METAMETRICS RESEARCH BRIEF

A Quantitative Task Continuum for K-12 Mathematics

Eleanor E. Sanford-Moore, Ph.D., Senior Vice President of Research and Development; Gary L. Williamson, Ph.D., Senior Research Scientist; Lisa Bickel, Director of C&D Services; Heather Koons, Ph.D., Director of Research Services; Robert F. Baker, Director of Analytical Services; Ruth Price, Quantile Curriculum Specialist

November 2014

OBJECTIVE

The two main objectives of this research were:

- to quantify the difficulty of mathematics lessons drawn from mathematics textbooks commonly used in the United States; and,
- to document the mathematical complexity of textbook lessons within and across grades.

Key Hypotheses:

1. The median mathematical complexity of textbook lessons consistently increases with grade;
2. Within grades, lessons vary in their mathematical complexity;
3. Both textbooks aligned with the Common Core State Standards for Mathematics (CCSSM) and those not aligned with CCSSM can be reflected in a common mathematics lesson continuum.

METHODS

Participants*:* The units of analysis were lessons extracted from selected textbooks used in grades K-12 in the United States. Core mathematics programs and publishers were identified based on market analysis (Education Market Research, 2014; pp. 381-416) and state textbook adoptions. Textbooks were selected, acquired, and assembled into a database for this study. Admissible texts consisted of only mathematics textbooks (i.e., not auxiliary books or supplemental instructional materials) used in the public schools in the United States, excluding texts designed specifically for honors programs and intervention programs. K-12 mathematics textbooks published before 2005 were excluded from the study to help ensure that the study results would represent the difficulty of mathematics lessons contained in textbooks published within the most recent decade. Textbooks published from 2005-2010 were regarded as "non-Common Core textbooks," whereas the majority of the textbooks with publication dates from 2011 to the present were categorized as "Common Core textbooks" (if confirmed by the publisher's claims and the date of publication provided on the publisher's website). High school texts were further categorized as appropriate either for a traditional mathematics sequence (i.e., Algebra 1, Geometry, Algebra 2) or for an integrated mathematics program (i.e., Integrated Math 1, 2, 3). Lessons were associated with year of schooling (i.e., grade) as identified in the textbooks' abstracts.

Procedure:

The Quantile® Framework for Mathematics was used to calibrate each lesson and assign a Quantile® measure to represent the difficulty of the mathematical skills and concepts associated with the lesson (MetaMetrics, 2009). Lesson measures were then analyzed by grade to determine a) how difficult lessons are in each grade, b) how much variability in difficulty exists within each grade, and c) how much lesson difficulty increases from grade to grade. The lesson continuum was further analyzed to determine whether there are differences in lesson difficulty depending on curricular design and intended use: i.e., whether the textbook was designed before or after the introduction of the CCSSM (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010) and also whether the textbook was designed for use with a traditional or integrated mathematics program.

Measures:

The Quantile Framework for Mathematics uses a Rasch measurement model to provide developmentally-scaled measurements for both students and mathematical tasks. More specifically, the Quantile scale measures a student's mathematics ability as manifested through his or her mathematics achievement; and, it also measures the difficulty of mathematics skills and concepts. It is important to note that students, skills and concepts are all measured using the same common scale, making it possible to manage mathematics instruction that is targeted to students' needs. MetaMetrics (2009, 2014) and Scholastic (2014) provide additional information about the Quantile Framework and the Quantile scale. Williamson and Simpson (2013) reported empirical evidence of the validity of the Quantile Framework for developmental growth analyses. Williamson (2013) further illustrated the usefulness of the Quantile scale for exploring student growth.

ANALYSES

Each grade-level collection of lesson measures was analyzed by curriculum category (i.e., Common Core, non-Common Core), instructional sequence at the secondary level (i.e., traditional sequence or integrated math sequence) and as one comprehensive (pooled) collection. Selected percentiles (5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th}) of each group of lessons were calculated. These were used to construct modified box-and-whisker plots for the grade-level distributions of mathematical complexity. The box-and-whisker plots were displayed graphically in grade-level order and visually inspected for monotonicity within each category (i.e., curriculum focus, instructional sequence and overall). Analysis of variance (ANOVA) was used to detect any significant differences between lessons from the Common Core textbooks and the lessons from the non-Common Core textbooks. Similarly, ANOVA was used to detect any difference between the lessons associated with the traditional instructional sequence versus the integrated mathematics sequence.

