agreement letter
- ☐ Add parent response to R-FEP forms
- ☐ Administrator & teacher sign the forms
- ☐ File original forms in students' CUMs
- ☐ Reclassify students in the database
- ☐ Send final reclassification letters to families
- ☐ Update CALPADS SINF for R-FEP's

Reclassification Procedures (EL to RFEP)
Updated 2018

Manzanita's REFP procedures are modeled on those of the WCCUSD RAP Center. To be reclassified a student needs:

1) An Overall CELDT Score of 4 or 5 (or ELPAC score of ~~???~~ 4)
2) Scores of 3 or higher in all CELDT subcategories (or ELPAC score of ???)
3) At or above the R180 cutoffs per grade level (see below)*
4) Grades at C or above (D okay in Math if CELDT overall is 5)
5) All 4's or 5's on the Student Oral Language Observation Matrix (SOLOM)
6) Approval of English teacher
7) Approval of parent or guardian

R180 Cutoffs
6th ≥ 822
7th ≥ 869
8th ≥ 900

* Prior to 1/1/18, scores from the Interim CAASPP ELA were used for this metric.

Layouts are in the school database:
RFEP Parent Agreement Letter
RFEP Reclassification Form
RFEP Final Reclassify Ltr
RFEP Certificate

Checklist

- Enter CELDT results into database
- In RFEP Tracking layout in database, mark students who have passed CELDT benchmarks
- Review Reading Inventory scores and mark those who have passed benchmarks
- Review report card data and mark those who have passed benchmarks
- Update SOLOM form on drive and ask teachers to give scores and mark those who have passed benchmarks
- Send parent agreement letter
- Administrator & teacher sign the forms
- File original forms in the CUM
- Reclassify students in the database
- Send final reclassification letters to families
- Update CALPADS SELA for R-FEP's
- Print R-FEP sticker for the CUM file (important!)
- Print RFEP Certificate for student

.OM Teacher Observation

dent Oral Language Observation Matrix

be used as part of the reclassification process for English Language Learners)

	1	2	3	4	5	Score(s)
:omprehension	Can not understand even simple conversation	Has great difficulty following everyday social conversation, even when words are spoken slowly and repeated frequently	Understands most of what is said at slower-than-normal speed with some repetitions	Understands nearly everything at normal speed, although occasional repetition may be necessary	Understands everyday conversation and normal classroom discussion without difficulty	
'luency	Speech so halting and fragmentary that conversation is virtually	Usually hesitant; often forced into silence because of language limitations	Everyday conversation and classroom discussion frequently disrupted by student's search for correct manner of expression	Everyday conversation and classroom discussion generally fluent, with occasional lapses while student searches for the correct manner of expression	Everyday conversation and classroom discussion fluent and effortless; approximately those of a native speaker	
/ocabulary	Vocabulary limitations so extreme that conversation is virtually impossible	Difficult to understand because of misuse of works and very limited vocabulary	Frequent use of wrong words; conversation somewhat limited because of inadequate vocabulary	Occasional use of inappropriate terms and/or rephrasing of ideas because of limited vocabulary	Vocabulary and idioms approximately those of a native speaker	
'ronunciation	Pronunciation problems so severe that speech is virtually unintelligible	Difficult to understand because of pronunciation problems; must frequently repeat in order to be understood	Concentration required of listener; occasional misunderstandings caused by pronunciation problems	Always intelligible, although listener conscious of a definite accent and occasional inappropriate intonation pattern	Pronunciation and intonation approximately those of a native speaker	
;rammar	Errors in grammar and word order so severe that speech is virtually unintelligible	Difficult to understand because of errors in grammar and word order; must often rephrase or restrict speech to basic patterns	Frequent errors in grammar and word order; meaning occasionally obscured	Occasional errors in grammar or word order; meaning not obscured	Grammar and word order approximately those of a native speaker	

dec. 2002 per WCCUSD. Jan. 2005

*Score:*____________