RESULTS & DISCUSSION

Lessons were extracted from 370 mathematics textbooks with publication dates ranging from 2005 to 2015. Of these, 65 texts were characterized as Common Core textbooks and 305 were classified as non-Common Core textbooks. There were 28,027 lessons altogether; each lesson was assigned a Quantile measure to represent its mathematical difficulty (i.e., the task difficulty of the skills and concepts associated with the lesson). The lesson measures were then subjected to two omnibus tests to determine whether there were overall differences in the average difficulty of lessons by either curriculum category or instructional sequence.

There was a small, statistically significant effect for instructional sequence, $F(1, 4916) = 3.88$, $p = 0.0488$, df = 1. Lessons from traditional texts were on average just 24.3Q higher in difficulty than lessons associated with integrated math texts. The $R^2 = 0.0007879$ for the ANOVA model confirms that the overall relationship between lesson difficulty and instructional sequence is very weak and statistical significance is mainly due to the large sample size.

Similarly there was a small, statistically significant effect for curricular focus, $F(1, 25955) = 22.32$, $p < 0.0001$, with lessons from Common Core textbooks being 35.2Q more difficult than lessons from non-Common Core textbooks. However, the $R^2 = 0.000859$ again indicates a weak overall relationship and statistical significance is again due primarily to large sample size. An inspection of the two continua in Figure 1 visually confirms little difference in difficulty of the lessons from Common Core and non-Common core textbooks. Furthermore, the small difference that exists appears to be isolated to just four grades (6-9). It is also notable that within each grade there is considerable overlap between the distributions of lesson measures from Common Core textbooks and the lesson measures from non-Common Core textbooks. These results provisionally lead us to conclude that a single continuum for mathematics lesson difficulty across grades is sufficient, based on pooling all lesson-measures together without regard to the curricular focus or instructional sequence attributed to the textbooks.

Figure 2 displays a single mathematics continuum of lesson difficulty across grades. Because of their relatively small *n*-counts, pre-algebra and pre-calculus lesson measures were dropped from Figure 2 to produce a more parsimonious summary of the mathematics lesson continuum. The lesson measures from integrated math textbooks were pooled with the lesson measures from traditional mathematics textbooks in grades 9, 10 and 11 for the display. Figure 2 is recommended as a convenient visual summary of the difficulty of mathematics lessons in textbooks that are commonly used in the United states. Table 1 provides the sample sizes for the lesson distributions and summarizes the percentiles of the distributions by grade as displayed in Figure 2.

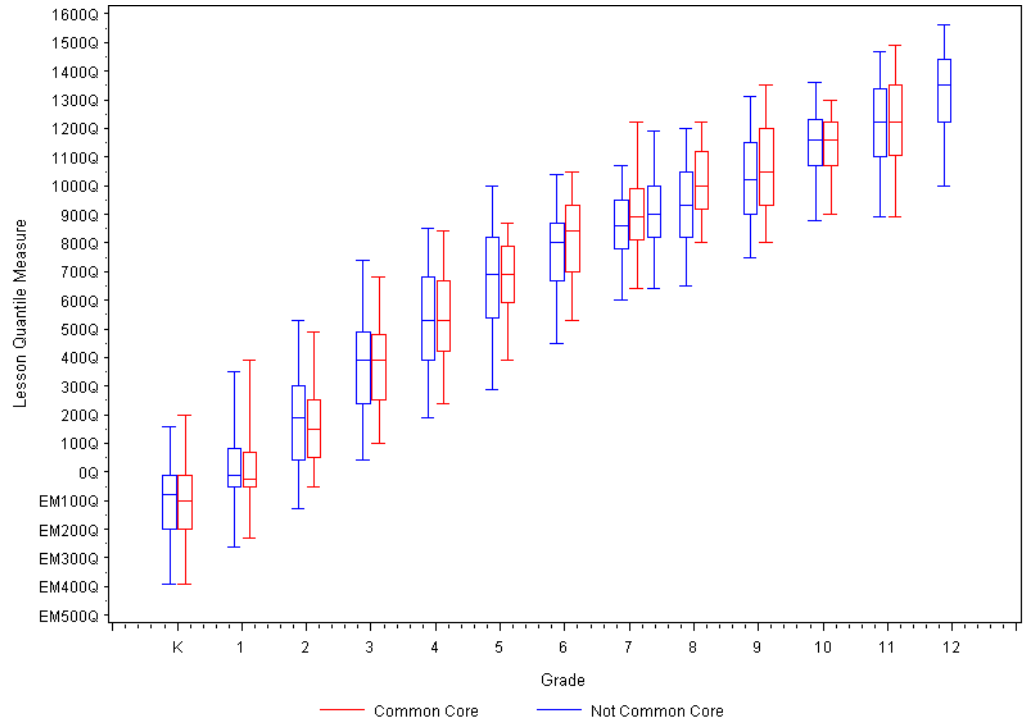
The analyses confirm our initial research hypotheses. First, the median mathematical complexity of textbook lessons consistently increases with grade. Generally, the largest grade-to-grade increases occur between Grade 1 and Grade 6. In those transitions, the grade-to-grade increases in mathematical complexity range from 120Q to 200Q with the largest increases being from Grade 1 to Grade 2 and from Grade 2 to Grade 3. Grade-to-grade increases during the middle and high school years are generally less than or equal to 70Q, except for the 140Q increase between Grade 9 (Algebra I/Math 1) and Grade 10 (Geometry/Math 2). The boundaries of the interquartile range also increase consistently from grade to grade, providing a visually consistent picture of increasing mathematical task complexity from earlier to later grades.

The second initial hypothesis was that lessons vary in their mathematical complexity within each grade. This is confirmed by the interquartile range (IQR) for each grade, which can be calculated from the percentiles in Table 1. (The IQR for each grade is calculated by subtracting the 25^{th} percentile from the 75^{th} percentile.) IQRs range from 130Q (in Grade 1) to 290Q (in Grade 4). The whiskers for each box in the graph display the variation between the 5^{th} and 95^{th} percentiles, further supporting the contention that there is great variability in lesson complexity within each grade.

Both non-Common Core textbooks and Common Core textbooks are reflected in a common mathematics lesson continuum displayed in Figure 2, confirming the third initial hypothesis. Even though textbooks aligned with CCSSM and those not aligned with CCSSM are reflected in a common mathematics lesson continuum, one distinction is noteworthy. The first of three key shifts identified in the CCSSM is greater focus on fewer topics in each grade level: “Rather than racing to cover many topics in a mile-wide, inch-deep curriculum, the standards ask math teachers to significantly narrow and deepen the way time and energy are spent in the classroom” (Common Core State Standards Initiative, 2014, “Greater focus on fewer topics,” para. 1). This can be seen in Figure 1 by comparing the lesson measures from the non-Common Core textbooks to the lesson measures from the Common Core textbooks. In Grades 2, 3, 4, 5, 6, 8, and 10, the range of the lesson measures within a grade is less for the Common Core textbooks than for the non-Common Core textbooks, indicating a narrower focus of topics.

The continuum displayed in Figure 2 is the primary result for the research summarized in this brief. The results provide educators a clear picture of the mathematical task continuum as it is currently manifested in lessons derived from commonly used textbooks in the United States. Although, statistical probability samples were not possible, the sample sizes are substantial. The textbooks and lessons that were selected for analysis may in fact represent a large proportion of the textbooks used in K-12 mathematics education because there are relatively few textbook publishers and states have limited options for textbook adoption.

Our understanding of the mathematical demands of instructional materials used in the public schools and in the postsecondary world is evolving. Schools have the option of supplementing regular textbooks and both texts and auxiliary mathematics instructional material can change over time. Although this research was limited to mathematics lessons from Grades K-12, the results provide a substantial baseline that can be built upon and expanded by future studies.



Note: Whiskers are 5th and 95th percentiles.

Figure 1. **Grade-level distributions of mathematics lesson complexity by curricular focus (Common Core vs. Non-Common Core). Pre-algebra lessons are displayed in the figure at grade 7.5; pre-calculus lessons are displayed at grade 12. There were no Common Core textbooks for pre-algebra or pre-calculus.**

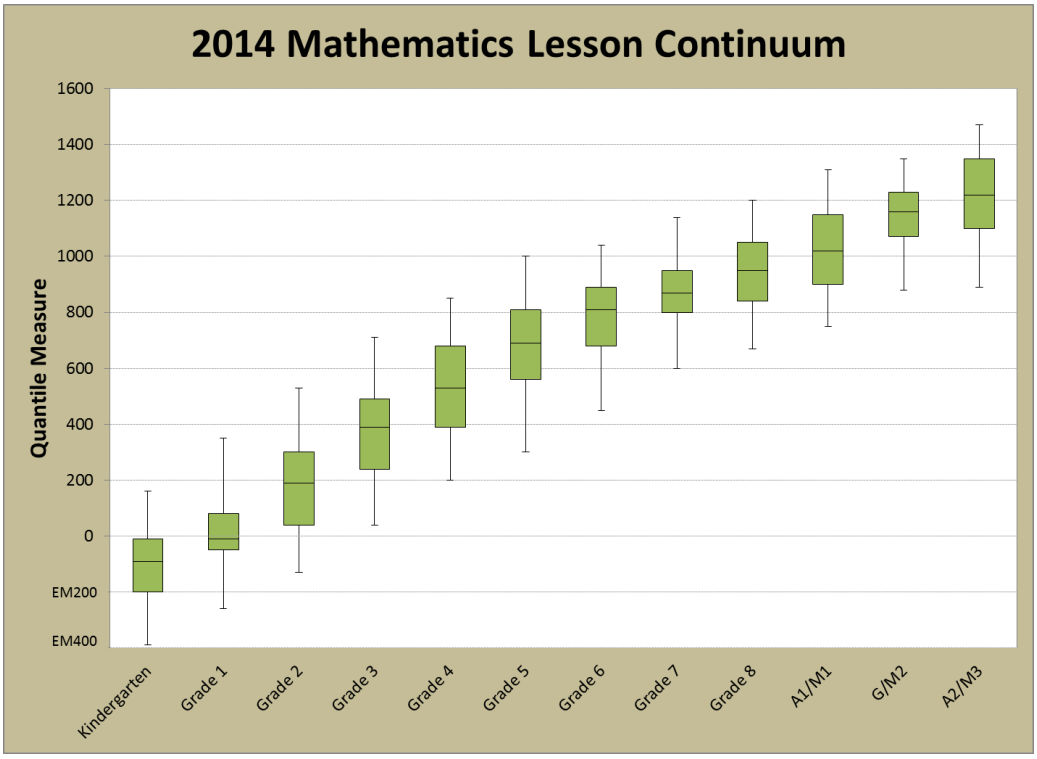


Figure 2. **The task difficulty of mathematics lessons (n=27,630) by grade. Box and whiskers represent the 5th, 25th, 50th, 75th and 95th percentiles of the grade-level distributions of lesson difficulty measures.**

Table 1

Number of mathematics lessons by grade with percentiles of the grade-level distributions of lesson-complexity measures.

Grade	N	5th	25th	50th	75th	95th
K	2,012	EM390Q	EM200Q	EM90Q	EM10Q	160Q
1	2,287	EM260Q	EM50Q	EM10Q	80Q	350Q
2	2,457	EM130Q	40Q	190Q	300Q	530Q
3	3,046	40Q	240Q	390Q	490Q	710Q
4	2,872	200Q	390Q	530Q	680Q	850Q
5	2,765	300Q	560Q	690Q	810Q	1000Q
6	3,266	450Q	680Q	810Q	890Q	1040Q
7	1,630	600Q	800Q	870Q	950Q	1140Q
8	1,926	670Q	840Q	950Q	1050Q	1200Q
9	2,109	750Q	900Q	1020Q	1150Q	1310Q
10	1,405	880Q	1070Q	1160Q	1230Q	1350Q
11	1,855	890Q	1100Q	1220Q	1350Q	1470Q
Total	27,630					

REFERENCES

Common Core State Standards Initiative. (2014). *Key shifts in mathematics*. Retrieved from http://www.corestandards.org/other-resources/key-shifts-in-mathematics/

Education Market Research. (2014). *The complete K-12 report® : Market facts & segment analyses, 2014.* Rockaway Park, NY: Author.

MetaMetrics. (2009). The Quantile Framework® for Mathematics: Linking assessment with mathematics instruction. Durham, NC: Author.

MetaMetrics. (2014). Linking the NC READY EOG Math/EOC Algebra I/Integrated I with the Quantile® Framework: A Study to link the North Carolina READY EOG Math/EOC Algebra I/Integrated I with The Quantile® Framework for Mathematics. Durham, NC: Author.

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

Scholastic Inc. (2014). *SMI College & Career*. New York: Author.

Williamson, G. L. (2013, April). *Mathematics growth conditioned on Algebra I course-taking behavior*. Paper presented at the annual meeting of the American Educational Research Association (AERA) to the Special Interest Group (SIG) on Longitudinal Studies, San Francisco, CA.

Williamson, G. L., & Simpson, M. A. (2013, April). *Statewide developmental growth curves in reading and mathematics: A practical strategy to lengthen empirical trajectories.* Distinguished Paper Presentation at the annual meeting of the American Educational Research Association, San Francisco, CA.

MetaMetrics®